MODELLING THE PRODUCTION POTENTIAL OF LAND FOR SUGARCANE IN THE KWA ZULU-NATAL MIDLANDS SUGARCANE BELT

by

EDGAR ALFRED BRÜGGEMANN
B.Sc. Agric. *summa cum laude* (Natal)

Submitted in partial fulfilment
of the requirements for
the degree of

MASTER OF SCIENCE IN AGRICULTURE
(AGRICULTURAL PRODUCTION)

in the
School of Agricultural Sciences and Agribusiness
Faculty of Science and Agriculture
University of Natal
PIETERMARITZBURG

February 2000

FINAL COPY

ABSTRACT

Commercial sugarcane records for 19 seasons from 146 fields were obtained from selected estates in the KwaZulu-Natal Midlands. The estates, located at Kranskop, Umvoti and Richmond, are representative of the higher-potential rainfed sugarcane production region of the Midlands sugarcane belt. Extensive editing and cleaning of the agronomic records was required.

Regression models were developed to determine which parameters of the field records were consistently associated with sugarcane yield (TCH) and could be used for yield predictions. Depending on the predictor variables selected, the best models based on 535 crop cycles accounted for 55% and 43% of the observed yield variation respectively. Linear regression was an appropriate analytical technique since the assumptions of normality and homoscedasticity were upheld and multicollinearity was not a problem in the models. The models were validated using an independent data set of 47 observations and satisfactory performances were confirmed. The 95% confidence limits of yield predictions for the population mean lie within 10% of long-term mean yields. These predictions could be useful for estate resource allocation and harvest planning.

Key physical field attributes associated with sugarcane yield were locality, aspect, altitude, soil type and effective rooting depth. Season and rainfall were important climatic variables. Of the factors influenced by management, sugarcane variety, plant / ratoon status, crop cycle, N and K nutrition and the topsoil Ca:Mg ratio were important yield predictors, depending on the equation used. The relative importance of individual predictors varies with the specific combination of resources for a particular observation.

The models were linked to a geographic information system to demonstrate an application of the models for yield prediction in response to spatial variables. These predictions showed that the models could be used at a general scale within estates to identify areas of differing production potential. Reliable yield predictions could not be made for individual fields and within-field resource variations could not be adequately accounted for.

PREFACE

This thesis documents research conducted at the University of Natal from February 1995 to April 1999 under the supervision of Mr JR Klug, assisted by Professor PL Greenfield and Mr HM Dicks.

I declare that the results contained in this thesis are from my own original work except where acknowledged. I also declare that these results have not otherwise been submitted in any form for any degree or diploma to any university.

E. A. Brüggemann
....
EA Brüggemann

February 2000

ACKNOWLEDGEMENTS

The assistance of the following persons and organisations is gratefully acknowledged.

Mr JR Klug for guidance, constructive criticism and assistance, as supervisor of this study.

Professor PL Greenfield for advice at critical times, as co-supervisor of this study.

Mr HM Dicks for guidance with statistical analyses, as co-supervisor of this study.

Mondi Forests for funding research expenses, providing the GIS data and allowing access to their field records. Mondi staff, particularly Messrs P Gardiner, R Leibnitz, B Alcock, and E Suleman and Mesdames W Morris, S Bernhard, R Phillips and A Moorehead are thanked for their assistance with these aspects of the project. The patient assistance of the estate managers, Messrs A Hoskin, J van der Ryst, D Chard and J Shange for the confirmation of the field records is gratefully acknowledged.

Staff of the School of Bioresources Engineering and Environmental Hydrology in the Faculty of Engineering, University of Natal, particularly Mr M Horan, for the GIS programming.

Spencer Holley Agronomic Services, particularly Mr G Moore and the late Mr J Spencer for the provision and explanation of the Mondi sugarcane field records.

Staff of the South African Sugar Experiment Station, especially Mr JH Meyer for explaining aspects of the Fertilizer Advisory Service and Drs NG Inman-Bamber and A Singels for explaining aspects of the CANEGRO model. The sugarcane extension officers, Mr DB Hellmann (Midlands South) and Mr QV Mann (Midlands North and Central) for explaining general sugarcane production practises in the Midlands.

The Grain Crops Institute of the Agricultural Research Council, particularly Dr MPW Farina for explaining aspects of the soil analyses performed for Spencer Holley Agronomic Services.

The Sugar Industry Trust Fund for Education for project funding in the form of a Postgraduate Scholarship.

The Foundation for Research and Development for project funding in the form of a one-year Master's Bursary.

My wife, Ilze, for support, encouragement and showing interest in my work; also for assisting with some of the data processing and proof-reading the various drafts of this document.

Any others whose names I might inadvertently have omitted.

Finally I wish to thank the good Lord, for revealing some of the knowledge of His creation and for solving numerous problems with this study.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

---

† This appendix comprises an electronic copy of the data analysed in this study. No specific reference is made to Appendix 13 in the text.

LIST OF ABBREVIATIONS

ACRU	Agricultural Catchment Research Unit
$(A\ell^{3+} + H^{+})$	Exchangeable acidity
AMBIC	Ammonium bicarbonate
ASI	Aluminium saturation index
a.s.l.	Above sea level
B	Bottomland
BRG	Bioresource group
BRU	Bioresource unit
C	Crest
CM	Crest followed by a midslope in the terrain sequence
CAD	Computer aided design
DAP	Di-ammonium Phosphate
DSSAT	Decision Support System for Agrotechnology Transfer
DTM	Digital terrain model
E_{ann}	Annual total evaporation
E_{act}	Actual total evaporation
E_{max}	Maximum total evaporation
EAI	Exchangeable aluminium index
ERD	Effective rooting depth
F	Footslope
FB	Footslope followed by a bottomland in the terrain sequence
FM	Footslope followed by a midslope in the terrain sequence
FAO	Food and Agriculture Organisation of the United Nations
FAS	Fertilizer Advisory Service – South African Sugar Association Experiment Station
FERTREC	Fertilizer Recommendation Service – Cedara, Department of Agriculture
FRS	Field Record System – South African Sugar Association Experiment Station

FSD	Forest industry soils database
GCI	Grain Crops Research Institute
GIS	Geographic information system
GPS	Global positioning system
IBSNAT	International Benchmark Sites Network for Agrotechnology Transfer
ICFR	Institute for Commercial Forestry Research
k_y	Yield response factor
LAI	Leaf area index
M	Midslope
MAP	Mono-ammonium Phosphate
MB	Midslope followed by a bottomland in the terrain sequence
MF	Midslope followed by a footslope in the terrain sequence
NN	Artificial neural network amalysis
NTE	Natal Tanning Extract company
OC	Organic carbon
PAS	Permissible acid saturation
PA.S*1	Plant autumn, harvest summer, one summer's growth
PA.S*2	Plant autumn, harvest summer, two summer's growth
PA.S*3	Plant autumn, harvest summer, three summer's growth
PA.S*4	Plant autumn, harvest summer, four summer's growth
PA.W*1	Plant autumn, harvest winter, one summer's growth
PA.W*2	Plant autumn, harvest winter, two summer's growth
PA.W*3	Plant autumn, harvest winter, three summer's growth
PA.W*4	Plant autumn, harvest winter, four summer's growth
PS.S*1	Plant summer, harvest summer, one summer's growth

PS.S*2	Plant summer, harvest summer, two summer's growth
PS.S*3	Plant summer, harvest summer, three summer's growth
PS.S* 4	Plant summer, harvest summer, four summer's growth
PS.W*1	Plant summer, harvest winter, one summer's growth
PS.W*2	Plant summer, harvest winter, two summer's growth
PS.W*3	Plant summer, harvest winter, three summer's growth
PS.W*4	Plant summer, harvest winter, four summer's growth
PDI	Phosphorus desorption index
PRF	Phosphorus requirement factor
R_a^2	Adjusted coefficient of multiple determination
RS.S*1	Summer ratoon, harvest summer, one summer's growth
RS.S*2	Summer ratoon, harvest summer, two summer's growth
RS.S*3	Summer ratoon, harvest summer, three summer's growth
RS.S*4	Summer ratoon, harvest summer, four summer's growth
RS.W*1	Summer ratoon, harvest winter, one summer's growth
RS.W*2	Summer ratoon, harvest winter, two summer's growth
RS.W*3	Summer ratoon, harvest winter, three summer's growth
RS.W*4	Summer ratoon, harvest winter, four summer's growth
RW.S*1	Winter ratoon, harvest summer, one summer's growth
RW.S*2	Winter ratoon, harvest summer, two summer's growth
RW.S*3	Winter ratoon, harvest summer, three summer's growth
RW.S*4	Winter ratoon, harvest summer, four summer's growth
RW.W*1	Winter ratoon, harvest winter, one summer's growth
RW.W*2	Winter ratoon, harvest winter, two summer's growth
RW.W*3	Winter ratoon, harvest winter, three summer's growth
RW.W*4	Winter ratoon, harvest winter, four summer's growth
Sc	Scarp
SASA	South African Sugar Association
SASEX	South African Sugar Association Experiment Station
SWDF1	Soil water deficit factor 1 – plant water demand exceeds water supply
SWDF2	Soil water deficit factor 2 – incipient water stress

TAW	Total available water
TCH	Tons cane per hectare
TCHM	Tons cane per hectare per month
TSH	Tons sucrose per hectare
TSHM	Tons sucrose per hectare per month
Y_{act}	Actual yield
Y_{max}	Maximum yield

1 INTRODUCTION

The South African sugar industry is one of the pillars of the agricultural sector in KwaZulu-Natal. Within this region sugarcane yields vary considerably, depending mainly on climate, physical land characteristics, soil fertility and management ability. Mondi Forests, a division of Mondi (Ltd) commissioned this study to evaluate rainfed sugarcane production in the KwaZulu-Natal Midlands using their field records and other secondary data. Mondi are one of the sugar industry's larger independent growers, producing about 350 000 tons of sugarcane as an enterprise secondary to their 5 million tons timber per annum.

Sugarcane (*Saccharum officinarum*) is produced along the KwaZulu-Natal coast at altitudes ranging from a few metres to about 1100 m above sea level (a.s.l.). This "sugarcane belt" has been loosely divided into Coastal (North and South Coast) and Midlands production regions. The intention of this division appears to have been to differentiate between the traditional Coastal sugarcane areas of KwaZulu-Natal and the higher altitude areas where commercial sugarcane production began during the early 1950's. Midlands sugarcane is exposed to weaker maritime influences than Coastal sugarcane and is usually grown at altitudes between 750 m and 1100 m a.s.l. (De Haas, 1981). In this region valley systems at altitudes less than 750 m a.s.l. usually experience frost annually and sugarcane is reliably produced only under irrigation, while at altitudes exceeding 1100 m a.s.l. cold stress and insufficient heat units are considered to severely restrict sugarcane production potential. Sugarcane grown in the Kranskop, Umvoti, New Hanover, Pietermaritzburg, Richmond, Camperdown and Eston districts of KwaZulu-Natal is considered to be Midlands sugarcane. These districts fall under the Midlands North-and-Central and Midlands South extension areas (Figure 1, areas 6 and 9).

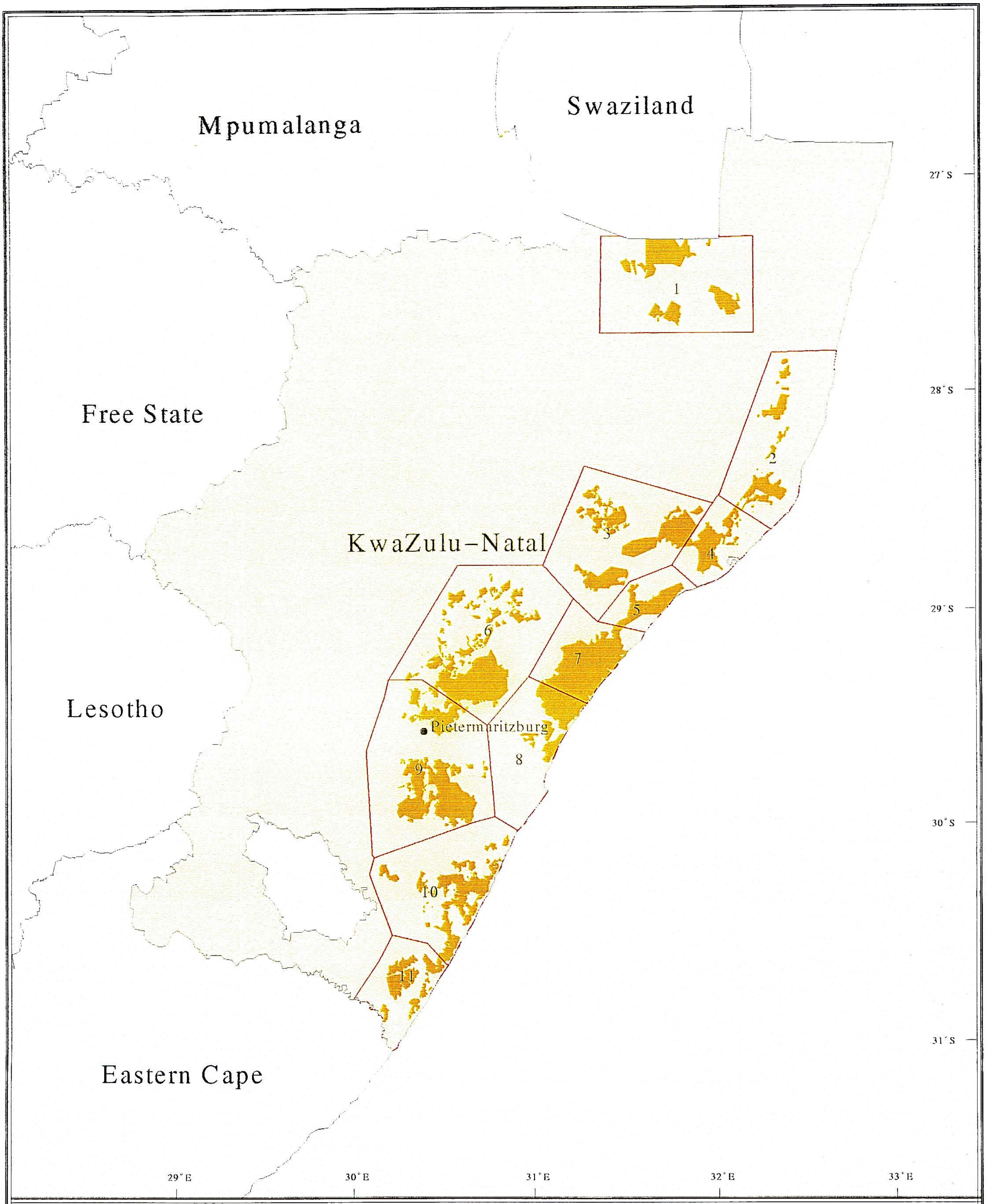


Figure 1 : Locality map of the KwaZulu-Natal Sugarcane mill extension areas.

During winter, low temperatures and very dry conditions severely restrict extension growth of sugarcane in the Midlands. For this reason Midlands sugarcane is harvested at 18 to 24 months of age and, unless annual sugarcane is produced under irrigation, frost-prone parts of estates are generally not planted to sugarcane because the crop stands on the field during at least one winter. Sugarcane is usually planted from September to December but planting may extend into January in some years. Frost prone areas are best planted during February and March to minimize production losses. The harvesting season usually extends from mid-April to mid-November (7 months), but may start as early as mid-March and continue until mid-December (9 months) during years of exceptionally large crops.

Mondi produce sugarcane at Kranskop, Umvoti and Richmond within the Midlands sugarcane region, and the field records from these estates formed the basis of this study.

The objectives of the study were to;

a determine key parameters strongly associated with actual sugarcane yields in field records from the three Midlands districts, and

b derive empirical models suitable for yield prediction and performance evaluation using estate records.

Relevant literature is reviewed in Chapter 2 to establish the scope of sugarcane field record analysis and approaches for analysing field records. The broadly agronomic sugarcane crop models applicable to South Africa are also considered in detail to determine their suitability for yield predictions at various scales and to determine which variables have or have not been used to predict yield. This information will be useful for selecting appropriate approaches to achieve objective (a) above. In Chapter 3 the production data from Mondi's sugarcane field records are described and evaluated. The available agricultural resource information is also described and limitations of the data for modelling applications (objective (b) above) are discussed. The identification and selection of parameters associated with actual yields, model building and validation are discussed in Chapter 4. An application of these models, linked to a geographic information system, is also described and demonstrates, by way of example, how managers can use the models to determine expected yields, thereby illustrating how objective (b) above can be achieved. In Chapter 5 selected aspects of the study are discussed in detail and the implications of the findings of the study are considered. The conclusions of the study and aspects which warrant further research are presented in Chapter 6 and an executive summary of this thesis is provided in Chapter 7.

2 SUGARCANE FIELD RECORD ANALYSIS AND YIELD MODELS

A number of mandatory records have been kept by South African sugarcane farmers as part of the quota and sugarcane payment systems. The lowest level of spatial aggregation to which these records refer is the commercial sugarcane field, and generally include field number, field size, crop class (ratoon status), age at harvest, variety, total tons cane and tons sucrose harvested. Although these records alone may provide useful data for analysis and interpretation (Hulbert and Harding, 1980), the South African Sugar Association (SASA) encouraged the keeping of more comprehensive farm records at a field level through the development and support of the Field Record System (FRS). Other record systems besides the FRS are available to the South African sugar industry, e.g. those of private consultants such as Spencer Holley Agronomic Services. While the information that can be derived from such records may facilitate improved farm planning and control, the data are usually available only in an unprocessed form and are underutilised because of the extreme difficulty experienced in observing and interpreting production trends. With the demise of the sugarcane quota system (1997) and the decommissioning of the FRS (1998) it is appropriate to review the parameters recorded in field records and used to guide management decisions.

2.1 SUGARCANE FIELD RECORD ANALYSIS IN SOUTH AFRICA

It has been customary to summarize field records in order to simplify interpretations because of the numerous factors involved. Summarizing usually reduces the data to a single factor, and yield differences between fields have often been ascribed to variation in this factor only (Hellmann *et al.*, 1995), e.g. yield expressed as a function of rain (Hellmann, 1988). Data manipulation and processing techniques have become increasingly sophisticated since the development of the computer, but the scope of formal analysis of South African commercial sugarcane field data has remained limited. Hoekstra (1976) analysed commercial field records, calculating standardized yields based on discounted cash flows to guide management when deciding whether to plough out a ratoon crop of sugarcane or not. Landrey *et al.* (1981) used field records to determine the most appropriate crop management strategies during times of quota restrictions, but were unable to account for yield variations at a field level. Tucker (1975), Hulbert and Harding (1980), Cluverwell (1984) and Hellmann (1988) have demonstrated the potential of commercial field data analysis for assessing farm productivity and the effect of agronomic practices (harvesting age and season) on sugarcane yield. These researchers noted the need to edit and correct field records prior to analysis. Their interpretations were all based on an analysis of the mean of each recorded variable, e.g. variety, crop age at harvest, sugarcane yield, or interaction tables limited to the investigation of only two variables at a time, e.g. crop age in relation to harvest month. Erroneous interpretations may be made when only the means of data are investigated. Hellmann (1988) found that a major source of these errors was firstly the wide range of most variables between individual fields comprising the mean, and secondly variability inherent in the data, reflecting natural field conditions, which could not be accounted for by the simplistic and limited analysis techniques used. Multivariate statistical techniques may be used to overcome this problem because instead of using only the mean of data, any number of independent variables (measured field conditions) can be accommodated in regression equations (Steel and Torrie, 1981). Such techniques have not been used for field record analysis in the South African sugar industry (Murdoch, 1996, personal communication [1]) but have been applied successfully in the South African timber industry site factor studies (Grey, 1983; Schutz, 1990; Schönau and Aldworth, 1991; Donkin, 1994) and commercial sugarcane field data analyses in Florida (Alvarez *et al.*, 1982) and in the Philippines (Early, 1980).

The two most recent formal analyses of commercial field records in the South African sugar industry include the use of some simple statistics (coefficients of variation, standard deviations, standard error of the difference of the means and confidence intervals) for comparisons between agronomic practices (Hellmann, 1993), and the use of a geographic information system for spatial presentation of data and database query facilities (Hellmann *et al.*, 1995). While these techniques allow for a more reliable interpretation of the differences between treatments, only one recorded variable (crop cycle) has been investigated (Hellmann, 1993) and no statistical analyses have been conducted to investigate which of the routinely recorded field variables best describe observed sugarcane yield variations. Such an analysis would be useful to focus

---

[1] Mr MG Murdoch, Head, Biometry and Data Processing
South African Sugar Association Experiment Station,
Private Bag X02, Mount Edgecombe, 4300

management attention on important field variables (or proxies of important variables) determining sugarcane yield. Since data collection is not income generating (for the farmer) only those variables that are required to facilitate responsible decision-making should be recorded. These objectives can best be achieved using crop models (France and Thornley, 1984).

2.2 CROP MODELS

Two general types of crop model can be developed and used (France and Thornley, 1984) viz. yield models and growth models. Yield models are usually derived from empirical functions which aim to predict yield at a specific location and summarize large quantities of data (field records) in terms of a few important parameters (Thornley, 1976; Steel and Torrie, 1981). Growth models simulate crop growth processes, and yield is calculated from the predicted growth.

While yield models predict crop yield, they cannot be used to explain why a certain yield was achieved because the relation between the dependent (yield) and independent variables (various recorded field conditions) is purely statistical. When an explanation becomes necessary a crop growth model is required to model the physiological response of the sugarcane plant to specified conditions (cause and effect relations). These models provide a powerful tool for evaluating crop management alternatives and can be used to identify factors limiting sugarcane production by comparing the predicted yield with the observed yield.

Crop (sugarcane) models have evolved to accommodate the needs of three fairly distinct user groups (France and Thornley, 1984) viz.;

a research scientists who aim to extend the boundaries of current knowledge,

b applied scientists, including agronomists and plant breeders, whose objectives are to improve the efficiency of production techniques, and

c farmers, advisors and extension staff, who are concerned mainly with improving production levels.

A single complex crop growth model may be able to handle this range of applications when adequate input data are available (Tovey and Bull, 1977), but because of the vastly different objectives, it is usually necessary to provide models geared to specific user requirements (France and Thornley, 1984). Models also need to be specified at a level that is compatible with the availability and scale of input information and it is imperative that the variables included in the model are easily accessible to the model users (Keig *et al.*, 1991). As a result it is not unusual to find a number of models of different complexity being used within an industry.

In sugarcane research, plant growth processes which are thought to influence crop yield are studied in ever increasing detail. Dynamic growth models are built to extend into areas where there is uncertainty and a lack of understanding. As a result much is known about photosynthetic and metabolic pathways and how these affect crop carbon balances, root extension, leaf initiation, enzyme kinetics, etc., but less is understood about the interaction of these processes to determine sugarcane yield. Modelling is done on a day by day basis at an individual plant level. Data requirements are therefore rigorous and include detailed and sophisticated measurements of environmental conditions (soil, site, climate and management) to which the sugarcane plant is exposed. This generally restricts commercial-scale applications of the models (France and Thornley, 1984).

The applied scientist requires less physiological detail than the researcher. Consequently models are used which simulate the reasonably well understood plant growth processes and describe areas of uncertainty with empirical functions. Depending on the application, either crop yield or crop growth models may be appropriate.

Research models are generally too demanding on input requirements to be of practical use to farmers (Thompson and Boyce, 1971a; France and Thornley, 1984) and yield expectations at a field level are determined very simplistically by the grower, based largely on his knowledge and experience and on as much information and advice as can be accumulated at the time. Thus decisions are presently characterised as being more subjective than objective (scientific). The complexity of the relations and interactions of the numerous field variables, as well as the large number of fields that must be evaluated simultaneously, suggest that a more formal assessment may be of value in providing improved information from field records (Alvarez *et al.*, 1982).

The agriculturalist requires a functional model which is easy to use and achieves a certain level of accuracy within a tested (though restricted) range for yield predictions at a field level. Planners and managers also require definitive answers about production levels as resources become increasingly limited. There is a need to evaluate the relative performance of fields, estates and districts between seasons as objectively as possible. Empirical relations (yield models) are usually adequate for these purposes (France and Thornley, 1984) and were recommended for agriculturalists by Thompson and Boyce (1971a). It is therefore likely that yield models will be more appropriate than growth models for achieving the broad objectives of this study which relate to the use of commercial field records for predicting sugarcane yields.

2.2.1 Yield models

Sugarcane yield may be predicted with empirical functions using variables readily available from growers' records and secondary sources that are correlated with yield, without examining plant growth processes. This modelling approach has been used by researchers to successfully predict sugarcane yield from mill district data in the Philippines (Bouldin, 1969, cited by Early, 1980; Panol, 1974), commercial field data in Florida (Alvarez *et al.*, 1982) and in the Philippines (Early, 1980), and experimental plot data (Thomas and Oether, 1976). Empirical models have also been widely used to determine the relation between crop water use (total evaporation) and sugarcane yields (Vazquez, 1970, cited by Kingston, 1994; Thompson, 1976; Shih and Gascho, 1980, cited by Kingston, 1994; Chang *et al.*, 1983, cited by Kingston, 1994; Kingston, 1994).

Regression analysis, using least-square or maximum likelihood estimates for model parameters, was used to derive these sugarcane yield models and has also been the most commonly used statistical technique in timber site factor studies (Carmean, 1975; Grey, 1983; Schutz, 1990). Linear regression or stepwise and best subsets multiple regression techniques were used depending on whether yields were predicted using correlations with univariate or multivariate sets of conditions. In these applications one of the most important tenets of regression analysis, viz. that independent variables are orthogonal of each other, is usually violated because recorded field variables, particularly soil and site properties, are highly covariate by nature (Donkin, 1994). Intercorrelation (multicollinearity) of the data causes the correlation matrix to approach singularity, and gives rise to imprecise estimates of the regression coefficients (Steel and Torrie, 1981). As a result, care must be exercised when evaluating the contribution of a partial regression coefficient to the observed variation in yield when multicollinearity is a problem in the models since the actual importance of the variable may be over- or understated. Although it is often difficult to determine the importance of the field variables independently of one another in multiple regression, principle component analysis and multiple discriminant analysis may be used to select uncorrelated site variables (Grey, 1983) together with ridge regression (Schutz, 1990) and principle components regression (Donkin, 1994) to improve these estimates.

Consequently, when the objective is to describe a whole-system response, and there is little interest in the reason why the system responds the way it does, empirical models can be applied very effectively (Thornley, 1976) and may be used as a basis for the evaluation of sugarcane yields both within and between seasons. Empirical modelling is an appropriate technique when models are developed and applied within a defined area for which a large number of suitable records are available.

Great care needs to be exercised in the causal interpretation of statistically significant relations (Schönau, 1987). While it is possible that a cause – effect relation between a variable included in the model and yield exists, it is equally possible that the variable acts as a proxy for another factor causally related to yield, or the association may be purely circumstantial (Schönau and Aldworth, 1991), i.e. it describes certain conditions well which are particular to the data set. Interpretations of the biological meaning of parameters included in a model are therefore tenuous until the trends and relations in question can be tested in controlled experiments or modelled by complex growth models, although the yield model output (predicted yield) *per se* is usually very useful to managers (Schutz, 1990).

Although no predictive accuracy norms are available for sugarcane models in the literature, models developed for the Philippines by Bouldin (1969, cited by Early, 1980) using soil properties, varietal differences, climatic factors and fertilizer inputs as predictors, accounted for 30 to 40% of yield variability. Similar statistics are reported by Phillips *et al.* (1989) for yield models from Hawaii based on climatic variables, altitude, soil properties and management practices. The greatest predictive accuracy reported for sugarcane yield models is a coefficient of multiple determination of 0.65 for Florida sugarcane, based on climatic factors, management practices, soil differences, past crop performance, varietal differences and agronomic inputs (Alvarez *et al.*, 1982). Grey (1983) suggests that for timber site factor studies a carefully chosen group of variables should explain from 65% to 80% of the variation in timber yield and should predict yields to within commonly acceptable management standards. According to Grey (1983), models failing to meet these criteria should be checked to ensure that important factors that explain tree growth have not been omitted, significant interactions and transformations have not been overlooked, and model specification is correct. Published sugarcane yield models for site-specific predictions consistently account for less variation than the target set by Grey (1983). However, an important difference between the timber site factor studies and the sugarcane models is that the former are generally based on sample plots, using estimated yield indices at reference ages, while the latter are usually based on actual yields. The considerable natural variation within sugarcane fields and large yield differences between successive harvests provide data that are inherently more variable and difficult to model than that of timber sample plots. Hence models with poorer predictive ability should, perhaps, be expected for sugarcane.

2.3 MEASURES AND MODELS OF SUGARCANE YIELD

Sugarcane yields are important because they determine enterprise income and have a direct bearing on variable production costs. Sugarcane growers are paid on the basis of sucrose yield for a season while important production costs (harvesting and haulage) are functions of cane (sugarcane stalk) yield. The two main components of sucrose yield are cane yield and sucrose concentration. A high cane yield is the result of a high biomass production which is caused by high temperature, optimum mineral nutrition and high precipitation. These environmental conditions result in a disproportionate sucrose yield because of a low sucrose concentration. This is because sucrose concentration increases with low temperatures, water stress and nutrient stress, at least during the ripening / maturation period (Van Dillewijn, 1952).

Since the sugar industry uses two indices of crop yield, sugar production may best be considered as the result of two component processes:

a the vegetative growth of the sugarcane plant (cane yield); and b the accumulation of sucrose (sucrose yield).

These distinct processes are best modelled separately as in the studies by Vazquez (1970, cited by Kingston, 1994), Panol (1974), Thompson (1976), Ometto (1977), Early (1980), Shih and Gascho (1980, cited by Kingston, 1994), Alvarez *et al.* (1982), Blume (1983), Shih (1988, cited by Kingston, 1994), Inman-Bamber (1991a), Keig *et al.* (1991) and Kingston (1994) because:

a given a set of environmental conditions, the one can be determined without reference to the other;

b they respond differently to changes in environmental conditions (Van Dillewijn, 1952);

c they change in different ways over time and as a result the relation between them is not uniform (Van Dillewijn, 1952); and

d individually estimated cane and sucrose functions may be more accurate and useful for purposes other than field yield estimates, e.g. a predictor of cane yield in a study to consider transport requirements (Alvarez *et al.*, 1982).

Commercial sugarcane yields are usually normalized by field area, yield being expressed on a per hectare basis, i.e. tons cane per hectare (TCH) and tons sucrose per hectare (TSH). This allows yield comparisons to be made between fields of different sizes. Production efficiency is usually expressed as an index of yield in terms of a limiting resource, e.g. time, expressed as tons cane per hectare per month (TCHM) and tons sucrose per hectare per month (TSHM). These age-corrected yields allow yield comparisons to be made between crops harvested at different ages and give an indication of sugarcane-farm productivity.

Crop yields in general depend on climatic conditions, and if there is no soil or nutrient stress they depend exactly on water utilization, which is in turn largely a function of solar radiation (Doorenbos and Pruitt, 1977; Doorenbos and Kassam, 1979). Sugarcane is phenologically simpler than other important graminaceous crops like rice, wheat or maize because the economic product (sucrose) is a direct product of photosynthesis. Thompson and De Robillard (1968) suggested the possibility of a linear relation between cane yield and crop water use (total evaporation) after collating 20 years' data from irrigation experiments in South Africa.

Thompson (1976) collected yield data from sub-tropical sugarcane production areas (Hawaii, Australia and Mauritius together with data from SASA experiments at Mount Edgecombe, Chaka's Kraal and Pongola), reporting cane yield from 91 observations in terms of total evaporation as a simple linear regression function from 600 mm to 3800 mm water, but found statistically significant deviations from linearity for the relation between sucrose yield and crop water use from 85 observations from 600 mm to 1800 mm water.

Kingston (1994) reports six linear functions for cane yield for annual sugarcane crops from Florida, Puerto Rico, Taiwan and Hawaii, and three linear functions for sucrose yield for up to 1600 mm total evaporation from Florida and Puerto Rico, establishing similar empirical models using Australian data (22 observations). A linear relation between cane yield and total evaporation is likely because cane yield is a major component of biomass yield (Kingston, 1994). A linear relation should therefore also apply to sucrose yield since it is primarily a function of cane yield (and also sucrose concentration). Helwig (1991) questioned the widespread assumption of linear crop-water functions on the basis that the full scope of the response range may not have been investigated. This could explain why Thompson (1976) obtained a non-linear sucrose yield function for crop water use to 1800 mm, while other researches reported linear functions for up to 1600 mm total evaporation. Nevertheless, water use efficiency data for sugarcane reported in the literature ranges from 7.4 – 16.9 TCH 100 mm^{-1} total evaporation and 0.5 – 1.94 TSH 100 mm^{-1} total evaporation (Kingston, 1994).

2.3.1 The Thompson yield models

The Thompson (1976) equations for cane and sucrose yield are very simple, based on the consumptive use of water by sugarcane and operating on the assumption that the total evaporation process does not distinguish between soil water evaporation and transpiration (Schulze *et al.*, 1995). Yields are predicted using the following equations:

$$\textit{Cane yield (TCH annum}^{-1}\textit{)} = 9.69\left(\frac{E_{ann}}{100}\right) - 2.4 \qquad \text{.........................}(1)$$

where E_{ann} = annual total evaporation (mm)

SE = 15.1 TCH annum^{-1}

n = 91

r^2 = 0.90

$$\textit{Sucrose yield (TSH annum}^{-1}\textit{)} = -22.65 + 4.923\left(\frac{E_{ann}}{100}\right) - 0.1419\left(\frac{E_{ann}}{100}\right)^2 \text{.........}(2)$$

where E_{ann} = annual total evaporation (mm)

SE = 3.0 TSH annum^{-1}

n = 85

R^2 = 0.61

Expected crop yield is determined from mean annual potential evaporation (Class A-pan), applying a crop factor of 0.8 to allow for periods of incomplete canopy, fallow periods, drying off periods, and possible differences between total evaporation from a fully canopied sugarcane crop and potential evaporation (Thompson, 1977). These equations have been used extensively to predict irrigated sugarcane yields at the estate level when calibrated for farm management ability (Thompson, 1977; Thompson and Harding, 1986). The cane yield model is especially robust and has formed the basis of the rule-of-thumb used in the South African irrigated sugarcane sector viz. 9 TCH 100 mm^{-1} total evaporation (experimental yields) and 6.5 TCH 100 mm^{-1} total evaporation (performance of an "average" commercial grower). Although the models have stood the test of time (Schulze *et al.*, 1995), they have a number of shortcomings.

Cane and sucrose yields are estimated for a 12 month (annual) period and not for the duration of a crop cycle, which falls mostly into the 18 to 24 month range in the Midlands. To obtain the yield for one harvest, the estimated yield has to be adjusted to account for the duration of the crop cycle. In addition, sugarcane is a dynamic crop with no set planting date or length of growing season, and the importance of starting and harvesting seasons of crop cycles in relation to both cane and sucrose yields has been demonstrated (Lonsdale and Gosnell, 1975; Landrey *et al.*, 1981; Sweet and Patel, 1985; Hellmann, 1988; Inman-Bamber, 1991a; Hellmann, 1993).

The equations were developed predominantly from irrigation experiments where water was not growth limiting. Midlands sugarcane is grown almost exclusively under rainfed conditions where water is invariably the single most limiting resource (Thompson, 1976; Inman-Bamber, 1995). Thus the models are considered to be more applicable to irrigated sugarcane than to rainfed production systems (Hughes, 1992). Additional model limitations are that similar cane yields are achieved with new sugarcane varieties but the sucrose percentage is usually improved, changing the sucrose yield relation (Smith, 1994), and that retarded growth occurs in spite of favourable growing conditions which cannot be accounted for, e.g. as a result of frost damage during the preceding winter, severe hail damage, etc.

2.3.2 The Smith yield model

In spite of the various limitations of the Thompson models, the Thompson (1976) cane yield equation forms the basis of the Smith (1994) yield prediction functions. Smith (1994) determines irrigated sugarcane yield potential according to the method described by Thompson (1977), and modifies this approach to estimate potential rainfed yields. Thus, rainfed cane yield is estimated, based on the norm of 9 TCH 100 mm^{-1} total evaporation (Thompson, 1977) but, because rainfall is generally less than potential evaporation in the sugarcane producing areas, total (mean) annual rainfall is assumed to be used by the crop. Further, because of inherently different yield potentials, the yield norm is adjusted according to production region (North Coast > South Coast > Midlands) and soil type (well drained deep clays > sands and shallow soils), 9 TCH 100 mm^{-1} rainfall being the greatest productivity allowed (Table 1).

Table 1: Annual incremental potential sugarcane yield per 100 mm water per hectare and regional sugarcane harvest cycles (after Smith, 1994).

Soil type †	North Coast (TCH 100 mm^{-1})	South Coast (TCH 100 mm^{-1})	Midlands (TCH 100 mm^{-1})
"Good" soil	9.0	7.5	6.5
"Poor" soil	7.5	6.5	5.0
Harvest cycle	12 - 14 months	14 - 18 months	18 - 24 months

† For sugarcane, "good " soils are considered to be deep (1000 mm), well drained soils, predominantly clays and clay-loams. "Poor" soils are defined as sands and shallow soils (<500 mm deep) (Smith, 1994).

To use the model, mean annual rainfall (dryland sugarcane) or mean annual potential evaporation (irrigated sugarcane) and the soil type of the study area need to be known. Cane yield is then calculated in three steps. First, the annual potential cane yield is calculated by multiplying rainfall (or evaporation) by the applicable TCH value from Table 1. Second, yield is corrected from the annual value (Table 1) for the applicable cutting cycle. Third, since the yields in Table 1 are potential (experimental) yields, they should be adjusted downwards using a factor to account for different levels of management ability.

The assumption that rainfall is 100% effective and that all of this water is transpired by the crop, is unreasonable when determining crop water requirements. It may, however, be acceptable for empirical yield functions, depending on the level of model sophistication because effective rainfall is difficult to measure. Thompson and De Robillard (1968) reported rainfall efficiencies of 85% for rainfed sugarcane (760 mm annum^{-1}) and found that rainfall efficiency declined with increasing precipitation. Since rainfed sugarcane is produced mainly in higher rainfall regions, a rainfall efficiency of 70% is usually assumed (Thompson and Boyce, 1968; Thompson, 1977). It has only become possible to improve these simple estimates by using soil water balance models (Inman-Bamber *et al.*, 1998). Since these balances require daily climatic input data and detailed measures of soil properties, they are probably too sophisticated to include in most of the crop yield models. For this reason effective rainfall can at best be estimated by reducing actual rainfall by a constant factor in yield models. This method of estimating the effective component of rainfall is unrealistic depending on the amount of rain, the season and the soil water content. In terms of statistical modelling, nothing is to be gained by reducing rainfall amounts by a constant factor. For this reason the use of total rainfall data to calculate yield is acceptable.

It is important to estimate total evaporation accurately when using the Smith (1994) yield functions because the model is driven by the consumptive water use of sugarcane. Since total rainfall is used in preference to estimated effective rainfall in the rainfed yield functions, the Smith (1994) model might be expected to overestimate yields. It is likely that these inaccuracies are off-set by the arbitrarily defined (reduced) regional and soil yield potentials, and otherwise adjusted using a factor to subjectively rate management ability. A major weakness of this model is that no validation or model calibration procedure has been published to indicate its performance. However, the model is simple and easy to use as a general rule-of-thumb, perhaps extending the Thompson rule of 9 TCH 100 mm^{-1} water for irrigated sugarcane to include the expected production attainable in the dryland sugarcane production areas of KwaZulu-Natal.

The model is suitable for regional modelling applications and initial enterprise planning at an estate level when rainfed yields are predicted using mean annual rainfall. Variable annual rainfall causes large variations in predicted yields and the model should not be used at the field level because yield differences between adjacent fields will be inadequately explained by differences in total rainfall. The model also suffers from the same shortfalls as the Thompson (1976) yield model (Section 2.3.1) on which it is based.

2.3.3 The ACRU sugarcane yield submodel

The Agricultural Catchment Research Unit (ACRU) modelling system (Schulze *et al.*, 1995) has a sugarcane yield submodel based on the Thompson (1976) yield models. The submodel is used in ACRU to predict irrigated and rainfed sugarcane yields on a catchment scale using crop plant and harvest dates together with the ACRU soil water budget. The soil water budget is determined by ACRU climate and soil parameters. These parameters include daily rainfall, daily maximum and minimum temperatures, soil type, texture, profile depth (A and B horizon), effective depth for storm flow generation, potential maximum rooting depth, effective total rooting depth, soil depth to which the majority of soil water extraction takes place for a fully grown crop, A – B response (movement of water from the A horizon to the B horizon), base flow response (movement of water from the subsoil to ground water), field capacity (A and B horizon), wilting point (A and B horizon), and porosity (A and B horizon). In ACRU, total evaporation is assumed to decline from the potential when total available water (TAW) has been depleted by 60% (Schulze *et al.*, 1995). Total available water may be modelled in terms of soil water relations and where different soil types or soils with slightly different properties are found within a catchment, these can be accommodated in repeat runs of ACRU to improve model resolution.

Many of the necessary soil parameters are not usually available to ACRU users and therefore default norms have been programmed for the South African soil families according to the Taxonomic Classification System (Soil Classification Working Group, 1991). Thus the model is simplified, and although model sensitivity is reduced, it provides a powerful tool for regional modelling. At a field-and-estate scale of modelling it is unlikely that such a model will adequately accommodate observed variations in yield, especially since the weaknesses of the Thompson yield models have been retained. In addition, the ACRU soil water budget is unable to account for the additional water available to crops through water conserving agronomic practices such as strip cropping and mulching. These husbandry practises are used extensively in the South African sugar industry and Thompson (1965) demonstrated that water conservation from trash mulching resulted in yield increases of the order of 9 TCH annum^{-1} under rainfed conditions in KwaZulu-Natal.

Models based on sugarcane thermal time [2] (base 12 °C) and accommodating ratoon status and water stress (Hughes, 1992) are currently being tested for ACRU (Schulze *et al.*, 1995) with a view to addressing some of shortcomings of the Thompson models for rainfed crops. Thermal time was included in the models being evaluated for ACRU to account for the different growth rates (and hence crop factors) for crop cycles starting in different seasons (Hughes, 1992). To calculate thermal time, daily mean temperatures are required. Thus yield is modelled on a daily

---

[2] Thermal time / growing degree days (°C d) / heat units:
The difference between mean daily temperature and the crop's temperature threshold for active growth.

time scale up to the stage of full canopy formation in these equations. Since the required daily climatic data are available within the ACRU modelling framework, the dynamic yield model is justified (Hughes, 1992). Because yield models predict yield for entire crop cycles only, static models are generally adequate (France and Thornley, 1984) and other less sophisticated approaches are reported in the literature to account for the effect of seasonally variable growth potential. These methods are discussed in Section 2.3.6.2.

Sugarcane yields are known to decline with successive ratoons and this yield decline is reflected in crop water use (Moberly, 1974; Thompson and Boyce, 1971b; Thompson, 1976; 1986). To account for this reduction in total evaporation and yield the models currently being evaluated for ACRU use a maximum crop factor (for sugarcane with a full canopy) that is progressively reduced for successive ratoons (Hughes, 1992).

Rainfed sugarcane frequently experiences periods of water stress in South Africa and during these periods total evaporation drops below the potential rate (Thompson, 1976; Inman-Bamber, 1995). In these situations the assumption of a constant relation between actual total evaporation and potential evaporation using a constant crop factor (Thompson, 1977) is unreasonable. It follows that to be able to predict sugarcane yields under rainfed conditions, a model needs to simulate the decline in total evaporation with the onset of water stress as well as its recovery once the stress has been relieved. The decline and recovery of total evaporation would have to be simulated for the duration of the crop cycle, but no such evaporation models have been developed (Hughes, 1992). To account for this water stress related decline in total evaporation, Hughes (1992) derived empirical crop factor reduction and recovery curves for ACRU from data (Moberly, 1974 and Thompson, 1986) for mature, water stressed sugarcane in lysimeter experiments. These reduction and recovery curves are included in the models being evaluated for ACRU (Schulze *et al.*, 1995).

2.3.4 The Doorenbos and Kassam yield model

A simplified equation describing the water : crop yield relation for 26 important irrigated crops, including sugarcane, was developed by Doorenbos and Kassam (1979). The models are used to evaluate the effect of water supply on crop yield for FAO[3] applications in the planning, design and operation of irrigation schemes. The equation takes the following form:

$$\left(1 - \frac{Y_{act}}{Y_{\max}}\right) = k_y \left(1 - \frac{E_{act}}{E_{\max}}\right) \qquad \text{..........(3)}$$

where Y_{act} = actual harvested yield (dependent variable)

Y_{max} = maximum harvested yield

k_y = yield response factor

E_{act} = actual total evaporation

E_{max} = maximum total evaporation

The model predicts sugarcane yield (Y_{act}) in terms of a benchmark yield (Y_{max}) and crop water use (E_{act}), expressed as an index relative to potential water use (E_{max}) to evaluate the level of water stress the crop experiences over the growing period. The equation can thus be used to assess crop yields under conditions of both adequate and limited water supply.

The maximum yield (Y_{max}) is determined for the most adapted sugarcane variety as dictated by climate (temperature and radiation) and the time taken to reach maturity, assuming that other factors are not limiting, e.g. water, fertilizer, pests and diseases. The intention is to set a benchmark for the maximum yield that can be obtained under actual farming conditions. Doorenbos and Kassam (1979) calculate sugarcane Y_{max} according to the agro-ecological zone method (Kassam, 1977) which was developed for yield assessments at a continental scale. Other methods of determining realistic estimates of Y_{max} may be used as appropriate for the scale of the investigation. Where economic conditions do not restrict production and in a constraint-free environment, $Y_{act} = Y_{max}$ when the crop's full water requirements are met (Doorenbos and Kassam, 1979).

Water stress is the only factor that can reduce Y_{max} yields in the Doorenbos and Kassam yield

---

[3] Food and Agriculture Organization of the United Nations.

model. The level of plant water stress is quantified by the rate of actual total evaporation (E_{act}) relative to the rate of maximum total evaporation (E_{max}). When crop water requirements are fully met by the available soil water supply, the actual crop water use is equal to the maximum (potential) crop water use, i.e. $E_{act} = E_{max}$; when the soil water supply is insufficient, $E_{act} < E_{max}$ and yield is reduced, i.e. $Y_{act} < Y_{max}$.

The yield response factor (k_y) is an empirically derived factor to describe the relation between relative yield decrease and relative total evaporation deficit. The value of k_y for sugarcane is based on the evaluation of research results over a wide range of growing conditions (Doorenbos and Kassam, 1979). A k_y value is given for both the total growing period and the individual growth phases of the crop which sometimes differ in their yield:water stress response. If $k_y < 1$ the decrease in yield is proportionally less than the extent of the water deficit, i.e. for a water stress insensitive crop / growth phase; when $k_y > 1$ the decrease in yield is proportionally greater than the level of water stress, i.e. a water sensitive crop / growth phase. For sugarcane $k_y = 1.2$ for the total growing period; $k_y = 0.75$ for the vegetative growth phase – establishment, tillering and stem elongation; $k_y = 0.5$ for the yield formation phase – some stem elongation and maturation; and $k_y = 0.1$ during the ripening phase (Doorenbos and Kassam, 1979). Although the cane yield of sugarcane is more or less directly proportional to crop water use (Thompson and De Robillard, 1968; Thompson, 1976) sucrose yield is not reduced by moderate water stress, and during the ripening period sucrose content (and therefore total sucrose yield) is increased by restricting the water supply (drying off the crop) (Van Dillewijn, 1952; Thompson, 1976).

Doorenbos and Kassam (1979) determine maximum total evaporation (E_{max}) from reference evaporation and crop coefficients appropriate for the stage of crop development. The method of determining reference evaporation may vary according to the available climatic data. To determine actual total evaporation (E_{act}) the level of soil TAW must be known. The soil TAW is determined as a function of the crop's rooting depth and the soil water content between field capacity and wilting point. A portion of the TAW is easily taken up by the plant so that $E_{act} = E_{max}$, but thereafter $E_{act} < E_{max}$. For sugarcane $E_{act} = E_{max}$ for soil water depletion levels ranging from 0.875 TAW when $E_{max} = 2$ mm d $^{-1}$ to 0.5 TAW when $E_{max} = 7$ mm d $^{-1}$ (Doorenbos and Kassam, 1979). Doorenbos and Kassam (1979) use this as a basis for determining irrigation system design capacities and crop irrigation schedules. Under rainfed conditions the soil water balance needs to be simulated to determine reliable estimates of E_{act}. The Doorenbos and Kassam yield model is being included in the ACRU agrohydrological model [4] where estimates of E_{act} can be made for the catchments being modelled.

The Doorenbos and Kassam yield model is a valuable tool for broad scale yield predictions because it applies to a wide range of agriculturally important crops. For sugarcane yield predictions the model limitations are similar to those of the Thompson yield models because yield is explained only in terms of crop water use. It is also critical that the user-defined maximum yield is accurate in order to ensure realistic model outputs. Maximum yield must be adjusted according to crop harvest ages and crop cycle starting and harvesting dates to account for inherently different yield potentials (Lonsdale and Gosnell, 1975; Landrey *et al.*, 1981; Sweet and Patel, 1985; Hellmann, 1988; Inman-Bamber, 1991a; Hellmann, 1993) making the model extremely clumsy to use for specific yield predictions at a field scale.

2.3.5 The CANEGRO sugarcane growth model

Although not a yield model, an evaluation of crop models suitable for field-level sugarcane yield predictions in South Africa would be incomplete without a careful consideration of this model. The South African Sugar Association Experiment Station (SASEX) has been developing a process level growth model (CANESIM, later called CANEGRO) during the last decade (Inman-Bamber and Thompson, 1989; Inman-Bamber, 1991b; Inman-Bamber *et al.*, 1993; McGlinchey *et al.*, 1995; Van Antwerpen *et al.*, 1996). The modelling project has received substantial research attention and as model development continues, sugarcane growth processes are being simulated with ever increasing detail and accuracy. Both long-term (using historic data or generated scenario data) and short-term crop production predictions (using current data) can be made (Inman-Bamber, 1991b; Inman-Bamber *et al.*, 1993).

---

[4] Mr NL Lecler, Research Fellow
School of Bioresources Engineering and Environmental Hydrology,
Faculty of Engineering, University of Natal,
Private Bag X01, Scottsville, 3209
Present address: Zimbabwe Sugar Association Experiment Station,
Private Bag 7006, Chiredzi, Zimbabwe

The CANEGRO model has been developed along the lines of the IBSNAT [5] models and has been validated over a wide range of climate and soil conditions in South Africa. The model comprises detailed balances for carbon, energy and water. Exchanges between these balances and the plant occur at the root / soil water and canopy / atmosphere interfaces. An outline of the CANEGRO model is presented in Figure 2. A selection of subroutines are available for simulating these balances depending on the detail of the available input data (Inman-Bamber *et al.*, 1993; McGlinchey *et al.*, 1995; Van Antwerpen *et al.*, 1996). Crop growth is modelled only for the NCo376 variety at the individual plant level on an hourly or daily temporal scale, depending on the modelling subroutines chosen.

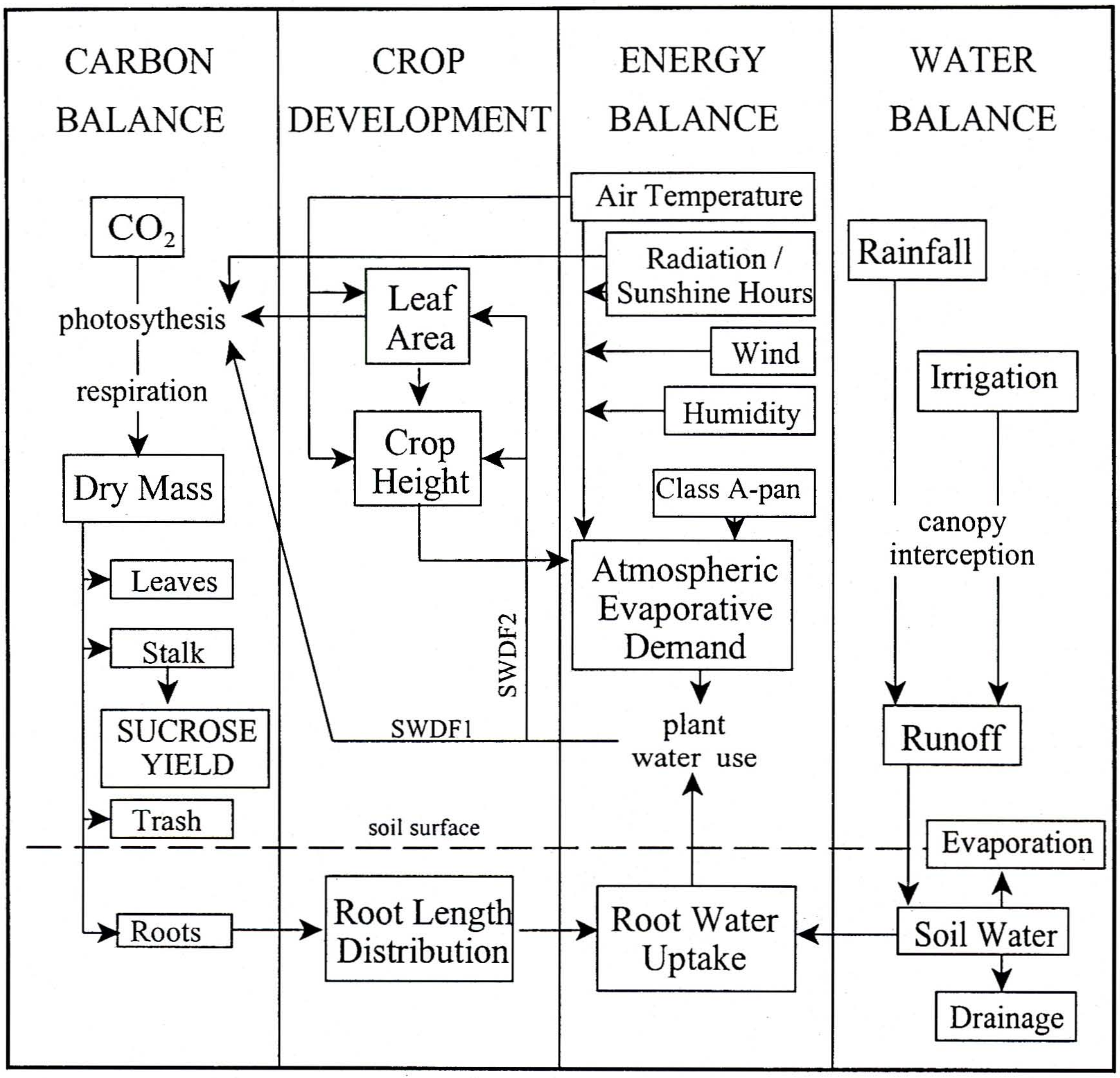


Figure 2: Flow chart of the CANEGRO model (adapted from Inman-Bamber *et al.* 1993 and McGlinchey *et al.* 1995)

---

[5] International Benchmark Sites Networks for Agrotechnology Transfer

The carbon balance simulates plant photosynthesis, respiration and partitioning of photosynthate. The rate of dry matter accumulation in the plant (including the roots) may be simulated using either a simple model of the daily plant carbon status according to the method of Inman-Bamber and Thompson (1989), based on the Lorber model (Lorber *et al.*, 1984) and the work of McCree (1970) and Heskith *et al.* (1971), or using a modification (Inman-Bamber *et al.*, 1993) of the Hedgerow model (Boote and Loomis, 1991) which accounts for hourly variations in sunlit and shaded fractions of the canopy from daily radiation data.

The mechanisms of dry matter partitioning within the plant are highly complex and because they are not yet fully understood, empirical associations between total dry mass and dry matter distribution ratios are used to partition photosynthate between the sugarcane plant's leaves, stalk and roots (Inman-Bamber and Thompson, 1989). Although the energy balance was calibrated using experimental data mainly from irrigated NCo376 crops (Inman-Bamber and Thompson, 1989) the model adequately accounts for dry matter accumulation in rainfed NCo376 crops (Inman-Bamber, 1991b).

Cane yield is determined empirically by CANEGRO in the energy balance as the component of dry matter partitioned to the stalk, multiplied by the stalk population which is reasonably predictable for mature sugarcane. A final stalk population of 133 000 stalks ha^{-1} is assumed for NCo376 (Inman-Bamber, 1991b). The model predicts stalk dry mass and the result therefore needs to be divided by the dry matter content of sugarcane stalks, about 29% (Glover, 1972), to derive cane yield (wet mass) as used in the sugar industry. Partitioning of dry matter fractions within the sugarcane stalk to brix [6] and juice purity [7] is also described empirically within the CANEGRO model using multiple regression equations (Inman-Bamber, 1991a). Sucrose yield is determined from the product of predicted brix and purity (dry matter sucrose content), multiplied by the predicted dry matter cane yield. Although the data sets used to derive the

---

[6] The term used for the solids concentration of an impure sucrose containing solution measured using a specially calibrated refractometer (Anon., 1977).
Brix = Total soluble solids = sucrose + non-sucrose

[7] The percentage of sucrose in the solids fraction of a sugar containing solution; i.e. the sucrose content of brix. Purity (%) = Sucrose / Total soluble solids (100)

empirical yield equations were relatively small (n = 33, $R^2 = 0.96$ for dry matter partitioning within the plant (Inman-Bamber and Thompson, 1989); n = 144 for dry matter partitioning within the stalk, brix equation $R^2 = 0.25$ and juice purity equation $R^2 = 0.34$ (Inman-Bamber, 1991a)), Inman-Bamber *et al.* (1993) found simulated cane and sucrose yields to be similar to field records on an irrigated estate when total precipitation was low, but that simulated yields overestimated actual performance when precipitation was adequate.

The energy balance simulates crop water demand as potential evaporation (atmospheric evaporative demand). A modified version of the Penman-Monteith evaporation equation (Monteith, 1965) is used with the profile equations (reference height = 10 m) of Monteith and Unsworth (1990) to estimate daily potential evaporation (Inman-Bamber *et al.*, 1993; McGlinchey *et al.*, 1995). Radiation may be measured or derived from daily sunshine hours using Thompson's (1986) calibrated Angstrom equation (Inman-Bamber and Thompson, 1989). When daily relative humidity and / or wind speed data are not available, class A-pan evaporation × 0.9 for a fully canopied crop (Thompson, 1976) is used to determine potential evaporation (Inman-Bamber, 1995).

The water balance simulates crop water supply. It is a modified version of the CERES-Maize crop growth model (Jones and Kiniry, 1986) water balance subroutine (Inman-Bamber, 1991b). This subroutine comprises a detailed soil water budget to determine soil TAW. Data defining the soil water holding and water release characteristics are required to depth for soil layers in 10 – 15 cm increments. Water inputs are obtained from rainfall, and also from irrigation when applicable, taking into consideration crop canopy water interception and runoff, while water is removed from the profile by transpiration, evaporation and drainage (Inman-Bamber, 1991b). The energy and water balances are closely linked in CANEGRO because plant water use is controlled by the energy balance when the soil water content is high, but when the water supply is limited water use is controlled by the water balance (Van Antwerpen *et al.*, 1993). Crop water stress is assumed to occur when the amount of water required by the energy balance exceeds the amount that the roots can absorb (Inman-Bamber *et al.*, 1993). This level of stress (coded SWDF1 in Figure 2) reduces photosynthetic activity thereby directly affecting biomass and sucrose accumulation. The CANEGRO model also gauges incipient water stress (coded SWDF2 in Figure 2) which is assumed to occur when soil and root water is less than twice the atmospheric demand (McGlinchey *et al.*, 1995). This represents the first stage of crop water stress which restricts cell expansion and the production of new leaf and stalk tissue. Transpiration is decreased from the potential rate for different levels of water stress (SWDF2 and SWDF1) and for stages of incomplete canopy as determined from estimates of leaf area index (LAI).

Crop development is modelled in CANEGRO according to the manner in which the crop is understood to interact with the carbon, energy and water balances. A detailed canopy routine calculates LAI and the height of the growing crop (McGlinchey *et al.*, 1995), while root development is simulated in terms of rooting depth, total root dry mass and root distribution within the soil profile (Van Antwerpen *et al.*, 1993). Research into the development of ratoon crops has received precedence over plant crops because about 90% of the sugarcane area in South Africa produces ratoon crops at any one time. No subroutine is therefore available to describe the development of sugarcane plant crops (Inman-Bamber, 1994a). For simulations of plant crops a constant period of 21 days is allowed from planting to germination and thereafter crop development is simulated as for NCo376 ratoon crops. The germination period is too long under ideal conditions and causes the model to underestimate crop light interception of plant crops (hence growth and development) up to the stage of full canopy (Inman-Bamber and Thompson, 1989).

Canopy development and LAI are determined by the simulated daily rates of tillering, leaf appearance, leaf extension and the size of each leaf. Thermal time (°C d) is used to predict the rates of these processes (Inman-Bamber, 1994a). The base temperature for tillering is 16 °C and stalk populations peak at about 500 °C d after ratooning and stabilize at about half the peak stalk population, after about 1200 °C d. Full canopy is assumed when 70% of the photosynthetically active radiation is intercepted and is associated with the onset of rapid tiller mortality immediately after the peak stalk population is reached. The base temperature for leaf emergence is 10 °C and two distinct growth stages are used to model this process (Inman-Bamber, 1994a). The phyllochron interval for each of the first 14 leaves is 109 °C d and 169 °C d per leaf thereafter. Daily leaf extension rate is calculated for each leaf according to thermal time in relation to levels of water stress which restrict cell expansion (SWDF2). Leaf area is calculated using polynomial functions for leaf width and length which in turn depend on the leaf extension rate (Inman-Bamber, 1991b). The maximum final leaf area increases linearly in the order of leaf emergence up to about 400 – 420 cm^2 for leaf 16 and is fairly constant thereafter (Inman-Bamber, 1994a).

The total number of green leaves per stalk is allowed to vary between 3 and 11 in CANEGRO and depends on the amount of soil TAW. Leaf senescence is related to leaf emergence and is accelerated during periods of water stress (Inman-Bamber, 1994a). The LAI for a fully canopied crop is allowed to vary between 2 and 4.5, depending on soil water content (Inman-Bamber, 1991b). Crop height is determined as a dynamic function of plant extension rate. The base temperature for extension growth is 10 °C and the rate of extension is assumed to be constant to a maximum height of 3 m since it is defined by a linear function (Singels, 1998, personal communication [8]). The rate of extension is reduced according to water stress (SWDF2).

Crop height and LAI are used in the energy balance to determine daily atmospheric evaporative demand. For the partitioning of potential total evaporation between transpiration and soil evaporation, the fraction of soil evaporation is calculated using soil water and a simulated LAI that includes senesced sugarcane leaves because these continue to shade the soil for a long time (Inman-Bamber, 1991b). The daily LAI (green leaves) is also used in the carbon balance where daily crop light interception for photosynthesis is modelled as a function of LAI (Figure 2), and canopy factors are included for water interception (rain and irrigation) in the water balance.

Root growth is simulated as a function of photosynthate allocation within the plant (Inman-Bamber and Thompson, 1989). The simulated fraction of total plant biomass in the roots decreases with increasing plant age (increasing total biomass) but is always greater than 12% of total plant dry mass (Inman-Bamber, 1991b, based on data of Van Dillewijn, 1952). Total rooting depth and root distribution vary from soil to soil and are defined empirically according to the roots' "affinity" for the defined soil layers of the profile (Inman-Bamber, 1991b). Root water use

---

[8] Dr A Singels, Head, Agronomy Department
South African Sugar Association Experiment Station,
Private Bag X02, Mount Edgecombe, 4300

is calculated with the Ritchie equation of the CERES crop models (Jones and Kiniry, 1986). The maximum water uptake allowed is 0.030 cm^3 cm^{-1} $root^{-1}$ d^{-1} and the root mass:length ratio used in CANEGRO is 300 cm g^{-1} (Van Antwerpen *et al.*, 1993). Root characteristics interact with the water balance to supply the crop with water, the rate of supply being limited either by the energy balance or the water balance, depending on the soil water content.

Model inputs include the following data:
Crop start date (plant / ratoon); crop harvest date; daily records of maximum and minimum temperatures, rainfall and irrigation (if applicable); daily records of either class A-pan evaporation or humidity, wind speed, radiation / sunshine hours; soil albedo, maximum soil evaporation, soil runoff category (according to three classes of cover), soil layers to total depth in 10 – 15 cm increments listing the master horizon, layer depth, drained lower limit, drained upper limit, root distribution, bulk density, clay%, silt% and the saturated hydraulic conductivity coefficient. The variety is assumed to be NCo376 and the crop is assumed to be a healthy ratoon, free of weeds and to have adequate nutrition.

The CANEGRO model can use meteorological data recorded by an automatic weather station to calculate the climatic parameters required to run the model (McGlinchey *et al.*, 1995). Where soil data are available the soil physical properties are usually not described in adequate detail for model inputs and modal profiles for a number of well documented soils are usually used in model runs (Inman-Bamber *et al.*, 1993). The model is intended for specialist applications in SASEX. A version of the CANEGRO model has been included in the DSSAT [9] crop growth computer models of the IBSNAT project (Tsuji *et al.*, 1994). The DSSAT shell has menu-driven routines which allow users to select and use any of the DSSAT components. The input file creation is simplified by using a database management system where the user enters the necessary data at various prompts. Because the model is essentially unchanged from the SASEX version, input data requirements remain highly detailed, but the system is simpler to use. The DSSAT CANEGRO version would therefore be the more appropriate one to consider for applications by third party users.

---

[9] Decision Support System for Agrotechnology Transfer

The CANEGRO model is driven by radiation (and temperature) and crop water use. The model does not provide for lodging and overestimates the water use of lodged crops. Neither frost damage nor low plant populations are accounted for (Inman-Bamber, 1991a). The assumption of a weed free crop, free of disease and optimally supplied with nutrients, seldom applies to commercial crops. These, and many other factors, including management ability, reduce commercial yields from the radiation-and-water limited yield potential predicted by CANEGRO (Inman-Bamber, 1995). It is probably for these reasons that Inman-Bamber *et al.* (1993) found that the model was able to predict commercial yields at a field level reasonably well when water was limiting, i.e. water supply was more yield-limiting than the "management factor" limited yield, but not when the water supply was adequate or abundant.

The model is a powerful tool for applied scientists because most of the "management" factors which potentially limit yields can be controlled in the research environment and sophisticated measurements of model input parameters can be made. Although the CANEGRO model has been applied at a commercial field scale to predict yields (albeit using modal norms for the soil input parameters) (Inman-Bamber *et al.*, 1993), the model is probably more useful for determining crop water use in irrigation schedules (McGlinchey *et al.*, 1995; McGlinchey and Inman-Bamber, 1996; Singels *et al.*, 1998), for determining target production levels (yield benchmarks) (Inman-Bamber, 1995; Inman-Bamber *et al.*, 1998) and for diagnostic purposes when attempting to explain why a certain yield was achieved (Hellmann, 1993).

A limitation to the application of the model to the Midlands is that CANEGRO applies only to ratoon crops of variety NCo376 which has been replaced by superior-yielding varieties, mainly N12, N16 and N22. While it is unlikely that the basic photosynthetic efficiencies of these varieties will be superior to NCo376, significant differences in leaf extension rates, final leaf area and rates of stomatal closure in response to water stress have been found for NCo376 and N12 (Inman-Bamber, 1994a) which should account for some differences in varietal yield potential.

The expertise gained in constructing a complex simulation model often enables the modeller to break it into simpler models geared to specific user requirements. Two yield models have been derived from the CANEGRO model in this way, viz. the IRRICANE model and the South

African radiation limited sucrose yield model.

2.3.5.1 *The IRRICANE model*

IRRICANE (Singels *et al.*, 1998) has been developed from CANEGRO for irrigation scheduling of sugarcane. This relatively simple model is driven by total evaporation and a soil water balance. The initial model set-up requires that the following information be specified for each irrigation management unit:

a soil profile total available water (TAW);
b lower and upper limits of the target soil water content;
c crop starting date;
d initial soil profile water content; and
e type of irrigation system.

Once these parameters have been defined, only the dates and amounts of irrigation water applied need to be supplied by the irrigator and the scheduling procedure is fully automated.

Reference sugarcane evaporation is calculated using automatic weather station data with the Penman-Monteith equation according to the method of McGlinchey *et al.* (1995). A water balance is maintained for a single layered soil profile accounting for effective water inputs, runoff, deep drainage and total evaporation (Singels *et al.*, 1998). Irrigation and rain water inputs are decreased by the estimated amount of water intercepted by the crop canopy. This potential interception is calculated as the product of canopy cover and 5 mm water. Runoff is simulated as all the water in excess of the saturated available water content, which is assumed to be 200% of TAW. Deep drainage is calculated assuming a drainage rate of 40% d^{-1} of the soil water in excess of TAW. Total evaporation is calculated as the sum of soil evaporation and sugarcane transpiration. Soil evaporation is derived using a function of reference sugarcane evaporation, days elapsed since a soil wetting event, according to the method of De Jager *et al.* (1987) and relative canopy cover. Relative canopy cover is calculated using a logistic equation of days after emergence, the parameters of the equation being derived from regressions using LAI data from CANEGRO simulations for annual, water-stress-free crops. Sugarcane evaporation is calculated using the relative canopy cover, a relative water stress index and reference sugarcane evaporation.

Model outputs include irrigation advice for the next 14 days, a daily water balance and estimates of current and future yield. Cane yield is predicted by an empirical equation based on cumulative transpiration as follows:

$$\text{Cane yield (TCH)} = -30.735 + 0.188\,x - 4\,(10^{-5})\,x^{2} \quad \ldots\ldots\ldots\ldots \quad (4)$$

where x = Cumulative sugarcane transpiration (mm)

$R^2 = 0.96$

The yield equation was developed by fitting a second order polynomial function to simulated stalk dry matter yields and simulated cumulative transpiration from CANEGRO predictions at four locations for four different annual crop cycles over 20 years (1977 – 1996) (Singels, 1998, personal communication [10]). The high coefficient of multiple determination therefore means that yields predicted by this simple equation are very similar to those predicted by the complex CANEGRO model, implying that under the specific irrigated cropping conditions for which IRRICANE was developed, there is little to be gained from using CANEGRO in preference to IRRICANE. The limited input data requirements of IRRICANE compared with CANEGRO clearly favour the use of the former model for irrigated sugarcane. The goodness-of-fit statistics do not, however, indicate how accurately the model is able to predict yield responses using real data sets and additional model testing is essential to demonstrate the adequacy of model performance

Besides the irrigation scheduling and yield prediction functions of IRRICANE, it can also be used as a strategy evaluation tool by scheduling irrigation automatically for a "season" according to various pre-defined rules and limits. These include irrigation cycle and water limitations, various soil water depletion and refill levels and using recent or long term mean daily crop water

---

[10] Dr A Singels, Head, Agronomy Department
South African Sugar Association Experiment Station,
Private Bag X02, Mount Edgecombe, 4300

use. Modelled yields for the various management options can then be used to select the most appropriate management strategy for that season.

Initial testing of IRRICANE has shown promising results as the model estimates soil water contents with an accuracy similar to CANEGRO, to within 3% to 6% of measured profile water content (Singels *et al.*, 1998). This is well within the accuracy of irrigation management. Although the model is still very new, it is being used on a limited scale in the sugarcane industry. As data on model performance accumulates it will be possible to establish whether the model is adequate for irrigation scheduling or whether additional calibration is required. Because of the largely empirical nature of IRRICANE, it is subject to the general limitations applicable to yield models which are discussed in Section 2.2.1. When correctly calibrated, IRRICANE and similar models have the potential of becoming simple yet extremely powerful crop management tools.

2.3.5.2 *South African radiation limited sucrose yield model*

At the highest level of production possible a crop's entire growth requirements are met optimally and yield is limited by radiation and temperature only. Inman-Bamber (1995) used the CANEGRO model to simulate radiation limited sucrose yield potentials for annual sugarcane crops at 32 locations in the South African sugarcane belt, ranging in latitude from 25.3 °S to 30.8 °S and altitude from 19 m to 1067 m a.s.l. Latitude and altitude accounted for 96% of the variation in CANEGRO-simulated sucrose yield using a multiple regression equation as follows:

$$\text{Radiation limited sucrose yield (TSH annum}^{-1}) = 0.18 + 2.9\,\theta - 0.0776\,\theta^2 - 0.0058\,A \quad ....(5)$$

where θ = Latitude (radians South)

A = Altitude (m a.s.l.)

SE = 0.72 TSH annum^{-1}

n = 32

R^2 = 0.96

Radiation varies mainly with latitude, decreasing with increasing distance from the equator. The amount of radiation intercepted by the crop depends on the development of the canopy which, in turn, largely depends on heat units when water and nutrients are adequate. Heat units vary mainly with altitude and, to a lesser extent, with latitude.

The yield equation is very useful for determining ceiling levels of sugar production without having to run the complex growth model repeatedly. Unfortunately the model cannot realistically be applied to Midlands rainfed sugarcane because the crop is not harvested annually, resulting in substantially improved radiation use efficiencies and hence potentially greater sucrose yields than predicted. Water is also more yield limiting than radiation in most seasons under rainfed conditions, although radiation limited yields may be approached in exceptionally wet years. It follows that rainfed yields are generally limited at a lower production level (water limited yield potential). The difference between potential evaporation and rainfall gives an indication of the extent to which sugarcane growth is limited by water, depending mainly on locality and soil type (Inman-Bamber, 1995).

2.3.6 Sugarcane yield benchmarks

Yield benchmarks for sugarcane may provide managers with production goals and be used to identify low performing (as opposed to low producing) fields, allowing for further investigations of yield limiting factors (Inman-Bamber, 1995). These standards can be used to help focus management attention on important production and productivity issues. Crop production systems are difficult to benchmark because of the unpredictable nature of climate and the large number of management and environmental factors impacting on the various plant growth processes (Inman-Bamber *et al.*, 1998).

Yield potential is defined by the radiation, temperature and the rate limits of physiological processes (Inman-Bamber *et al.*, 1998). If water is limiting, the attainable yield is lower than the potential yield. Similarly, if crop nutrients are limited or weeds, disease and pests are prevalent, actual yields will be lower than attainable yields. Year to year variation in potential yield as a result of radiation and temperature variation is low in the South African sugarcane belt but the range in attainable yields is large in rainfed production areas because of the high variability in rainfall (Inman-Bamber *et al.*, 1998).

Three general approaches to yield benchmarking have been used in the South African sugar industry. These include the use of crop models to predict potential and attainable yields, yield indices as a relative comparison of production and productivity between different yield observations, and physical sampling of the crop to estimate cane yields.

2.3.6.1 *Yield predictions*

The first attempts to benchmark South African sugarcane yield potential were made by Glover (1972) for Mount Edgecombe, based on simple derivations of photosynthesis, photosynthetic efficiency and respiration of sugarcane. More recent efforts have involved a systems approach using the CANEGRO model (Inman-Bamber, 1995). One difficulty with yield benchmarking is that an absolute yield potential cannot be defined for a given location because of the influence of harvest age, crop cycle season and variety (Inman-Bamber, 1995). For this reason dynamic models such as CANEGRO are required to accommodate most of these variables (Inman-Bamber *et al.*, 1998). Attainable yields are similar to potential yields when good irrigation is practised (Inman-Bamber *et al.*, 1998). Under rainfed conditions in South Africa Inman-Bamber (1995) estimated attainable yields to be 10% to 90% lower than yield potentials, depending on locality and soil type.

Inman-Bamber *et al.* (1998) found that CANEGRO-simulated mean attainable yields were often similar to top actual yields during dry seasons in South Africa, but during wetter years actual yields were much lower than simulated attainable yields. This trend is consistent with the findings of Inman-Bamber *et al.* (1993) where CANEGRO predicted commercial field yields reasonably well when water was limiting, but when the water supply was adequate or abundant, actual yields were much lower than those predicted. A reasonable explanation for this is that in dry seasons rainfall limits yield sufficiently to mask other limitations, but in wet seasons the rainfall limitation is largely removed and management factors become limiting.

The South African sugarcane yield models reviewed from the available literature all aim to predict attainable yields, driven primarily by estimates of crop water use. Since these yields are difficult to achieve under commercial conditions, it has become customary to reduce predicted attainable yields by an arbitrary management factor to estimate realistic actual yield-targets as a fraction of those that are attainable. Attainable yields are generally achieved only under exceptional management in experimental trials and are reduced by 30% to estimate the expected "average" grower performance (Thompson, 1977; Thompson and Harding, 1986). This 30% management factor for the Thompson (1976) yield models is also applied to CANEGRO simulated yields for commercial predictions (Inman-Bamber *et al.*, 1993) and the ACRU sugarcane yield submodel (Schulze *et al.*, 1995). Smith (1994) uses a more detailed management factor concept, reducing predicted attainable yields by 10%, 20%, 30% and 40% to estimate the likely actual yields for excellent, very good, good and average management respectively. While this approach may improve the predictive accuracy of the models when using commercial data, it does not help focus management attention on important production parameters.

2.3.6.2 *Yield indices*

Commercial sugarcane yields are generally substantially lower than attainable yields (Inman-Bamber *et al.*, 1998) and for this reason it is customary to express yields as indices of various production parameters that are perceived to be yield limiting. Liebig's law of the minimum suggests that there is always some factor limiting actual yield. It follows that yield needs to be expressed in terms of the limiting factor to allow for a meaningful interpretation of crop yield responses to inputs (Inman-Bamber, 1995). Penning de Vries and Van Laar (1982) proposed a hierarchy of production systems based on factors that limit yield. At production level 1 yield was limited only by radiation and temperature (potential yield); at production level 2 growth was limited by a shortage of water (attainable yield); other lower production levels involved shortages of both water and nutrients.

For cane yields (expressed as TCH or TCHM) to be compared meaningfully, it is necessary to consider the season in which the crop cycle was started and completed because crops harvested at different times of the year, and growing through different seasons, have vastly different growth potentials (Lonsdale and Gosnell, 1975; Landrey *et al.*, 1981; Sweet and Patel, 1985; Hellmann, 1988; Inman-Bamber, 1991a; Hellmann, 1993; Inman-Bamber, 1995). Two static modelling approaches have been developed in the South African sugar industry for this purpose. Sweet and Patel (1985) and Tobin and Ellis (1988) derived factors to weight the relative seasonal influence on cane yield for yield comparisons using small data sets. These correction factors apply only to the data from which they were derived and are subject to annual variation. Alternatively, Landrey *et al.* (1981) and Hellmann (1988; 1993) grouped crop cycles of similar season and duration and treated these separately when working with large data sets. Since attainable yields are not constant between years because of climatic variations, Hoekstra (1976) developed a climatic correction factor to allow for a fair evaluation of relative field performance when comparing cane yields achieved in different years.

Sucrose yields may be compared on the basis of age corrected relative sucrose yields alone because almost all the within-season variation of sucrose content (Van Dillewijn, 1952) is accounted for when sugarcane sucrose content is expressed as an index relative to the expected mill average sucrose content for the entire season. For this reason relative sucrose yield is a more reliable index of performance than actual sucrose yield when comparing yields from fields harvested in different seasons (Hellmann, 1993).

South African sugarcane yields have been expressed relative to individual plant nutrients (N, P and K) in a production function format (Prins *et al.*, 1997) to evaluate production in terms of input costs. The index of TCH 100 mm^{-1} (effective) rainfall is often used because water frequently limits attainable yield in rainfed production systems. Although meaningful relations between yield and sugarcane water use have been derived (Thompson, 1976; Kingston, 1994), the yield:water use relation may become misleading when other production factors are more yield limiting than water (Inman-Bamber *et al.*, 1993; Inman-Bamber *et al.*, 1998). For this reason care needs to be exercised when evaluating productivity using yield indices because relations may be masked by other more yield-limiting factors (Inman-Bamber *et al.*, 1998). This is especially important when comparisons are made between different seasons.

2.3.6.3 *Yield sampling*

The yield of a standing crop of sugarcane may be estimated by physically sampling and weighing the plant material. This technique is used in agronomic trials to evaluate relative crop performance between treatments and by extension officers when comparing the performance of newly released sugarcane varieties between different growers. Growers themselves may use this technique to estimate the size of the standing crop shortly before harvesting.

The technique involves counting the number of stalks along a 1 m row length. This number of stalks is then removed from the field, either harvesting all the stalks along a 1 m row length, or by selecting representative stalks from the entire stalk population in the field. The stalks are topped at the natural breaking point and weighed to determine the stalk yield per metre row length. This yield is multiplied by the total row length planted per hectare (calculated from the mean row spacing) to determine the cane yield of a standing crop. Yield sampling is a very simple technique and only a tape measure and scale are required. It is important to determine the field row spacing and stalk population accurately. It is also critical to ensure that the harvested stalks are representative of the field to avoid sampling bias. Unrepresentative sampling causes this technique to fail dismally.

2.4 SCALE CONSIDERATIONS AND MODELLING

The spatial scale to which models apply is of critical importance in determining which parameters are used to explain yield. When yield differences between continents and regions are to be explained, crop water use should be an important explanatory variable – especially where the observed range in rainfall is large. Because water is considered to be the single most limiting resource in rainfed sugarcane production in South Africa (Thompson, 1976; Inman-Bamber, 1995) differences in crop water use should also be highly correlated with yield at the district level. It is at this scale of modelling that the Thompson yield equations, and the models based on the Thompson (1976) study, have been applied most successfully (Thompson, 1977; Thompson and Harding, 1986; Smith, 1994; Schulze *et al.*, 1995).

When the scale of modelling is reduced to a commercial estate-and-field level, yields are unlikely to be affected by only a single factor (Hellmann *et al.*, 1995). At this scale the Thompson (1976) – type models which depend only on crop water use would be expected to perform poorly because the rainfall between adjacent fields is likely to be very similar, if such records are available, and only weakly correlated with yield as a result. A further complication arises from the fact that since rainfall cannot be reliably predicted, it is not a suitable explanatory variable for predictive models. It would also be unreasonable to ignore the vast natural variations of site quality occurring between fields at this scale of modelling (Hellmann, 1988). Where yield differences do occur between fields additional explanatory variables should therefore be sought to account for the observed variation in yields (Hellmann *et al.*, 1995).

Empirical models rely on the statistical relations between independent variables and yield. These relations may change, depending on the scale of modelling. It is therefore essential to ensure that these models are not used under conditions different to those for which they were derived. Although the Doorenbos and Kassam yield equation may be appropriate for use at a regional scale, model resolution is entirely dependent on the spatial scale of the estimate of maximum harvested yield (Y_{max}). When the agro-ecological zone method (Kassam, 1977) is used to estimate Y_{max}, the model should be used only for investigations at a continental scale. At the other extreme, crop models such as CANEGRO simulate growth processes at the individual plant level and yield predictions are based on a "scaling up" of the expected individual plant performance. The spatial resolution of this model is consequently limited primarily by the resolution of input data. While the CANEGRO model (Inman-Bamber, 1991b) has been used selectively to predict attainable yields at a field scale (Inman-Bamber *et al.*, 1993), no suitable yield model is available for a general estate-and-field scale application in South Africa and no empirical models based on agricultural resources and crop management decisions such as those of Bouldin (1969, cited by Early, 1980) Alvarez *et al.* (1982), and Phillips *et al.* (1989) have been developed.

Since the primary influences on sugarcane growth (solar radiation, temperature, water supply and nutrient supply) are difficult to measure at a commercial field scale and are generally not available from field records, proxies correlated with these factors are sought, e.g. latitude and altitude as proxies for radiation and temperature (Inman-Bamber, 1995). Factors that make a potentially important contribution to the observed variation in sugarcane yields must be included in the modelling process. These are generally selected on the basis of an understanding of physical and biological processes underlying site productivity, the findings of others, and on practical considerations such as ease of measurement and procurement costs of the data. While it is possible to measure and describe a large number of variables for each field, screening the data set for suitable factors (Grey, 1983), it is imperative that the variables included in the model are easily accessible to the model user (Keig *et al.*, 1991).

2.5 GEOGRAPHIC INFORMATION SYSTEMS IN MODELLING

Variables that are included in models to account for yield differences between fields may be related to the estate resource base using a geographic information system (GIS). The reason for using a GIS is its ability to combine spatial objects (polygon, line and point data) from maps with attributes of these features (Papajorgji *et al.*, 1994) giving a spatial dimension to the data so that features and attributes at one site can be compared with information at the same site or at other locations. This assists in determining explanatory variables and helps identify possible geographic reasons for yield variations (Hellmann *et al.*, 1995).

Within a GIS, sugarcane fields are presented as polygons referenced in space, each with a unique identifier (field number). The GIS can link to databases containing attribute data for these fields, e.g. agronomic records, provided the numbers identifying each field in the database correspond exactly to those used in the GIS. This has been done in the local sugar industry using data from the SASA Fertilizer Advisory Service (Schroeder *et al.*, 1994) and the Field Record System (Hellmann *et al.*, 1995). Fields with similar resources and management histories can be identified using a GIS and the yields compared, based on the performance of analogous sites or evaluated more thoroughly using a sugarcane growth model.

The integration of crop models with GIS and spatial analysis applications has many advantages (Johnson, 1994). For data input, GIS can be used to provide parameters such as longitude, latitude, altitude, aspect, slope and soil type (Tim *et al.*, 1992) and model output can be aggregated across space, displaying a thematic map of crop yield (Papajorgji *et al.*, 1994). Managers of the areas represented on a map can often recall important observations or decisions that explain apparent anomalies (Yost and Pandutama, 1990). In addition, the model output may be analysed by interrogation of the GIS to find all areas that match a set of user-specified conditions, providing a valuable tool for decision makers.

According to Tim *et al.* (1992) the potential of linking GIS and models has been recognised for several years and can be achieved at three levels of integration viz. *ad hoc* linkage, partial linkage and full integration. The *ad hoc* approach is the most commonly used form of GIS and model linkage because generally no software modifications are required. Both the GIS and the model function separately, simply exchanging files or being linked by file exchange protocols. The model reads certain input data from the GIS and produces output in a format that allows processing and display with the GIS. Partial linkage is achieved where the model is written in a GIS application language, e.g. ARCINFO Arc Macro Language, or developed using tool kits that provide GIS functionality in addition to interface components for modelling. Fully integrated systems are the most expensive to develop. In this approach the model and the GIS are developed in close interaction, such that the model becomes one of the many analytical functions within the GIS.

The IBSNAT models have been combined with GIS because of the need to expand the application of crop growth models to large areas in which both the environment and management practises vary substantially. Initially linkage was achieved on an *ad hoc* basis, but the new versions of the models are partially integrated with ARCINFO (Papajorgji *et al.*, 1994). Preliminary work with the CANEGRO model linked to a GIS on an *ad hoc* basis has resulted in the production of a map of radiation-limited sugarcane yield (Wallace, 1996).

One major problem encountered when modelling sugarcane production using a GIS is the degree to which reliable, location-specific estimates of key input parameters can be made (Johnson, 1994). Most resource data are collected as point samples but for a raster GIS each cell needs an own value, and with vector-based systems discrete boundaries need to be defined between mapping units which are often gradual in nature. Interpolation and averaging are therefore used extensively to define values and delineate boundaries, causing inaccuracies in the data. For this reason uncertainty (conceptional, locational, descriptive and meta-uncertainty), scale and data availability are a concern where GIS is used in combination with modelling (Johnson, 1994).

All input data need to be available and specified at a spatial scale commensurate with the level of generalization of the model to achieve a sensible model output (Burrough *et al.*, 1988). Traditional arable management has treated fields uniformly and has tended to ignore the inherent spatial variability within fields (Blackmore, 1994). Agricultural inputs are applied uniformly across a field and records of yield at a field level assume uniform production from the entire area. It follows that where sugarcane production is modelled using field records, spatial analysis of resources beyond the field polygon level is meaningless because differential responses to the resources have not been recorded.

It is reasonable to expect variation in some of the natural resources at a scale within field management units. It follows that when field records are analysed, impure "mapping units" of some natural resources, e.g. soil type, may distort yield relations. For this reason it is important to determine those resources / factors which are highly correlated with yield to enable a sensible delineation of management units. Improved crop management and greater modelling accuracies could be achieved in the long-term by recording yields from discrete resource combinations – necessitating field boundaries to coincide with important resource changes. Alternatively, precision farming technology can be used for site-specific management using detailed yield maps (where sugarcane is harvested mechanically), or near infra-red aerial photographs of fields (where sugarcane is harvested manually) (Meyer *et al.*, 1997), to determine yields beyond the field level. These crop performance data can be related to resource differences within the field, e.g. localized soil fertility differences, and could be used to determine and apply appropriate quantities of agricultural inputs according to the production potential of different field sections using a global positioning system (GPS) controlled applicator (Blackmore, 1994). Although this technology is not yet commercially available in the South African sugar industry it is being evaluated (Meyer *et al.*, 1997).

Both field record analysis and crop modelling have received research attention in the South African sugar industry. The yield models that have been developed aim to predict either potential or attainable yields and therefore do not rely on input data from commercial field records comprising data of actual yields. Mondi require an objective method for predicting actual sugarcane yields for economic planning and a means of evaluating their performance within and between seasons. Ideally this should be achieved using data already available for their sugarcane enterprise. No suitable yield model is available for applications based information from field records and it would be preferable to avoid collecting the detailed data required to run the CANEGRO model (Section 2.3.5). For this reason the suitability of agricultural resource and production data available for Mondi's Midlands rainfed sugarcane enterprise was evaluated for yield modelling applications.

3 AGRICULTURAL RESOURCE AND PRODUCTION DATA

Natural resource data for the Mondi sugarcane enterprise in the Midlands were obtained from the Mondi GIS database, the bioresource units (BRUs) of KwaZulu-Natal (Camp, 1999) and a limited amount of field work. Agronomic production data were obtained exclusively from Mondi sugarcane field records. In order to establish the suitability of the Mondi records for statistical analysis, the field record system was reviewed to establish the source and details of each record, field records were updated and all available data were confirmed and corrected as far as possible. Since the agronomic records are available at a field level it was important to specify other data at this scale wherever possible (Burrough *et al.*, 1988).

Mondi out-sourced the maintenance of sugarcane field records on their Midlands estates to Spencer Holley Agronomic Services in 1986. This organisation operates as agronomic consultants in KwaZulu-Natal, specializing in sugarcane, maize, pasture (dairy) and avocado production. Spencer Holley consolidated all the Mondi / NTE [11] sugarcane field records. Although the earliest of these records date back to 1976, comprehensive records are available only from 1985 to 1997 (13 years).

The field records were extracted from an essentially paper-based record system for which the computer software package QUATTRO PRO was used for data presentation purposes only. Estate managers were responsible for the initial recording of all data (rainfall records, field operations and crop husbandry) on worksheets in the Spencer Holley estate record books. These records provide a chronology of enterprise management because the data were recorded as field operations were completed. Spencer Holley typed these handwritten estate records into QUATTRO PRO for each season and also used the data to maintain a record of the management history for each field. As a result, the field record system contains much duplicated information. The spreadsheet format was structured in document form to allow for easy block-printing, relevant footnotes for each page being placed within the print block. Although this data structure

---

[11] Natal Tanning Extract company, now incorporated in Mondi Forests, a division of Mondi Ltd.

allowed for the production of neat documents, a database structure was not considered and consequently electronic processing and manipulation of the data for this study was tedious. Within the spreadsheets only a limited amount of space was allocated to each field and the oldest records were deleted sequentially to make space for new entries. Such a practice defeats the purpose of keeping long-term records. All Mondi records discarded in this manner were recovered from hard copies of the data. Some of the improvements suggested as a result of this study have recently been incorporated in the extensively revised Spencer Holley record system.

It was necessary to check all the data thoroughly, correcting recording errors and inconsistencies as far as possible. In addition, the database structure had to be reworked to present the records in a form that could be manipulated electronically and used for data input in statistical analysis and GIS applications. This involved the consolidation of all the field data from separate record sections and the incorporation of footnote data into appropriate columns. The process was time consuming and highlighted the importance of using a database structure appropriate for the intended purposes when data collection begins. All duplicated records were used to check the data for recording accuracy. Wherever possible, data discrepancies and inconsistencies were corrected with the assistance of the estate managers responsible for the various sugarcane sections.

3.1 LOCALITY

Mondi produce sugarcane at Kranskop (Salem, Sutherlands and Elandsvlei estates), Umvoti (Canema estate) and Richmond (Uplands and Greenhill estates) in the Midlands. Insufficient natural resource data were available for the Elandsvlei estate and therefore this sugarcane was excluded from the study. Since the estates are large and comprise numerous sub-divisions, only the property names of those sub-divisions producing some sugarcane are presented in Table 2. The general locality of the estates is shown in Figure 3.

3.1.1 Kranskop

Salem and Sutherlands are located near Kranskop, about 85 km north-east of Pietermaritzburg (Figure 3). Salem is situated about 57 km inland from the Indian Ocean and Sutherlands approximately 53 km from the Indian Ocean. The 31 sugarcane fields at Salem lie within an area of 5.5 km^2, at approximately 29.01 °S, 30.93 °E while the 24 sugarcane fields on Sutherlands are concentrated within an area of 2.4 km^2 at approximately 29.07 °S, 30.94 °E. All the sugarcane fields lie within the Moist Midlands Mistbelt bioresource group (BRG 5) (Camp, 1999) and the entire area falls within BRU 212 – Kranskop (Appendix 1).

Table 2: Property names of Mondi land holdings on which sugarcane is produced in the Midlands (excluding Elandsvlei).

District	Estate	Property Name
Kranskop	Salem	Sub. 10 of Elands Kop 1377 Sub. 18 (of 13) of Elands Kop 1377 Sub. 1 of Driefontein 10147 Rem. of Driefontein 10147 Sub. 7 of Scottsdale 1961
	Sutherlands	Sub. 2 of Jammer Daal 1941
Umvoti	Canema	Rem. of Burleigh 1863 Sub. 1 of Canema 1881 Sub. 2 of Mistley 2034 Sub. 3 (of 1) of Mistley 2034 Sub. 4 (of Lot 4) of Mistley 2034 Rem. of Sub. 1 of Groene Kop 2048 Sub. 3 of Groene Kop 2048 Rem. of Sub. 4 (of Lot 1) of Groene Kop 2048 Sub. 5 (of 4) of Groene Kop 2048 Rem. of Groene Kop 2048
Richmond	Uplands	Rem. of sub. 168 of Harmony 1472 Sub. 177 (of 169) of Harmony 1472 Sub. 86 of Harmony 1472 Sub. 87 of Harmony 1472 Sub. 88 of Harmony 1472 Sub. 90 of Harmony 1472 Sub. 91 of Harmony 1472 Sub. 92 of Harmony 1472 Sub. 93 of Harmony 1472 Sub. 94 of Harmony 1472 Sub. 95 of Harmony 1472 Sub. 96 of Harmony 1472 Sub. 98 of Harmony 1472
	Greenhill	Sub. 436 of Beaulieu Estate 1412 Sub. 64 of Kruys Fontein and Weltevreden 826

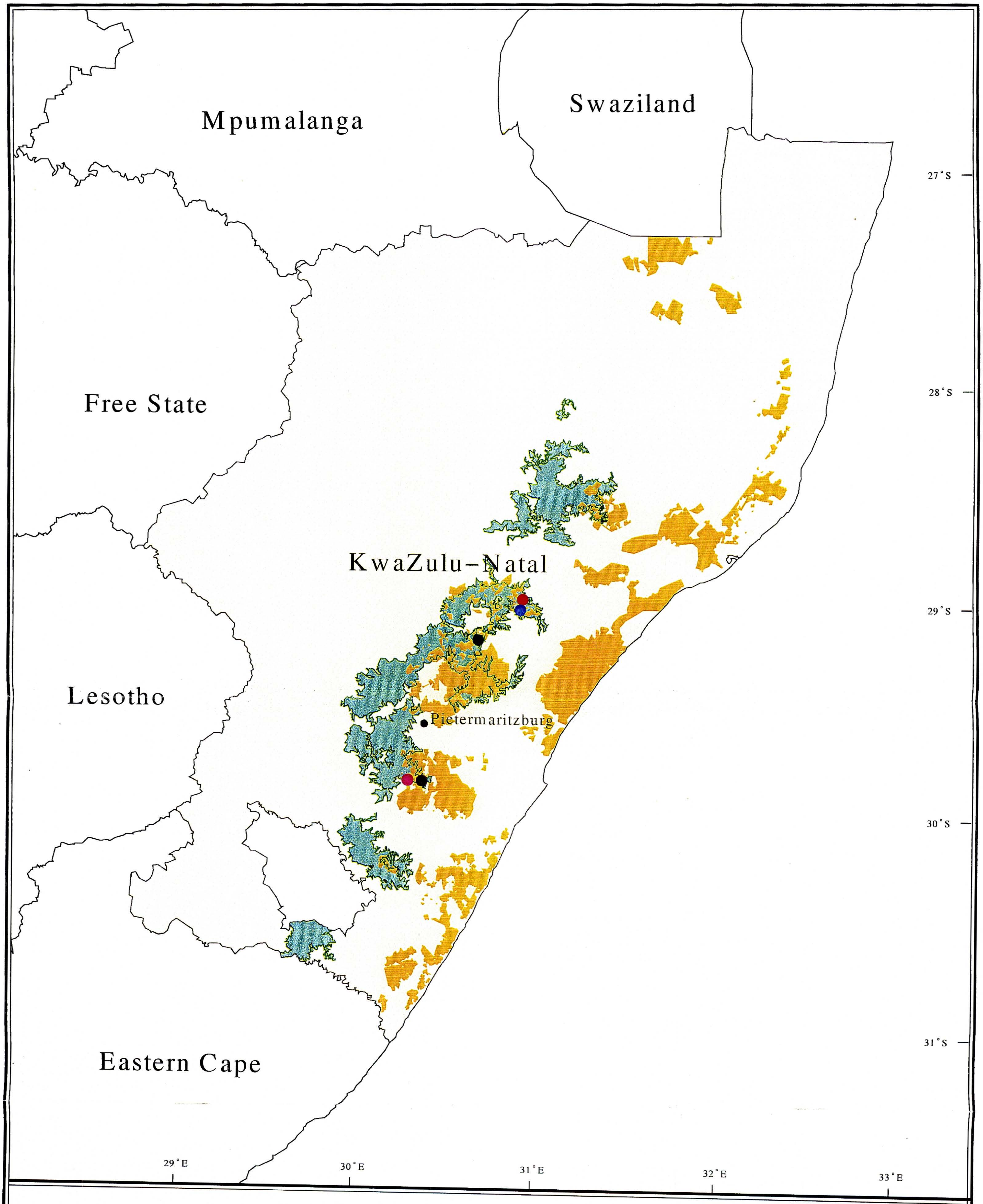


Figure 3 : Map of KwaZulu-Natal showing the location of Mondi estates on which sugarcane is produced in the Midlands (excluding Elandsvlei).

3.1.2 Umvoti

Canema is located near Seven Oaks, about 50 km north-east of Pietermaritzburg (Figure 3). The estate lies about 64 km inland from the Indian Ocean. Sugarcane is produced on 34 fields within an area of 10.8 km^2 at approximately 29.20 °S, 30.68 °E. Although the sugarcane fields all lie within BRG 5 (Camp, 1999) most of the sugarcane area falls into BRU 427 – Wartburg while a smaller portion lies in BRU 428 – Harden Heights (Appendix 1).

3.1.3 Richmond

Uplands and Greenhill are located near Richmond, approximately 25 km south of Pietermaritzburg (Figure 3). The estates lie about 55 km inland from the Indian Ocean, Greenhill being situated about 3 km west of Uplands. Most of the Richmond sugarcane is produced on Uplands, the 46 sugarcane fields lying within an area of about 9.0 km^2, at approximately 29.87 °S, 30.37 °E. Four additional fields (AS21, AS22, AS23 and AS24) were planted to sugarcane during 1997 and 1998 but these were excluded from the study because no yield data were available. On Greenhill sugarcane is produced on 11 fields within an area of about 1.2 km^2, at approximately 29.86 °S, 30.29 °E. The Richmond sugarcane fields lie predominantly in BRG 5 (Camp, 1999) and only a portion of fields AS09 and AS19 on Uplands are classified as BRG 3 (Moist Coast Hinterland Ngongoni Veld). The Moist Midlands Mistbelt fields fall into BRU 330 – Byrne while the Coastal Hinterland section on Uplands is in BRU 400 – Baynesfield (Appendix 1).

It follows that, in terms of general agricultural resources, this study focuses on Midlands sugarcane production in BRG 5 (Figure 3) and excludes the lower potential BRG 3 sugarcane regions of the KwaZulu-Natal Midlands region.

3.2 CLIMATE

Climate is an important determinant of yield potential (Inman-Bamber, 1995; Inman-Bamber *et al.*, 1998). Rainfall is the only climatic variable recorded on the Mondi estates. When other climatic data are required, these need to be obtained from weather stations. The names of weather stations with a similar climate, nearest each estate and with appropriate climatic data for the study period (1976 – 1997), are listed in Table 3.

Table 3: Details of weather stations nearest the estates with representative temperature, evaporation and radiation data for the study period (1976 – 1997) and the approximate distances from the estates. (Source: Computing Centre for Water Research)

District	Estate	Climatic Data	Weather Station	Co-ordinates		Altitude (m a.s.l.)	Distance from Estates (km)	Recording Period	Good Data (years)
				Lat. (°S)	Long. (°E)				
Kranskop	Salem and Sutherlands	Temperature and Evaporation	0302658 Hazyview Kranzkop	20.58	30.52	1120	9.3	1983 – 1991	3
			0270481 Waldeke Hermannsburg	29.01	30.47	1140	13.8	1976 – 1991	7
			0271038 Mapumulo	29.08	31.02	716	15.8	1977 – 1992	0
Umvoti	Canema	Temperature and Evaporation	0270189 Mispah – Greytown	29.04	30.37	1010	7.2	1976 – 1986	8
			0270194 Seven Oaks Ryhill	29.14	30.37	1066	6.3	1970 – 1992	8
			0270399 Mowbray	29.09	30.44	1167	7.5	1971 – 1974	1
Richmond	Uplands and Greenhill	Temperature and Evaporation	0239585 Baynesfield Estate	29.45	30.20	808	20.3	1975 – 1991	11
			0239563 Ivanhoe Richmond	29.53	30.19	860	3.4	1979 – 1982	2
			0239829 Tala Valley	29.49	30.20	692	10.4	1974 – 1977	0

Since the spatial variation of temperature and evaporation is much less variable than that of rainfall, the use of nearby "off farm" data is sufficiently reliable. Unfortunately the quality of the available data is poor and records are not available for the entire study period, especially from 1985 to 1997 for which comprehensive field records are available (Table 3). The closest weather stations with appropriate records are also located too far away from the estates, especially at Richmond (Table 3), to use their data with confidence. Schulze (1982) found that temperature correlated well with latitude and altitude and used these relations to produce medians of daily maximum and minimum temperatures, and median A-pan equivalent potential evaporation for each month at a 1' x 1' of a degree interval for southern Africa (Schulze and Maharaj, 1991). The BRU temperature and evaporation data (Appendix 1) are the average of these 1' × 1' gridded values within each unit. Although heat units are calculated from daily maximum and minimum temperatures, the BRU data (Appendix 1) were calculated using the monthly means of daily median maximum and minimum temperatures (Camp, 1999).

The daily / monthly median climatic data available for the relevant BRUs and 1' × 1' grids are useful for comparing the general climate of the study area, and the 1' × 1' gridded data probably reflect the actual climate of the estates more accurately than the data from the nearest weather stations. These data are, however, of no value for modelling sugarcane yields because when median values or other summary statistics are used, temperature and evaporation become a function of crop cycle starting month and age at harvest only. Consequently these data were not considered as predictors of sugarcane yield.

Radiation is a primary determinant of crop growth but few weather stations record solar radiation or daily sunshine hours, which can be used to calculate radiation (Thompson, 1986). The Computing Centre for Water Research have no collated records of radiation or sunshine hours. The seasonal sunshine hours recorded for the relevant BRUs (Appendix 1) are based on a subjective interpolation between sunshine-recording stations (Camp, 1999) and should therefore be used as a general guide only.

Latitude and altitude were specified for each estate as proxies for radiation since radiation limited yield is closely associated with these site characteristics (Inman-Bamber, 1995), as are temperature and evaporation (Schulze, 1982). These site characteristics remain constant from year to year and therefore have a similar effect to using medians of data, except that the values of latitude and field altitude are not incorporated in other variables such as crop cycles. The use of latitude and altitude as proxies for radiation and temperature is justified on the basis that they are easily determined and Inman-Bamber *et al.* (1998), using 25 years' climatic data, found the year to year variation in simulated potential (radiation and temperature limited) yield to be low in South Africa. Differences in mean CANEGRO-simulated potential yields of the best and the worst years of 25 were only 15 TCH (Inman-Bamber *et al.*, 1998).

3.2.1 Rain

In the Midlands water is considered to be the single most yield-limiting resource for rainfed sugarcane production (Thompson, 1976; Inman-Bamber, 1995). Spencer Holley maintained a record of estate rainfall using data provided by the sugarcane managers. The daily rainfall record of some estates seldom corresponded to the recorded monthly rainfall totals and other data summaries (annual totals, graphs, etc.). A closer investigation of the data revealed that a number of stations had been used to compile the data set – rainfall recorded at the managers' homesteads together with records from any one of the local estate offices.

For the purposes of this study, estate rainfall stations which were considered to have the most representative, and accurate records for the sugarcane producing areas of the estates were selected. Rainfall records from the Salem, Sutherlands, Mistley and Uplands estate offices were used for the sugarcane on Salem, Sutherlands, Canema and Uplands / Greenhill respectively (Appendix 2). A separate station was not used for the Greenhill sugarcane because estate managers agree that the overall climate and rainfall of this area is more similar to the climate on Uplands than that recorded at the Greenhill estate office.

Monthly rainfall data for the four selected stations were accessed at the ICFR [12]. Records were available for two periods, 1947 to 1960 and 1976 to 1994. Since the oldest sugarcane records date back to 1976, only the rainfall records from 1976 onwards were extracted. These records

---

[12] Institute for Commercial Forestry Research

were checked against Mondi's original paper copies where these were still available. A number of discrepancies were found, mainly where the handwriting was illegible. Where daily rainfall records were available for these months, totals were checked and corrected accordingly. Otherwise the monthly records for the surrounding Mondi stations were compared, using the entry which was more similar to these records. Mondi maintain an electronic record of monthly rainfall data from 1989 onwards for all their rainfall stations. The rainfall records from this database for the four selected stations were checked using those daily records which were still available on the estates. A number of records could not be checked because the daily rainfall records had been lost and these entries were assumed to be correct.

Besides errors of addition and transcription which were corrected as far as possible, additional problems with the estate rainfall records include the manner in which the data were recorded. Initially only whole millimetres of rain were recorded but in recent years many rainfall events have been recorded to the nearest 0.5 mm. The accumulation of half-millimetres over the year and the rounding-up of monthly rainfall totals has the potential to introduce "additional" rainfall when compared with the initial recording method. Although the total amount of rainfall recorded is considered to be accurate (clerks are responsible for the estate office rainfall records) the daily rainfall distribution is somewhat unreliable because rain falling over weekends was only recorded on Mondays. A similar situation applies to public holidays.

Schedules of monthly rainfall (to the nearest millimetre) were compiled for the four stations from January 1976 to December 1997 (Appendix 2). Statistics were calculated using the 22 years' data and these are presented in Table 4. On average, Kranskop is the wettest and Umvoti the driest of the three districts. Relevant monthly rainfall was related to each crop cycle in the field records. In addition to the rain total for each crop cycle and rain for the first 12 months' growth of each crop cycle, rain for the two months before the start of each crop cycle was calculated because of the association Panol (1974) found in the Phillippines between cane yield and rainfall for the eight weeks prior to the start of crop cycles. Similarly, the rain for the month before harvesting was calculated because of the relation Panol (1974) found between sucrose yield and rainfall during three weeks prior to harvesting.

Table 4: Mean monthly and mean annual rainfall for Salem, Sutherlands, Mistley and Uplands based on 22 years' data (1976 – 1997). Coefficients of variation (%) included in brackets.

District	Estate	Station	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Annual
Kranskop	Salem	Salem estate office	149 (49)	153 (64)	116 (51)	46 (56)	25 (97)	21 (127)	21 (135)	39 (68)	68 (164)	123 (36)	134 (52)	143 (49)	1039 (26)
	Sutherlands	Sutherlands estate office	163 (42)	155 (65)	122 (52)	51 (63)	28 (91)	24 (137)	21 (129)	41 (59)	73 (157)	127 (42)	137 (57)	151 (47)	1090 (24)
Umvoti	Canema	Mistley estate office	130 (47)	109 (50)	125 (40)	36 (58)	15 (87)	16 (141)	16 (157)	34 (92)	47 (107)	102 (42)	114 (45)	126 (38)	870 (22)
Richmond	Uplands / Greenhill	Uplands estate office	128 (41)	104 (55)	125 (60)	55 (43)	25 (73)	18 (119)	22 (189)	36 (93)	68 (144)	114 (38)	125 (62)	130 (62)	949 (27)

3.3 SOILS

Differences in soil water and nutrient storage / release characteristics are expected to influence the relative yields from different fields. Unlike rainfall, the physical soil resource remains constant from year to year and therefore cannot be used as a variable to account for yield differences for successive harvests from the same field. Some soils data from a general purpose free-survey conducted by Spencer Holley were available in the field records. Soils were classified according to the South African Binomial Classification (Soil Classification Working Group, 1977). Where fields comprised more than one soil series, these were simply listed without indicating the proportionate field area of each series. No attempt was made to group soil series with similar physical properties, and important data such as total profile depth were omitted.

Mondi sugarcane fields were surveyed by Keith Snyman and Associates [13] according to FSD [14] norms (Anon., 1994) to provide soils data for this study because the available soils data were severely limited. A digital map of soil units was provided, based on profile descriptions from a 150 m grid-survey conducted across the entire sugarcane area. Soil units were separated from one another by the logical grouping of pedological characteristics determined during the soil survey, each polygon being allocated a unique alpha-numeric code. Soil units were grouped at two levels. At the primary level, soil units were classified according to ten broad soil groups (Table 5) to separate well drained and poorly drained soils and to distinguish between structured and apedal soils. At the secondary level of classification, prevailing pedological properties were used to subdivide these soil groups into soil zones (Table 6). Sugarcane is produced almost exclusively on group A soils (Table 5), predominantly on soil zones Ak, Al, Aa, An, Ab and Ad (Table 6).

For each soil unit the area, soil form, dominant soil families, subdominant soil families, effective rooting depth (ERD), ameliorated ERD, depth limiting material, wetness hazard, cultivation factors and surface features, lithology, organic carbon content, compaction, clay percentage and

---

[13] Keith Snyman and Associates, Soil Scientists and Land Use Planners
PO Box 2, Mntunzini, 3867

[14] Forest Industry Soils Database

total available water (TAW) were described or estimated.

Table 5: The 10 soil groups used in the FSD soil survey system (Anon., 1994).

Soil group symbol	Soil group description
A	Red and yellow apedal soils
B	High chroma plinthic soils
C	Hydromorphic soils
D	Duplex soils
E	Melanic, vertic and red structured soils
F	Lithosols or skeletal soils
G	Podzolic soils
H	Other youthful soils
I	Miscellaneous land classes
J	Organic soils

The soils data were incorporated in the GIS database and the soils map was used with the coverage of sugarcane field boundaries to define soil properties at a field level. The proportional field area (%) of each soil unit, soil zone and soil group was calculated, ignoring polygons contributing less than 0.5% to the total field area. This filter helped to simplify the data, especially where soil changes essentially coincided with field boundaries. A total of 58 soil units (10 soil zones) were identified for the sugarcane fields, individual fields comprising one to seven different units. For this reason the proportional field area of each soil unit was used to calculate field area-weighted means of the quantitative soil properties (ERD, clay content and TAW). Area-weighted classes for cultivation factors (profile gravel and rocks), texture, lithology and organic carbon content were also defined at a field level.

A dominant field soil type was defined for each field, based on the soil zone classification, to allow the effect of different soils on field yields to be investigated. An area-weighted soil type could not be calculated because these groupings are qualitative. Soil types were defined as orthic red, orthic yellow, humic red and humic yellow soils groups with a purity of 70%, i.e. fields classified into one of these four groups should comprise at least 70% of soil zone Ab, Ad, Ak and Al respectively, including the contribution of the undifferentiated soil zones (Table 6).

Table 6: Soil zone symbols, related soil types and soil forms used in the FSD soil survey system (Anon., 1994).

Soil Zone	Soil Type (Soil Zone group shown in brackets)	Soil Forms †
	RED AND YELLOW APEDAL SOILS (A)	
Aa	Undifferentiated humic soils	Kp, Ma, Ia, Lu, Sr, No
Ab	Red apedal dystrophic soils	Hu
Ac	Undifferentiated red and yellow apedal dystrophic soils	Hu, Cv, Gf
Ad	Yellow apedal dystrophic soils	Cv, Gf
Ae	Red apedal eutrophic soils	Hu, Ky, Py, Gr
Af	Red apedal mesotrophic soils	Hu
Ag	Yellow apedal mesotrophic soils	Cv, Gf
Ah	Undifferentiated red and yellow apedal mesotrophic soils	Hu, Cv, Gf
Ai	Yellow apedal eutrophic soils	Cv, Mp, Ak
Aj	Undifferentiated red and yellow apedal eutrophic soils	Hu, Cv, Gf, Ky, Py, Gr, Mp, Ak
Ak	Humic soils / red subsoils	Ia
Al	Humic soils / yellow subsoils	Ma, Kp
Am	Humic soils / cutanic subsoils	Lu, Sr
An	Humic soils / rocky material	No
	HIGH CHROMA PLINTHIC SOILS (B)	
Ba	Plinthic dystrophic soils - red	Bv, Bd
Bb	Plinthic dystrophic soils - yellow	Av, Gc, Pn
Bc	Plinthic eutrophic soils - red	Bv, Bd
Bd	Plinthic mesotrophic soils - red	Bv, Bd
Be	Plinthic mesotrophic soils - yellow	Av, Gc, Pn
Bf	Plinthic eutrophic soils - yellow	Av, Gc, Pn
	HYDROMORPHIC SOILS (C)	
Ca	Undifferentiated hydromorphic soils	Ka, Kd, Lo, Wa, We
Cb	E-horizon hydromorphic soils	Kd, LO, Wa
Cc	Non-E hydromorphic soils	Ka, We
	DUPLEX SOILS (D)	
Da	Red duplex soils	Sw, Va, Km
Db	Non-red duplex soils	Sw, Va, Es, Km, Ss, Se
Dc	Undifferentiated duplex soils	Sw, Va, Es, Km, Ss, Se
	MELANIC, VERTIC AND RED-STRUCTURED SOILS (E)	
Ea	Margallitic soils	Bo, My, Mw, Ik, Wo, Sn, Im
Eb	Mesotrophic red-structured soils	Sd
Ec	Vertisols	Ar, Rg
Ed	Undifferentiated melanic, vertic and red-structured soils	Sd, Bo, My, Mw, Ik, Wo, Sn. Im, Ar, Rg
Ee	Eutrophic red-structured soils	Sd
	LITHOSOLS (F)	
Fa	Undifferentiated lithosols	Gs, Ms, Cf, Dr, Cg, Kn
Fb	Soft lithocutanic soils	Gs, Cf
Fc	hard lithocutanic soils	Gs, Cf, Cg
Fd	Lithosols with hard rock	Ms, Dr, Cg, Kn
	PODZOLIC SOILS (G)	
Ga	Podzols on wet, unconsolidated material	Lt, Wf
Gb	Podzols on non-wet, unconsolidated material	Pg, Cc
Gc	Podzols with placic pan or saprolite	Gk, Hh, Ts, Jb
	YOUTHFUL SOILS (H)	
Ha	E-horizon sands with pale topsoils	Fw
Hb	Undifferentiated sands and other soils	Fw, Kd, Vf, Ct, Lo
Hc	E-horizon sands on yellow apedal or neocutanic soils	Ct, Vf, Kk
Hd	Red neocutanic soils	Tu, Et, GM, Ou, Oa, Mu, Ad, Pr, Tr, Ag
He	Non-red neocutanic soils	Tu, Et, GM, Ou, Oa, Mu, Ad, Pr, Tr, Ag, Br
Hf	E-horizon sands with dark topsoils	Fw
Hg	Regic sands	Nb
	MISCELLANEOUS LAND CLASSES (I)	
Ia	Alluvial / colluvial deposits	Du
Ib	Surface rock (60-80%), outcrops and boulders with miscellaneous soils	Miscellaneous
Ic	Surface rock (>80%), outcrops and boulders with little soil	Miscellaneous
Id	Man-made soil deposits	Wb
	ORGANIC SOILS (J)	
Ja	Organic soils	Ch

† Soil forms classified according to the Taxonomic System for South Africa (Soil Classification Working Group, 1991).

All fields for which a dominant soil type could not be defined using these criteria were investigated. These fields invariably comprised similar proportions of either orthic or humic red-and-yellow soils. For this reason two additional soil types were defined, viz. orthic red / yellow and humic red / yellow soils, comprising similar proportions of either soil zones Ab and Ad or Ak and Al (Table 6). Because dominant field soil types were defined somewhat arbitrarily, two additional soil groupings were defined for use as possible yield predictors. These soil groups were classified according to the dominant soil colour (red, yellow or red / yellow) and according to humus content (orthic or humic).

During the spatial analysis of soil distribution between fields "sliver polygons" were created automatically where the soils map did not overlay perfectly with the field boundary map. Only the sugarcane field soils were surveyed and therefore no soils data were available for land adjoining the sugarcane areas. Although both coverages represent the full extent of the sugarcane fields, they do not overlay perfectly and this highlights one of the sources of inaccuracy (Johnson, 1994) that can exist in GIS data. Where sliver polygons without soils attribute data were created, the pedologist's soils field boundaries, demarcating the end of their survey (not the soil unit boundary), were extended manually to coincide with the field boundary coverage.

3.4 FIELD AREA

Field areas need to be determined accurately for the meaningful interpretation of field records and the efficient management of field operations. Field yields are normalized by area, expressed as TCH and TSH, to allow for comparisons to be made between different fields based on measures expressed in the same units. Similarly, the quantities of agricultural inputs required per field are determined by multiplying the recommended application rate per hectare by the field area.

Three different sets of field areas were used for the Mondi sugarcane enterprise since 1987, in spite of field layouts remaining largely unchanged. The SASA field areas, determined for the sugarcane quota system, were used until 1992. These were determined to 0.1 ha and represented nett arable areas (Appendix 3). All roads, breaks and waterways were excluded from reported field areas. Thereafter Mondi's Midlands estates, including the sugarcane areas, were remapped from new stereoscopically corrected aerial photography, using a computer aided design (CAD) program, MICROSTATION, to capture the spatial data. These areas excluded field boundary roads / waterways and main roads, but the internal field infrastructure (roads, breaks and waterways) was shown as a line only, without taking the area of these features into consideration. The CAD areas (Appendix 3) therefore represented gross arable areas and were therefore generally larger than those determined by SASA. Recorded field areas changed extensively and field management was adjusted accordingly. These area changes distorted the data recorded in the field records because there were no changes in the physical field boundaries. During 1996 the CAD maps were converted into ARCINFO coverages for the Mondi GIS program. While the conversion from one software package to the other did not change field areas, subsequent processing of the spatial data in the GIS caused further area changes. This processing included the exclusion of all recorded field infrastructure from the field cropping areas using an 8 m buffer, irrespective of the real widths of the various features. The 8 m wide field infrastructure was probable realistic for Mondi's timber enterprise but grossly overestimated the widths of the sugarcane field infrastructure. Consequently the GIS areas (Appendix 3) were considerably smaller than those reported by SASA. Depending on the set of field areas used, Mondi's Midlands sugarcane area (excluding Elandsvlei) ranged from 894.8 ha (GIS areas) to 1 064.0 ha (CAD areas).

Among the more serious data distortions caused by the area changes were that the agricultural inputs were applied according to the SASA areas while yields were calculated using CAD areas for some crop cycles, and that field yields (per hectare) from before and after the mapping changes appeared to be different although total field production remained fairly constant. Such distortions render the data useless for deriving information to guide management decisions. It was therefore essential to verify and correct the sugarcane field areas. A GPS survey of nine selected sugarcane fields was conducted on a portion of Canema at Umvoti. A larger number of fields could not be surveyed because the GPS equipment was available for only one day. The limited time during which a suitable satellite constellation was available for surveying precluded sample-surveys of fields at Richmond and Kranskop because of the long distances that needed to be travelled between estates.

The GPS aerial was mounted in the centre of a vehicle roof before travelling along all the boundaries, infield roads, breaks and waterways of the selected fields. The widths of these features were measured by tape measure and recorded. The survey data were imported as a GIS coverage after correcting the co-ordinate readouts for satellite distortion. The survey lines were compared with the original CAD lines. These were found to correspond well with each other, confirming the general accuracy of field infrastructure captured on the CAD maps. It was concluded that the differences between SASA, CAD and GIS field areas (Appendix 3) were caused primarily by differences in the assumed widths of the field infrastructure.

Estate managers at Kranskop and Richmond were consulted regarding the width of their sugarcane field infrastructure. The same infrastructure specifications as those measured at Canema apply to all estates, with the exception of Sutherlands, where infield roads and breaks are 5 m wide. These specifications, presented in Table 7, were used to buffer all the field infrastructure recorded on the CAD maps to the required widths in the GIS field boundary coverage. The features were all excluded from the calculated field areas (nett arable areas), presented as the "New GIS" areas in Appendix 3.

Table 7: Sugarcane field infrastructure specifications used on Mondi's Midlands estates.

Feature	Width (m)
Main roads	8
Waterways and field boundary roads	5
Infield roads and breaks †	3

† This infrastructure is 5 m wide at Sutherlands.

Since the GPS survey data from the map validation study on Canema represent the most accurately defined field boundaries available, the survey lines were buffered to the measured widths of the field infrastructure, and the resulting coverage compared with the New GIS field boundary coverage (Figure 4). This check confirmed the general accuracy of the digital field maps and suggested that the New GIS areas could be used with confidence. Across all 146 fields, the New GIS areas (974.3 ha) were more similar to the SASA areas (998.3 ha) than either the CAD areas (1064.0 ha) or the GIS areas (894.8 ha). For individual fields this trend was, however, not consistent (Appendix 3). The SASA, CAD, GIS and New GIS areas for the eight fields surveyed on Canema were compared with the GPS areas, as shown in Table 8. The New GIS areas agreed more closely with the actual field areas (GPS) than the original areas. Because the New GIS areas were the most accurate estimates of actual field area, all the field records were reworked using the New GIS areas to remove the area related data distortions introduced over the years.

Table 8: Field areas for the nine sugarcane fields surveyed at Canema.

Field number	Original areas (ha)			Revised areas (ha)		GPS Comment
	SASA	CAD	GIS	New GIS	GPS	
DS10	5.4	5.4	4.5	5.0	5.0	
DS12	6.7	6.6	5.3	6.0	6.1	Recently replanted.
DS16	7.5	8.0	6.3	7.2	7.0	Infield road not shown on CAD map.
DS18	8.3	8.4	6.6	7.5	7.7	
DS19	7.9	8.2	6.4	7.2	7.2	
DS22	9.7	9.8	7.5	8.7	8.8	
DS23	10.4	8.0	6.9	7.4	7.5	
DS25	18.0	21.1	18.2	19.8	19.6	
ES01	4.7	4.5	3.7	4.0	4.3	
Totals	78.6	80.0	65.4	72.8	73.2	

For field DS16, one infield road was not captured on the CAD maps. For this reason the road was not shown on the New GIS coverage (Figure 4) and was also not excluded from the calculated field area. These mapping errors were not corrected and, provided they were made infrequently, do not affect field areas substantially (Table 8). When sugarcane fields are replanted, the re-working and improvement of field conservation structures often causes field area changes. Field DS12 was replanted during 1997 and the changes in field layout are shown in Figure 4 where the position of infield roads determined by GPS is different to that captured for the CAD maps used to generate the New GIS coverage. The field area of DS12 increased marginally (1.7%) at replanting because the number of internal roads remained unchanged.

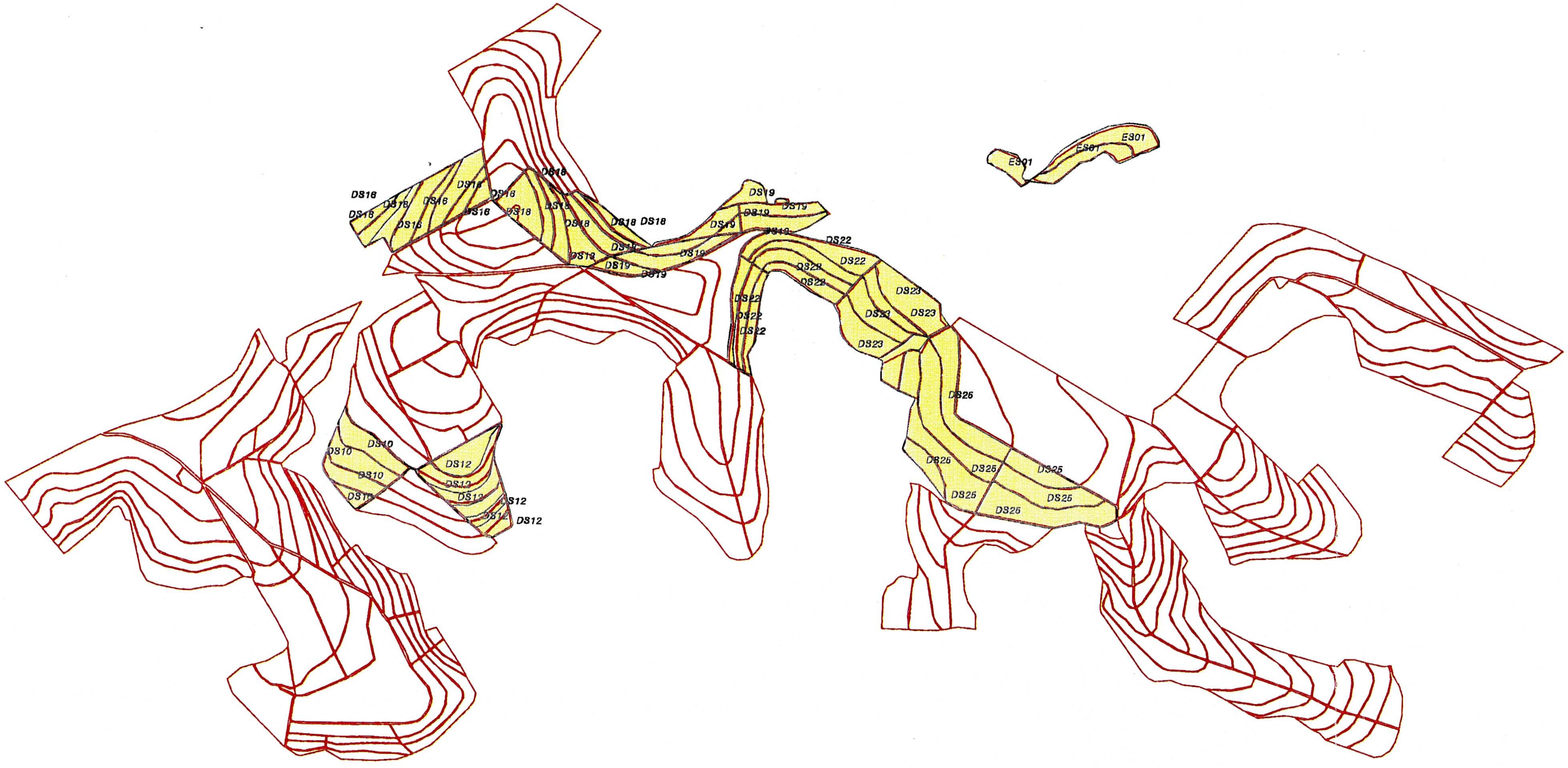


Figure 4 : Canema map validation.

3.5 FIELD ATTRIBUTES

Physical field properties remain constant across crop cycles and, like the soils resource, are important only when investigating production differences between fields. No field attributes were recorded by Spencer Holley. The selection of field attributes to be included in the field records was objective in the sense that only properties that influence the biology related to sugarcane yield were considered. They were not intended to represent a complete list because other considerations, e.g. the perceived importance as a yield determining factor and ease of procurement, determined which ones were included with the field records.

3.5.1 Aspect

The primary determinants of sugarcane yield, radiation, temperature and crop water supply, are all influenced by field aspect to some extent. The overall temperature and rainfall regime differ markedly between northern / western, and southern / eastern aspects. In the southern hemisphere, northern / western aspects are both hotter and drier than southern / eastern aspects, which tend to be cooler and moister. Field aspect was included in the field records given the importance of water for determining sugarcane yield in the Midlands (Thompson, 1976; Inman-Bamber, 1995; Inman-Bamber *et al.*, 1998).

The Mondi GIS database included a coverage of the Surveyor General 10 m interval contour data for the sugarcane fields. A digital terrain model (DTM) was generated from this coverage using a 20 m^2 grid. The GIS query functions were used to generate a coverage of four aspect classes (northern, eastern, southern and western) using the DTM. Assuming that north represents 0°, the northern aspect class was defined from 315° to 45°; eastern aspects ranged from 45° to 135°; southern aspects from 135° to 225° and western aspects from 225° to 315° (Figure 5). The proportional field area (%) represented by each aspect class was determined using the GIS and included in the field records. Aspect classes accounting for less than 0.5% of the field area were ignored, as for soil units in the field soils classification.

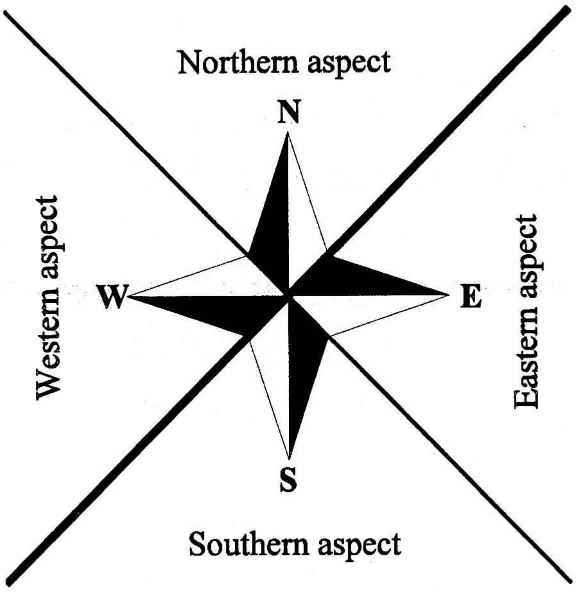


Figure 5: Diagram showing the four aspect classes used to classify field aspect.

The dominant aspect class, defined for each field as the aspect representing the largest proportion of the field area, was used to investigate the effect of different aspects on field yields. The aspect groupings based on four aspect classes were not entirely satisfactory because three of the four aspect classes were represented in many fields. For this reason a second field aspect group, based on two dominant classes, viz. northern and western aspects (225° to 45°), and southern and eastern aspects (45° to 225°) (Figure 5), was defined to correspond with the expected moisture and temperature gradients in the southern hemisphere and tested as a yield predictor during model building.

3.5.2 Slope

Field slope should be included with the field records because it influences the effectiveness of rainfall and thus crop water supply, and may also influence the incidence of frost damage (Mann, 1991). Flat fields are easier to work than steep fields and as a result crop husbandry standards are usually superior on these fields. Besides the biological influences field slope may have on sugarcane yields, it also affects field areas. All the field areas determined from the various mapping systems (Appendix 3) are projected areas, and the actual land surface area is not measured. The greater the field slope, the greater the land surface area relative to the (orthogonal) projected area.

A slope map was generated from the DTM created for the field aspect analysis. Slope classes were defined according the Cedara Land Capability Classification (Smith, 1993) for BRG 5. The computer generated slope map was too detailed to be applied at a field scale because of the numerous slope breaks identified within most of the fields in spite of the general terrain being fairly uniform. For this reason the GIS Zonal Statistics function was used to calculate the area weighted mean field slope (%) for each field and included with the field records.

3.5.3 Altitude

Inman-Bamber (1995) found yield potentials to be closely related to latitude and altitude due to the association of these parameters with radiation and temperature (Schulze, 1982). Since radiation and temperature are difficult to measure at a commercial field scale, the inclusion of field altitude, latitude and longitude as proxies for these primary influences of plant growth may be useful, especially since these parameters also provide the basis for spatial analyses using GIS.

The area weighted mean field altitude (m a.s.l.) was determined from the DTM using the GIS Zonal Statistics function. These values were included with the field records to be tested as yield predictors during model building. Since latitude does not vary much between fields at the estate level, a single value of latitude and longitude was determined for each estate (Section 3.1) and the respective values included with the field records.

3.5.4 Landscape position

The position a field occupies in the landscape influences the extent to which it receives and sheds drainage water. The field terrain unit may therefore affect crop water supply, thereby influencing sugarcane yields. A terrain unit is any part of a landscape with homogenous form and slope. The 5-unit landscape classification (Land Type Survey Staff, 1988) was used to classify all sugarcane fields. According to this classification, a landscape may comprise some or all of the following kinds of terrain unit: crest / plateau (C), scarp (Sc), midslope (M), footslope (F) and floodplain / bottomland (B). The delineation of terrain units is somewhat subjective, being defined on the basis of marked changes in the general pattern and density of relief, slope and drainage; and using variable classification criteria for the units from one landscape to another. For this reason they should be used in conjunction with more quantitative slope and altitude data.

Most of the Mondi sugarcane is produced on M and C unit fields in single phase landscapes, and on F unit fields in multi-phase landscapes in the Midlands. A simple specification of a field's terrain unit is inadequate for defining its position in the landscape when a terrain unit in the conventional sequence (C, M, F, B) is absent, and in multi-phase landscapes where a footslope can be situated above a midslope. For this reason both the terrain unit of the field and the unit occurring below the one specified for the field were included in the field records. Thus fields in a M or F position, adjacent to a bottomland (B), classified as MB or FB, occupy a lower lying position in the landscape than M or F fields in a multi-phase landscape, classified as MF or FM. In this manner all sugarcane fields were classified as one of four terrain classes viz. CM, FM, MB, MF.

3.5.5 Frost

It has been estimated that from a quarter to a third of the Midlands sugarcane crop is affected by frost every year (De Haas, 1981). From time to time much more widespread frosts are experienced and larger areas of sugarcane are periodically subjected to frost. Severe frost permanently damages the apical meristem of sugarcane and forces growers to harvest the crop prematurely, causing serious yield losses (Inman-Bamber, 1991a). The injury caused by frost is characterised by browning of the canopy. The leaf sheaths and the apical meristem of the stalk may also be destroyed. Wilson (1960) recognised four different categories of frost damage to sugarcane. These categories are:

a where the fully developed or exposed parts of leaves are killed, but the innermost and covered parts of the spindle leaves are undamaged and remain green;

b where, in addition to (a), the innermost leaves of the spindle are killed but the apical meristem and the basal parts of other leaves in the spindle are unaffected;

c where, in addition to (a) and (b), the apical meristem is killed; and

d where, in addition to (a), (b) and (c), the lateral buds on the stalk are killed but the stool remains alive.

As long as the apical meristem of the sugarcane remains undamaged, growth resumes with the onset of favourable growing conditions, although the damage caused by frost can usually still be seen for several months after it has occurred (Roth, 1966). The yield loss varies depending on the extent of the damage (within categories (a) and (b)), but at least stalk growth continues. Under rainfed conditions the extent of the growth setback also depends on the degree of drought coincident with the frost (Wilson, 1960). When the apical meristem is killed (category (c) and (d) damage), sugarcane that has formed some stalk must be harvested, irrespective of its age, to avoid serious production losses (Mann, 1991).

Frost damage was not routinely recorded for all crop cycles, but the incidence of frost was sometimes noted where Spencer Holley staff and estate management considered yields to have been reduced by frost. These data are inadequate to investigate the effect of frost on sugarcane yield because sugarcane age at the time of frosting and the severity of the damage were not recorded. For this reason frost damage was surveyed on the five estates during August 1995 to rate the frost susceptibility of fields, following the lowest winter temperatures recorded during the history of sugarcane production in the Midlands (Mann, 1995, personal communication [15]). Vast areas of sugarcane sustained severe frost damage and damage also occurred in scattered areas at lower elevations, where frost has seldom been known to occur previously.

Frost damage to sugarcane on the five estates was mapped according to three classes. Frost class one included all sugarcane that had retained a green canopy in spite of the exceptionally cold winter. Since a healthy canopy of green leaves remained available to commence photosynthesis immediately with the onset of favourable growing conditions in spring, no yield reduction caused by frost was expected. Frost class two included all sugarcane that had been damaged by frost but where the stalks had not been killed, i.e. damage categories (a) and (b) (Wilson, 1960). Although the extent of yield loss that should be expected differs between categories (a) and (b), in both cases green leaf first has to be produced from plant energy sources before photosynthesis and growth can be resumed in spring. The time taken for active growth to resume, and therefore the loss of yield, will increase with the severity of frost damage which varies from year to year. Frost class three included all sugarcane that had been killed by the 1995 frosts, i.e. damage categories

---

[15] Mr QV Mann, Midlands North and Central Extension Officer,
South African Sugar Association Experiment Station,
Private Bag X02, Mount Edgecombe, 4300
Present address: PO Box 65, Dalton, 3236

(c) and (d) (Wilson, 1960). No distinction was made between damage categories (c) and (d) because these differences are only important for the short-term management of the crop (Roth, 1966).

Boundaries between the three classes of frost-affected sugarcane were recorded independently of field boundaries. Frost damaged sugarcane occurring in isolated frost pockets and in long strips where the natural drainage of cold air was obstructed by trees on the lower sugarcane boundary – a common problem on all five estates – was also recorded. However, the sugarcane field is the lowest level of spatial aggregation for which yield and agronomic data are available. For this reason frost affected sugarcane was re-classified on a field-by-field basis, ignoring small isolated areas of damage. This classification was included with the field records even though every crop cycle from a field may not necessarily have been affected by the severity of frost damage defined for the fields because more appropriate records were not available.

The 1995 field frost classification should be used as a rating of the relative susceptibility of fields to cold stress rather than an absolute definition of frost damage potential because of the variable distribution of frost. The distribution of frost and the severity of crop damage are affected by air movement, altitude, slope, aspect, landscape position, temperature (and the duration of the freeze), sugarcane variety, crop age, crop condition and the presence or absence of trash (Wilson, 1960; Roth, 1966; De Haas, 1981; Mann, 1991). Slight differences in wind drift and direction affect the distribution of frost (Mann, 1991) and managers at Richmond reported that areas which are usually frost prone remained green throughout the 1995 winter. These statements are supported to some extent by the Spencer Holley records for Richmond where nine of the 31 frost class one fields, one of the 14 frost class two fields and two of the 11 frost class three fields were affected by frost during 1992. However, the 1992 winter was not unusually cold and the severe drought will probably have contributed to the distribution of the frost damage, especially since drought stress has been observed to cause sugarcane to become more susceptible to frost (Wilson, 1960).

3.5.6 Proximity of trees

Spencer Holley noted the adverse impact trees had on sugarcane growth where they occurred along a field boundary. During the 1995 frost survey, severe frost damage was also associated with trees on the lower field boundary where the natural drainage of air was retarded along a strip at the edge of the fields. For these reasons the presence or absence of trees along any field boundary, and also along the lower field boundary, was determined and the classification included in the field records.

3.6 FIELD HUSBANDRY

Sugarcane is an intensively managed field crop and yields are known to vary substantially between individual producers as a function of the type and timing of their field husbandry practises and management ability in general (Hellmann, 1993; Neen *et al.*, 1994). Some of the field husbandry practises advocated by Spencer Holley are discussed, specifically considering those that differ between the Mondi estates, and others which influence the nature of the field data.

3.6.1 Row spacing

Sugarcane on the Mondi estates is planted using a 1.0 m row spacing although a 1.2 m row spacing has been used on most of the estates in the past. At Canema various row spacings, ranging from 0.9 m to 1.4 m were used before Mondi introduced their "standard" 1.0 m row spacing. The row spacing used was not recorded in the field records for plant crops and therefore this information is not available for the field history.

Results of row spacing experiments in KwaZulu-Natal under favourable growing conditions showed a persistent trend towards higher yields at closer spacings (Thompson and Du Toit, 1967). Increased cane yields were associated with higher stalk populations although the stalks were both thinner and lighter than those from widely spaced rows (Boyce, 1968). Sucrose yield gains of up to 10% could be achieved with 0.9 m rows compared with 1.37 m rows, mainly because of the improved radiation use efficiency of the crop (Inman-Bamber, 1996). Sugarcane grown in narrowly spaced rows is more susceptible to drought than that grown in wider rows and consequently rainfed yields may be lower at closer spacings in certain seasons (Thompson and Du Toit, 1967). Although it is important to consider differences in row spacing when comparing sugarcane yields, this could not be done using the data available for this study.

3.6.2 Variety selection

When commercial sugar production started in the Midlands, mainly NCo376 and NCo298 sugarcane varieties were planted. These are no longer the recommended varieties and N12, which produces cane and sucrose yields superior to the old varieties, is now the dominant variety in the Midlands. Recent progress in plant breeding has resulted in the release of new, further improved varieties of which N31 is the most promising for rainfed production systems in the Midlands (Nuss, 1999, personal communication [16]).

Mondi have been slow to replace the old sugarcane varieties on their Midlands estates. At Richmond, NCo298 was the dominant variety produced until 1997 and at Umvoti and Kranskop a number of fields still grow NCo376. Where new sugarcane varieties were planted, N12 was used predominantly, but N16, an early maturing variety, was sometimes planted on humic soils and frost prone fields while N21, a hard-stalk variety, has recently been planted on fields prone to theft and animal damage.

Varietal differences in total yield and the cane:sucrose yield relation make it essential to take variety into consideration when comparing yields. Since the varieties have not been planted at random in the fields on the Mondi estates it is not possible to investigate site – variety interactions to guide management in the selection of the most suitable varieties for different sections of the estates.

3.6.3 Seed removal

While the partial harvesting of sugarcane from fields for commercial production is unusual, it does occur where only a portion of a field is cut for seed or destroyed by wild fire. Until such a field is replanted, it is effectively growing two (or more) concurrent crop cycles each of which need to be recorded separately, i.e. treated as separate fields with the same history but different current crop status, management requirements and inputs. Besides the difficulties of recording partial field harvests in the field records, the areas of field subsections were only crudely

---

[16] Dr KJ Nuss, Head, Pant Breeding Department
South African Sugar Association Experiment Station,
Private Bag X02, Mount Edgecombe, 4300

estimated by the estate managers and consequently most of the recorded field data were severely distorted by grossly incorrect area estimates. Such fields are difficult to manage because the planning of all field operations is complicated, and the management information that can be obtained from the field records is, at best, misleading. It was not possible to correct these records because no maps show the field sections harvested at specified stages.

3.6.4 Harvesting schedules

Spencer Holley used the soils data from their general purpose soil survey (Section 3.3) to recommend sugarcane harvesting schedules. Fields with weak soils (shallow and sandy soils; soils with high subsoil Aℓ causing shallow rooting) should be harvested during the early part of the harvesting season (autumn and early winter) because the sugarcane on these soils would usually experience the most severe water stress on the estate causing a relatively high sucrose content during a naturally low-sucrose-content period (Van Dillewijn, 1952). Spencer Holley considered another benefit to be that these relatively low potential soils would retain an improved water status during the dry season through reduced transpiration once the canopy was removed, thereby allowing a stronger ratoon crop to develop. During late winter, before the first spring rains, fields with humic soils should be harvested. At this stage the standing crop would have ripened as much as possible, reaching its peak sucrose content for the season (Van Dillewijn, 1952). During the remainder of the harvesting period all other fields due to be harvested should be cut. These would generally comprise fields on the high potential, non-humic soils. The purpose of these recommendations was to maximise the sucrose content of the sugarcane harvested, and so maximize estate income from the relative sucrose payment system throughout the harvesting period. It follows that fields were not harvested in a random manner and this may or may not be important when analysing the data.

3.6.5 Tillage methods

The importance of tillage methods was illustrated by Keig *et al.* (1993) using tillage practice as an explanatory variable for sugarcane yield in Australia. Alvarez *et al.* (1982) also list tillage as an important factor influencing sugarcane yield in Florida. While Spencer Holley specified certain tillage operations and various types of cultivation for the incorporation of fertilizers and lime, these instructions were not retained in the field history. Consequently no records are available for tillage and cultivation practises.

Recently a number of ratoon crops, especially those at Richmond, were ripped on the interrow to help reduce soil compaction. According to Spencer Holley this should extend the life expectancy of the roots and facilitate the movement of fertilizer (specifically lime, gypsum and phosphorus) into the soil along the rip lines. Turner *et al.* (1992) showed that such husbandry practices substantially reduce sugarcane yields because of damage to the roots, the severity of yield loss increasing as the width of the interrow decreases. Although no appropriate records are available, it would be interesting to investigate yield trends associated with this change in crop husbandry.

3.6.6 Ratoon status

Sugarcane yields are expected to decline with successive ratoons (Hoekstra, 1976; Tobin and Ellis, 1988) and the crop is generally ploughed out and re-established after 5 ratoon crops although it is common practice to produce between 4 and 8 ratoon crops (Inman-Bamber and Stead, 1990). Mondi usually produce a plant crop and up to 9 ratoons because Spencer Holley consider replanting to be an expense justified only when varietal changes are recommended, for rectifying serious subsoil $A\ell$ toxicity problems, or where sugarcane has a major disease problem. Crop husbandry practises are therefore directed at sustaining the longevity of the crop. Where yield reductions are "caused" by low plant populations, Spencer Holley recommend gapping up, and where soil compaction problems are likely, deep ripping on the interrow is advocated.

3.6.7 General

Occasional comments on low plant populations and animal damage were included in the field records. Unfortunately these factors were not recorded consistently and the severity of the problem was not quantified and, while it is clear that yields are variably affected by these factors, the available information is of little value for yield predictive purposes. The incidence of lodging was also rarely recorded, although the timing and severity of lodging may severely reduce yields. Thompson (1976) showed that total evaporation is reduced by about 25% in a lodged sugarcane crop and as a result, yield losses increase substantially as the time between lodging and harvesting increases. No suitable records were available to account for lodging on the Mondi estates.

3.7 ANALYSES

Spencer Holley required one representative soil and leaf sample for each crop cycle from every field to be analysed. This policy implied that all fields were sampled annually – either for soil or leaves, since sugarcane is grown over an 18 to 24 month cycle in the Midlands. These analyses were all included in the field records. In spite of this sampling policy numerous analytical results were found to be missing from the field records. Some of these were recovered from the SASEX central database but many other analyses could not be traced.

Soils were sampled before the fields were harvested to ensure that the analyses were returned from the laboratories in good time to order the required fertilizers. As a result, no soil samples were taken from fields that were harvested for unforeseen reasons. Leaf samples were usually taken during February and March. For analytical results to be meaningful, the samples must be taken from actively growing sugarcane between three and nine months old (Anon., 1991) and leaf samples were usually not taken when these sampling conditions were not met, i.e. crops were drought stressed or too old at the time of sampling.

A number of superfluous soil analyses were included in the records where soil samples were taken from fields scheduled for harvesting, but which were carried over to the following season for unforeseen reasons. These analyses were removed from the field records.

3.7.1 Soil analyses

Topsoil samples (0 – 150 mm) were analysed for each crop cycle to monitor soil crop nutrient levels. Subsoil samples from fields were analysed less regularly to monitor subsoil acidity, Aℓ, and the movement of Mg and K, caused by gypsum applications. Soil samples were analysed either by the SASEX Fertilizer Advisory Service (FAS) laboratory or the Grain Crops Institute (GCI) soil laboratory at Cedara. The soil analysis techniques used by the GCI are identical to those used by the Cedara Fertilizer Advisory Service (FERTREC) (Manson *et al.*, 1993), but differ substantially from those used by the FAS (Wood, 1990). The analysis techniques used by the two laboratories are compared in Table 9.

Table 9: A comparison of soil analyses and soil analysis techniques used by the FAS (Wood, 1990) and the GCI (Manson *et al.*, 1993).

ELEMENT	ANALYSIS	DETAILS	FAS	GCI
Nitrogen	Nitrogen mineralization category	Four categories of soil potential to mineralize soil organic matter	Near infra-red spectroscopy (Meyer *et al.*, 1983)	NOT DONE
	Organic carbon (OC)	Estimate of OC% requested for humic and humic phase topsoils	NOT DONE	Near infra-red spectroscopy (soil organic matter calculated empirically from this estimate)
Phosphorus	Phosphorus content	Estimate of plant extractable soil P	Modified Truog procedure (Meyer *et al.*, 1989)	AMBIC extraction (Hunter method)
	Phosphorus fixation	A measure of the availability of added P in soil with a P fixing capacity	Three categories (Wood, 1990) of the phosphorus desorption index (PDI) (Reeve and Sumner, 1970)	Phosphorus requirement factor (PRF) (the quantity of P required to increase the soil test level by 1 mg L^{-1})
Potassium	Potassium content	Estimate of plant extractable soil K	Atomic absorption from 1 M ammonium acetate extraction	AMBIC extraction (Hunter method)
Calcium and Magnesium	Calcium and Magnesium content	Estimate of plant extractable soil Ca and Mg	Atomic absorption from 1 M ammonium acetate extraction	Atomic absorption from 1 M potassium chloride extraction
Zinc	Zinc content	Estimate of plant extractable soil Zn	Displacement with 0.01 M EDTA	Displacement with 0.01 M EDTA
Sulphur	Sulphur content	Estimate of plant extractable soil S	Turbidimetric determination from 0.5 M ammonium acetate extraction (Meyer et *al.*, 1989)	NOT DONE
Aluminium	Aluminium content	Index (FAS) / estimate (GCI) of soil Aℓ content	Exchangeable aluminium index (EAI) 0.2 M ammonium acetate extraction (Reeve and Sumner, 1970)	Exchangeable acidity ($A\ell^{3+} + H^{+}$) 1 M potassium chloride extraction + titration to phenolphthalein endpoint
	Aluminium toxicity	Estimate of phytotoxic soil Aℓ levels	Aluminium saturation index (ASI) (Schroeder *et al.*, 1995)	Acid saturation (% total cations) (Farina *et al.*, 1980)
Clay	Clay content	Estimate of sample clay content	Near infra-red estimate; two categories (<30%; >30%)	NOT DONE, derived empirically from sample bulk density

Historically, soil samples were analysed by the FAS, but recently preferential use has been made of the GCI laboratory, primarily because of a lower cost per sample analysed and a faster sample / analysis turn-around time. Spencer Holley used the FERTREC approaches for correcting Aℓ toxicity and soil acidity (Manson *et al.*, 1993) and followed the liming philosophies of Dr MPW Farina [17], which are based largely on maize requirements in the KwaZulu-Natal Midlands, and require an estimate of soil acid saturation which is not determined by FAS (Table 9).

Since nutrient threshold values for sugarcane in KwaZulu-Natal have been determined only for FAS nutrient extractions, GCI soil nutrient contents were converted to FAS equivalents to determine the plant nutrient requirements. Analyses at the GCI are conducted on a soil volume basis (mg L^{-1}). The FAS laboratory analyses soil on a mass basis (mg kg^{-1}), expressing results in parts per million (ppm). The analytical results are reported in units understood by farmers (kg ha^{-1}), assuming that 1 ppm is equivalent to 2.25 kg ha^{-1} on the basis of:

a a soil working depth of 150 mm; and

b an industry-average soil bulk density of 1.5 g cm^{-3} (irrespective of the actual sample density).

The conversion of GCI analyses to FAS equivalents therefore involves a conversion from a volume to a mass basis (multiplying by sample density) and the use of a factor to account for the different quantities of plant nutrient extracted by the laboratories, depending on the analytical techniques used. Spencer Holley use factors developed by the Triomf fertilizer company to account for these extraction technique differences.

While it is possible to develop such conversion factors (Meyer, 1997, personal communication [18]), the accuracy of the Triomf factors has not been statistically proven and none of the relevant research has been published. However, no alternative conversion factors are available and no data

---

[17] Dr MPW Farina, Grain Crop Institute
Agricultural Research Council, Cedara

[18] Mr JH Meyer, Head, Crop Nutrition and Soils Department,
South African Sugar Association Experiment Station
Private Bag X02, Mount Edgecombe, 4300

are available to refute the Triomf factors. For this reason the GCI analyses, expressed in FAS equivalents were retained in the field records, although the analysing laboratory was noted as an additional record to allow the separate evaluation of analyses.

3.7.1.1 *Plant nutrient requirements*

Sugarcane N fertilizer requirements were determined according to FAS guidelines (Anon., 1991) from soil form and soil mineralization potential data. Since soil mineralization potential is not determined by the GCI, Spencer Holley used values determined by the FAS for previous samples when the N requirement determination was based on analyses from the GCI.

The P fertilizer requirements were determined according to FAS guidelines for single crop recommendations for the Midlands (Anon., 1991), considering soil P content (kg ha^{-1}) and the phosphorus desorption index (PDI). Different quantities of soil P are extracted by the two soil laboratories (Table 9), and a factor of 1.76 was used to convert GCI P analyses (first converted to a mass basis) to FAS (ppm) values. All soil P levels reported in the field records were expressed as FAS equivalents (kg ha^{-1}). Since GCI analyses do not include an estimate of PDI, data from previous FAS analyses for each field were used when requirements were determined using GCI analyses.

Potassium fertilizer requirements were determined according to FAS guidelines, based on the soil K content (kg ha^{-1}). Optimum soil nutrient levels vary according to soil texture (Anon., 1991). Soil K content was reported in FAS equivalents (kg ha^{-1}) in the field records. Although the two laboratories use different analytical techniques for the determination of K, no conversion factor (other than a conversion to a mass basis) was used when relating GCI K test values to FAS equivalents. This is supported by the findings of Farina and Channon (1991) that, for highly weathered soils, the quantities of soil K extracted by the Hunter method (GCI) are similar to those determined using neutral ammonium acetate (FAS).

Clay content was not calculated using the empirical function based on sample density for GCI analyses (Manson *et al.*, 1993) and FAS clay contents from previous analyses were used when K requirements were determined from GCI analyses. Soil clay content is classified according to only two broad classes by the FAS (Table 9) because the near infra-red calibrations for clay are not very reliable. These classes are adequate for determining fertilizer rates according to FAS norms but the clay contents estimated by experienced pedologists during the 150 m grid survey are believed to be more representative of field conditions. For this reason they were tested as yield predictors in preference to the soil analytical values during model building.

Soil Ca and Mg were recorded as FAS equivalents in the field records. Theoretically, similar quantities of Ca, and also of Mg, should be extracted by the two laboratories because the cations are displaced from the cation exchange sites by NH_4^+ and K^+ of equal concentration (Table 9). However, factors of 0.76 and 0.71 were used to convert GCI soil test values (expressed on a mass basis) to FAS (ppm) values, for Ca and Mg respectively. The FAS recommendations (Anon., 1991) for correcting soil deficiencies in these nutrients were followed.

The methods used by the FAS and the GCI to measure soil Zn are essentially the same since the FAS use 0.01 M EDTA to displace soil Zn while the GCI use AMBIC solution which contains 0.01 M EDTA (Table 9). Analytical results from the two laboratories were therefore used interchangeably, a conversion from a volume to a mass basis being used to express GCI analyses in FAS (ppm) equivalents. Where large quantities of lime are applied, induced trace element deficiencies have been observed (Reeve and Sumner, 1970; Moberly and Meyer, 1975). For this reason the Zn threshold value varies according to levels of applied lime and clay content (Anon., 1991).

Only the FAS analyse soil samples for S (Table 9). The importance of soil S content for reducing the toxic effect of $A\ell^{3+}$ in sugarcane has been established (Schroeder *et al.*, 1993) and is especially important in soils with a high organic matter content. Deficiencies in Midlands soils are rare and soil S levels were not recorded in the field records. However, in recent years Mono-ammonium Phosphate (MAP) and Di-ammonium Phosphate (DAP) have replaced Single Superphosphate as the primary P fertilizer, and while Superphosphate contains substantial amounts of S (from the sulphuric acid reactions), both MAP and DAP contain none. As a result no indirect additions of S have been made to many fields and deficiencies are expected in some soils of low S mineralization potential, i.e. those with a low soil organic matter content.

3.7.1.2 *Aluminium toxicity and soil acidity*

Aluminium toxicity is a serious problem in many Midlands soils and is often associated with a high P fixation capacity and soil acidity (Reeve and Sumner, 1970). The FAS laboratory measures pH in water, and soil acidity was recorded using these units in the field records. The GCI pH values, measured in a 1 M potassium chloride (KCl) solution, were converted to $pH_{(Water)}$ equivalents using a common rule-of-thumb conversion, $pH_{(KCl)} + 1$. Soil acidity is important in determining the availability of plant nutrients in the soil solution because above $pH_{(Water)}$ 5.3 the solubility of phytotoxic $A\ell^{3+}$ decreases rapidly (Jenny, 1961).

Aluminium toxicity is evaluated on the basis of the relative concentration of $A\ell$ in the soil using either the aluminium saturation index (ASI), calculated from FAS analyses, or the acid saturation, calculated from GCI analyses (Table 9). The FAS and GCI laboratories use very different analytical approaches for assessing soil $A\ell$ levels (Table 9), and recent research (Schroeder, 1997) shows that any relation between the exchangeable aluminium index (EAI) and exchangeable acidity ($A\ell^{3+} + H^+$) would probably be curvilinear in nature. Separate records of EAI, determined by the FAS, and ($A\ell^{3+} + H^+$), determined by the GCI, were kept in the field records.

Where the field records contained entries for both EAI and ($A\ell^{3+} + H^+$) for one soil analysis report, no sample density had been recorded, indicating that the sample had been analysed by the FAS. In these cases EAI and ($A\ell^{3+} + H^+$) consistently differed by a factor of 90, i.e. a conversion of EAI (ppm mass) to c $mol_c\ kg^{-1}$. This value merely expresses the EAI in different units and, although the values were recorded in the ($A\ell^{3+} + H^+$) column, they do not correspond to GCI ($A\ell^{3+} + H^+$) equivalents because:

a no conversion from a mass to a volume based measure was made (sample density was not recorded for the FAS samples); and

b no conversion factor was applied to account for the different extraction techniques used by the two laboratories.

For these reasons all ($A\ell^{3+} + H^+$) values for FAS samples were removed from the field records.

Acid saturation is calculated as the percentage of total cations (effective cation exchange capacity) occupied by ($A\ell^{3+} + H^+$) (Manson *et al.*, 1993). For GCI analyses, total cations are calculated as the sum of soil Ca, Mg, K and ($A\ell^{3+} + H^+$), all expressed in c mol_c L^{-1}. Total cations are not determined by the FAS and cannot be calculated from the individual soil cation test levels because only a partial extraction of Aℓ is made. However, "total cations" were calculated and recorded for most of the FAS analyses in the field records, expressing soil Ca, Mg and K levels in c mol_c kg^{-1} and adding to this sum ($A\ell^{3+} + H^+$) incorrectly calculated from the EAI. The "total cations" were in turn used to calculate "acid saturation" for FAS analyses. The "acid saturation" recorded for FAS analyses in the field records is in fact the ASI which differs substantially from the GCI acid saturation (Schroeder *et al.*, 1995). Extensive cleaning of these data was therefore necessary, separating ASI entries for FAS analyses from acid saturation records for GCI analyses. The "total cations" calculated for the FAS analyses were all deleted from the field records because they cannot be compared to the GCI total cations since:

a no correction factors were used to account for the different proportions of soil cations extracted by the two laboratories;

b no conversion from a mass basis to a volume basis was made (sample density was not recorded for the FAS samples); and

c the EAI was incorrectly converted to represent ($A\ell^{3+} + H^+$).

3.7.2 Leaf analyses

Leaf samples were all analysed by FAS, using near infra-red reflectance for N determination (Meyer, 1983), and X-ray spectroscopy for P, K, Ca, Mg, S, Zn, Cu and Mn determinations (Wood *et al.*, 1985). In the Midlands, foliar threshold values for N range from 1.6% to 1.9% (Anon., 1991) depending on whether the samples are from plant or ratoon crops and according to the month of sampling. The P foliar threshold for variety N12 is 0.16%, and 0.19% for all other varieties (Anon., 1991). For the other nutrients, foliar thresholds (K: 1.05%, Ca: 0.15%, Mg: 0.08%, S: 0.12%, Zn: 13 ppm, Cu: 3 ppm and Mn: 15 ppm) are constant across sampling months, plant and ratoon crops, and all varieties (Anon., 1991).

Foliar nutrient levels of N, P, K, Ca, Mg, S and Zn were expressed as a percentage of the relevant nutrient threshold to allow direct comparisons across all N and P analyses to be made. The date of sampling was recorded together with each leaf analysis because Meyer *et al.* (1989) noted the importance of considering *inter alia* the age of sugarcane and the month of sampling when interpreting foliar analyses according to nutrient threshold values.

Various soil and foliar nutrient ratios were calculated using the analytical data and included with the field records. These ratios are useful for describing nutrient balances and interactions. Nutrient ratios were recorded in association with the absolute concentration of the respective nutrients to avoid the information loss that is associated with the exclusive use of ratios.

3.8 AGRONOMIC INPUTS

A detailed record of all agronomic inputs was kept in the field records. Most of these data were duplicated in the record system because they were recorded in the estate manager's seasonal field operation section and the individual field history section. All the available data were checked thoroughly and the duplicated entries were compared as a control for transcription errors. Managers responsible for the Kranskop sugarcane from July 1992 onwards were consulted to correct discrepancies found in their sugarcane records for this period. A similar data validation procedure was conducted with managers responsible for the Umvoti sugarcane from November 1994 and the Richmond sugarcane from September 1994. In addition, detailed notes of field husbandry operations from the 1989 / 1990 seasons onwards were available for some fields to confirm the Spencer Holley records. Where problems in field records could not be resolved satisfactorily with the estate managers or for which no additional documentation explaining the actual field inputs could be found, these records were highlighted to ensure that they were not used indiscriminately during model building.

3.8.1 Fertilizer

Both the fertilizer types and amounts applied were recorded. The number of bags of fertilizer applied to a field were usually recorded together with the actual application rate per hectare. For this reason it was possible to account for the effect of area changes on fertilizer rates, caused only by changes in the mapping systems used (Section 3.4). Fertilizers were applied by hand along every second interrow using 500 mL oil tins to spread the material along a specified row length, determined according to the required application rate. This application method was used in preference to machine applications, primarily because it was less costly but also because application rates were supposedly more accurate since fertilizers were being applied along a specified row length rather than being spread over an inaccurately determined field area. In this context it is particularly worrying that fertilizer inputs per field changed more or less proportionally to the variously reported field areas (Section 3.4) although physical field boundaries remained unchanged. Managers insist that the recorded number of fertilizer bags were applied to each field and, after correcting the field areas, this caused actual fertilizer application rates to vary substantially from those recommended by Spencer Holley.

When making fertilizer recommendations, the full spectrum of accumulated analytical information and the history of each field was studied and considered by Spencer Holley staff. This was considered to be important because, in spite of every effort being made to ensure that the soil and leaf samples were representative of the field, research in the United States of America has indicated that 80% of the variability in soil test data is inherent in the sample itself (Hauser, 1973). As a result the quantities of fertilizer recommended, especially those of K and P, sometimes differed from the theoretical requirements determined according to FAS norms.

The total inputs of elemental nutrients (N, P, K and Zn) per field were calculated using the field records of fertilizer brands and the number of bags applied. Fertilizer rates were calculated using the New GIS field areas (Appendix 3). The Zn content of zincated fertilizers was not always recorded in the field records and for some crop cycles duplicated records corresponded exactly, except as to whether or not the applied fertilizers contained Zn. Where such discrepancies were found the data were retained in the field record, noting the uncertainty about the Zn fertilizer input.

Since fertilizer input requirements were determined primarily as a function of soil nutrient levels, "total" P, K and Zn nutrition for a crop cycle were calculated as the sum of topsoil P (kg ha^{-1}) and applied P fertilizer (kg ha^{-1}), topsoil K (kg ha^{-1}) and applied K fertilizer (kg ha^{-1}) and topsoil Zn (kg ha^{-1}) and applied Zn fertilizer (kg ha^{-1}) respectively. In many situations these nutrient levels may provide more meaningful information about crop cycle nutrition levels than either soil nutrients or fertilizer applications alone because low fertilizer inputs are generally made when soil nutrients are high and *vice versa*.

3.8.2 Lime and gypsum

Soil acidity and Aℓ toxicity are generally ameliorated using calcitic or dolomitic lime. Spencer Holley believe that the SASEX lime recommendations are too low for the Midlands, and do not accept the experimental results on which their recommendations are based. For the period to which the field records apply, recommended lime inputs were therefore determined using the FERTREC method (Manson *et al.*, 1993) for both FAS and GCI soil analysis results. The FERTREC lime requirement was calculated as the amount of lime required to neutralize exchangeable acidity in excess of the crop's acid saturation tolerance. No acid saturation threshold value (permissible acid saturation – PAS) has been determined for sugarcane, and lime requirements for sugarcane are therefore determined subject to guessing an appropriate PAS.

Spencer Holley used a PAS of 10% for sugarcane, irrespective of whether the acid saturation – GCI analyses, or the ASI, incorrectly calculated as an acid-saturation-equivalent for FAS analyses, was used. This is equivalent to the PAS used for perennial ryegrass and winter wheat which are extremely intolerant of Aℓ toxicity, compared with sugarcane. At this level of liming negative yield responses to lime were recorded for NCo376, N16 and N12 in SASEX experiments (Schroeder *et al.*, 1995). These experimental results have been used to derive an ASI threshold of 40% for N12, equivalent to a PAS of about 59%, and 20% for all other varieties, equivalent to a PAS of about 39% (Schroeder, 1997). Recent revisions of the FAS lime recommendations (Anon., 1996) have further reduced the average amount of lime recommended by the FAS.

Spencer Holley considered it preferable to correct soil Aℓ problems when replanting fields, but recommended topdressing with a combination of lime and gypsum where Aℓ problems persisted after replanting. For plant crops the lime requirement was calculated using a PAS of 10% and this quantity of lime was ploughed into the soil during land preparation. Where subsoil acidity was considered to be high (no consistent criterion was used) and gypsum was expected to have an ameliorative effect (gypsum appears to be effective only in naturally acidic soils with a high sesquioxide content), various quantities of gypsum were recommended in addition to the lime required to reduce acid saturation in the top 200 mm of soil to 10%. The quantity of gypsum applied was determined on an *ad hoc* basis because no scientific methods exist to calculate soil gypsum requirements nor to predict its ameliorative effect.

Deep application of lime or phosphogypsum, and the surface application of phosphogypsum for the amelioration of Aℓ toxicity in the subsoil, are not recommended by SASA (Anon., 1991) because experimental results have consistently shown no significant yield responses to these practices (Meyer *et al.*, 1991; Turner *et al.*, 1992). However, improved root development and considerable yield responses to lime and gypsum have been recorded in Midlands maize crops (Buyeye *et al.*, 1985; Farina and Channon, 1988; Shainberg *et al.*, 1989) and as a result of these findings there is strong support for such ameliorative practices in the Midlands.

Although the FERTREC methods to determine lime requirements were used, their approach for topdressing established pastures (ratoon crops) with lime (Manson *et al.*, 1993) was not followed. For low-number ratoon crops where Aℓ problems persisted after planting, the lime requirement was calculated using a PAS of 10%, but dolomitic lime was applied up to a maximum of only 2 t ha^{-1} because of the limited movement of lime in the soil. The balance of the calculated lime requirement was applied as gypsum (up to 5 t ha^{-1}) because of its superior mobility in the soil, being substituted for lime on a 1:1 basis. This practice is not based on any scientific principles, and in some cases gypsum may have been applied to ratoon crops on soils where no ameliorative response should be expected. In an effort to maintain topsoil Mg at an acceptable level where gypsum topdressing was recommended, Spencer Holley specified that the gypsum be applied first, and that the 2 t ha^{-1} dolomitic lime be applied only once the gypsum had moved into the soil.

3.8.3 Herbicides

The weed species causing a problem in fields of young sugarcane were recorded in the field records together with the herbicide brands and application rates used to control them. The date of treatment and weather conditions at the time of application, which affect the effectiveness of chemical weed control, were also recorded. Plant crops usually received a short-term pre-emergence herbicide treatment and a long-term post-emergence herbicide treatment, while ratoon crops received only the long-term treatment. Where inadequate weed control was achieved, fields were usually re-sprayed with additional chemicals. Hand-hoeing to control weeds was not recorded.

The amounts of active herbicide ingredients applied to each crop cycle were calculated using the data from the field records. This was necessary because a number of the 29 brands used over the study period contain the same active ingredients albeit in different concentrations. Fields that were sprayed more than once to achieve satisfactory weed control generally received greater herbicide inputs than those sprayed only once. These data are fairly meaningless because of the limited period for which weeds are controlled, generally ranging from four to fourteen weeks. The amounts of the various active ingredients applied are also not directly comparable in terms of their weed control characteristics. As a result there was no meaningful method of summarizing the data and for this reason only the most commonly used active ingredients were considered. For plant crops paraquat (Gramoxone ®, Paraquat ® and Skoffel ®) was selected and for ratoon crops only diuron (Diuron ® and Impi ®) and hexazinone (Velpar ®) were included with the field records.

3.9 AGRONOMIC OUTPUTS

The sugarcane yield data recorded in the field records were obtained from mill consignment notes which are also used to calculate relative sucrose deliveries which form the basis of sugarcane payments to growers. Although each estate supplied only one mill for an entire season, for the period to which the field records apply, sugarcane from Kranskop was crushed at the Darnal and Glendale mills, sugarcane from Richmond was crushed at the Mount Edgecombe and Noodsberg mills while that from Umvoti was crushed only at the Noodsberg mill.

3.9.1 Cane yield

The total tons of sugarcane harvested for a field, TCH and TCHM were recorded for each crop cycle in the field records. All recorded cane yields from the five estates were verified using the estate crop estimate books from the 1990 milling season onwards; older records were not available. Since the Spencer Holley records were generally accurate, the pre-1990 yields were assumed to be correct.

Each consignment (truck-load) of sugarcane is weighed at the mill weigh-bridge. A record is also kept of the field from which each consignment originates and field cane yields are compiled using the relevant weigh-bridge data. The accuracy of these yield data depend entirely on the correct field number being recorded for each consignment. Additional sources of inaccuracy arise where consignments comprise sugarcane from two fields – a common occurrence when the last sugarcane from one field and the first sugarcane from another are transported to the mill as a mixed load. It is customary to allocate a portion of these mixed loads to each field and except for very small fields, these sources of inaccuracy are negligible because of the large number of truck-loads per field (Hellmann, 1993). Since fresh-mass cane yields are recorded and the dry matter content is not reported, seasonal changes in sugarcane moisture content cannot be separated from reported yields.

3.9.2 Sucrose yield

The relative sucrose content (% fresh mass), TSH and TSHM were recorded for each crop cycle in the field records. Approximately 66% of consignments are sampled at the mill to determine the absolute sucrose content of the sugarcane. The relative sucrose content is calculated on a weekly basis using the mean of the absolute sucrose content, weighted by the cane yield to which each sucrose analysis applies, using the following formula:

$$\begin{array}{c}\textit{Relative sucrose content}\\ \textit{for the week}\end{array} = \begin{array}{c}\textit{Mean absolute sucrose}\\ \textit{content for the week}\end{array} - \begin{array}{c}\textit{Actual mill-average absolute}\\ \textit{sucrose content for the week}\end{array} + \begin{array}{c}\textit{Expected mill-average}\\ \textit{sucrose content for the season}\end{array}$$

The sucrose deliveries (sucrose yield), on which monthly payments to the grower are based, are calculated as the product of weekly sugarcane deliveries (cane yield) and the weekly relative sucrose content.

Most of the natural seasonal variation in the absolute sucrose content of sugarcane (Van Dillewijn, 1952) is accounted for when the relative sucrose content is used. Although relative sucrose content is a more reliable index of performance than actual sucrose when comparing sucrose yields from fields harvested in different seasons (Hellmann, 1993), a number of serious problems are associated with the use of relative sucrose for yield comparisons. The expected mill average sucrose for the season is adjusted from time to time during the harvest season and as a result, relative sucrose content is an unreliable index of sucrose yield. In addition the mill-average sucrose content also varies from mill to mill and from season to season, distorting the data when relative sucrose yields from sugarcane crushed at different mills and during different seasons are compared. These factors must be borne in mind when using the sucrose yield data recorded in the field records.

Relative sucrose is an economic yield index, used to smooth the natural variations in the actual sugarcane sucrose content. Payments are based on the calculated relative sucrose content to encourage growers to deliver sugarcane to the mills rateably during the harvest season, irrespective of the actual sucrose content. Thus the sugarcane sucrose content is overstated during the months of naturally low sucrose content and understated during months of naturally high sucrose content, with the aim of expressing sucrose yield for the season as a function of cane yield only. At the end of the milling season accounting procedures are used to correct differences between the expected and actual mill-average sucrose content for the season to ensure that all payments are equitable.

Absolute sucrose, either considered in conjunction with harvest season or corrected for seasonal concentration differences using appropriate correction factors consistently, would provide more meaningful data for agronomic comparisons of sucrose yield. Although absolute sucrose content is measured and recorded on consignment notes by Cane Testing Services, only cane yield and relative sucrose content were recorded by Spencer Holley. The reason for omitting the remaining data from the field records is not clear. Absolute sucrose yields for the five Mondi estates were not captured in the SASA central industry database because Mondi do not participate in the FRS and have otherwise not specifically requested the consignment data to be captured (Hofmeyr, 1997, personal communication [19]). Consequently the relative field sucrose data recorded by Spencer Holley are the only sucrose yield data available.

---

[19] Mr K Hofmeyr, Infowave (Pty) Ltd
SASA Database Managers
PO Box 5341, Durban, 4000

The relative sucrose content recorded in the field records was calculated as the mean of the weekly relative sucrose content for those weeks during which each field was harvested. A cane-yield-weighted-mean of relative sucrose content was expected to be more accurate than the numeric mean of relative sucrose. This was tested using mill returns for Canema from the 1997 harvest season which had been captured in spreadsheet format on the estate. Both the numeric mean of the relative sucrose content (numeric mean) and the cane-yield-weighted-mean of relative sucrose content (weighted mean) were calculated (Appendix 4). The differences between the two calculated values are extremely small and confirm that the sucrose data recorded in the field records were determined from the mill returns with acceptable accuracy. The probable reason for the small difference between the two methods of calculating field sucrose content is that the weekly allocation of tons cane delivered is fairly constant throughout the harvesting season and, because consignment pay-loads are fairly uniform, a similar number of consignments is sampled for sucrose each week.

It was concluded that, in general, cane yields were more reliably recorded than sucrose yields.

3.10 CROP AGE

Crop age is incorporated in common sugarcane production indices (TCHM and TSHM). Besides providing the basis for cropping system productivity evaluations, crop age also influences foliar nutrient concentrations which are monitored to ensure adequate crop nutrition.

The crop cycle start and harvest dates were recorded as months only in the field records: the month of planting / harvesting was recorded as the month in which the task was completed. This recording time-scale was considered appropriate since larger fields may take a week or longer to harvest or plant. Separate months were not recorded for harvesting / ratooning because ratoon crop cycles start immediately after a field has been harvested.

3.10.1 Age at harvest

The cane yield of sugarcane generally increases as the age of the crop increases but the rate of growth decreases over time (Van Dillewijn, 1952). Consequently it is important to maximise

TCHM and TSHM rather than only TCH and TSH. To achieve this objective in the Midlands, sugarcane should usually be harvested between 18 and 22 months of age (Hellmann, 1993).

An assumed date, constant for all crop cycles, had to be used when estimating crop ages and secondary time-related factors, e.g. rain per crop cycle, because the exact dates of planting and harvesting were not recorded. As a result, crop age at harvest may be under- or over-estimated by up to one month, e.g. a crop ratooned during August 1990 and harvested during September 1992 (25 months) may range in age from 24 months (31 August 1990 to 1 September 1992) to 26 months (1 August 1990 to 30 September 1992).

All crop cycles were assumed to start (planted or ratooned) on the first day of the month recorded in the field records. Except for plant crops which are preceded by a fallow period, this assumption implies that harvesting the previous crop was completed on the ratoon day – effectively having grown to the end of the previous month. This approach was preferred to one of using the middle-of-the-month (fifteenth day) for the crop cycle start and harvest dates because only monthly rainfall records were reliably available for all estates (Section 3.2.1), and crop ages would not be determined with greater accuracy.

Since all crop cycles were assumed to start on the first day of the month recorded (even though the actual starting date may have been somewhat later in the month) rainfall allocated to the early stages of the crop cycles may be overestimated slightly. Similarly, the rainfall for the end of the crop cycles may be underestimated slightly because all crop cycles are assumed to be harvested on the first day of the month recorded, i.e. the last rainfall allocated to the cycle is for the previous month. Bearing in mind the limitations of the data and the potential of the crop to exploit the water, this allocation of rainfall to successive crop cycles is considered to be realistic.

3.10.2 Age at leaf sampling

Leaf samples from Midlands sugarcane must be taken between 3 and 9 months of age for the analytical results to be meaningful (Anon., 1991). Although Spencer Holley recorded the date of leaf sampling, the age at leaf sampling could be overestimated by up to 1 month because only the cycle starting month was recorded, e.g. a crop growing from August 1990 and leaf-sampled on

20 March 1991 may range in age from 6 months and 20 days (31 August 1990 to 20 March 1991) to 7 months and 20 days (1 August 1990 to 20 March 1991).

Only the month of leaf sampling was retained in the field records because the SASEX age classes are based on complete months, and Meyer *et al.* (1989) noted the importance of considering the month of leaf sampling in addition to crop age when interpreting leaf analyses. Since the starting date of crop cycles was assumed to be the first day of the month, leaf analyses were also assumed to apply to the first day of the month of sampling. Thus the crop age at leaf sampling is not underestimated, although it could be overestimated by up to one month. This approach was preferred to one of classifying leaf samples taken during the first half of a month, e.g. 1 March to 15 March, as the first day of that month (1 March), and those taken during the second half of a month, e.g. 16 March to 31 March, as the first day of the following month (1 April) because the potential overestimate of crop age at sampling was increased to one and a half months. Analyses were always considered in association with the crop age at sampling and the month of sampling. For this reason leaf samples taken from crops older than 10 or 11 months were not excluded from the records.

3.11 CROP CYCLES

Sugarcane harvested at different times of the year, and growing through different seasons, has vastly different growth potentials (Lonsdale and Gosnell, 1975; Landrey *et al.*, 1981; Sweet and Patel, 1985; Hellmann, 1988; Inman-Bamber, 1991a; Hellmann, 1993). While the sucrose yield of winter-harvested sugarcane is generally superior to that of summer-harvested crops (because of a higher sucrose concentration caused by natural ripening during winter), the regrowth potential of the winter ratoon crop is inferior to that of the summer ratoon crop. Inman-Bamber (1994a) found that the time taken to reach full canopy was affected significantly by variety and ratoon date main effects, and the variety × ratoon date interaction. The unfavourable winter growing conditions cause root and tiller mortality resulting in the establishment of smaller and weaker plant populations than for summer ratoons. Winter crops take up to three times longer than summer crops to reach full canopy and cane yields from crops starting in spring / summer are generally greater than those starting in autumn / winter (Inman-Bamber, 1994a).

Seasonal differences in sucrose yields for individual mills are largely accounted for when relative sucrose is considered in preference to absolute sucrose yields (Hellmann, 1988; 1993), but for cane yield comparisons to be meaningful, it is necessary to consider the season in which the crop cycle was started and harvested. Although sugarcane is normally harvested for milling from March / April to November / December in the Midlands, the field records include crop cycle starting and harvest dates spanning the entire year – where fields were harvested during January and February, either for seed or crushing at a mill also receiving Coastal sugarcane (Mount Edgecombe, Darnal and Glendale mills). This meant that crop cycles had to be defined for 12 starting and harvest months. Using an approach based on the work of Hellmann (1993), eight broad types of crop cycles were defined, viz.;

a crops ratooned during summer and harvested in summer (RS.S),
b crops ratooned during summer and harvested in winter (RS.W),
c crops ratooned during winter and harvested in summer (RW.S),
d crops ratooned during winter and harvested in winter (RW.W),
e crops planted during summer and harvested in summer (PS.S),
f crops planted during summer and harvested in winter (PS.W),
g crops planted during autumn and harvested in summer (PA.S),
h crops planted in autumn and harvested in winter (PA.W).

High radiation and temperature are required in addition to an adequate water supply for vigorous sugarcane growth. The duration of the active growing period varies from one season to another. In the Midlands the growing season is expected to start some time between August and October (Appendix 1 and 2). Consequently the crop cycles were defined on the assumption that 1 September was the start of summer across all estates for all years. Favourable growing conditions generally continue until the end of March, after which expected rainfall drops markedly (Appendix 2) and temperatures begin to decline (Appendix 1), inducing the sugarcane ripening phase. The end of summer was therefore assumed to be 1 April (31 March) for all estates. On the basis of these assumptions, the Midlands region experiences a seven month summer from September to March and a five month winter from April to August. For sugarcane ripening, both cool and dry conditions are required and the summer / winter harvest month derived from the crop cycle classification broadly corresponds with the natural low sucrose content months

(March, April / May, October, November, December) and the natural high sucrose content months (May / June, July, August, September).

The crop cycles were defined by distinguishing between ratoon crops and plant crops because of the different growth patterns exhibited by these crop classes (Thompson, 1988; Hellmann, 1993). Ratoon cycles (R) starting between 1 September and 1 March were classed as summer-start (S) crops and those from 1 April to 1 August as winter-start (W) crops. For plant crop cycles (P) a distinction was made between plantings before and after the Christmas shut-down. Plantings made between 1 September and Christmas (December) were grouped as summer-start (S) crops because they are generally expected to reach full canopy before winter. Plantings made after Christmas (January) to April were grouped as autumn-start (A) crops and are generally not expected to reach full canopy before winter. Similarly, crops harvested between 1 October (30 September) and 1 April (31 March) were classed as summer-end (.S) crops. Those cycles harvested from 1 May (30 April) to 1 September (31 August) were classed as winter-end (.W) crops [20] .

In addition to these broad crop cycle classes, the periods of active growth to which the crop is exposed generally ranges from one to four summers (*1S, *2S, *3S, *4S). The sugarcane age at harvest recorded in the field records ranged from nine months (frosted sugarcane) to 40 months. Since the Midlands summer is assumed to extend over seven months, all crop cycles growing through four to 10 summer months were grouped as a one-summer's-growth cycle (*1S), crops growing through 11 to 17 summer months as a two-summer's-growth cycle (*2S), crops growing through 18 to 24 summer months as a three-summer's-growth cycle (*3S) and crops growing through 25 to 31 summer months as a four-summer's-growth cycle (*4S) (Appendix 5). Crop cycles calculated to have been harvested after less than four summer months' growth were investigated because a harvestable yield cannot be produced over such a short growth period in

---

[20] For harvesting, the summer is assumed to extend from 1 October to 1 April because only the month of harvesting / ratooning was recorded in the field records (Section 3.10.1), and since new crop cycles are assumed to have started on the first day of the month recorded in the field records, the previous crop cycle is assumed to have grown to the end of the preceding month.

the Midlands. All the possible crop cycles resulting from this classification procedure are presented in Appendix 5 together with the total number of summer and winter months accumulated for each of these cycles. Detailed schedules for each crop cycle are also presented.

Crop cycles were classified according to the season in which they were started and harvested, and the expected number of months of active growth (total summer months, Appendix 5) for each crop cycle were also recorded with the field records. This classification is useful for grouping cycles which experienced similar growth patterns for meaningful yield comparisons and has been used successfully by Landrey *et al.* (1981) and Hellmann (1988; 1993).

4 SUGARCANE YIELD MODELS

The field records were used to compile a database in which each crop cycle comprised an observation. All incomplete crop cycles were excluded from the database. These comprised most of the first crop cycles recorded for each field, which had harvest data only, and the current crop cycles, which had records of agronomic inputs only. Where obvious recording errors could not be corrected with the assistance of the estate managers, or where it was impossible to remove bias from the data (predominantly crop cycles affected by seed-cane harvests), these cycles were also excluded. In addition, all crop cycles from three fields at Umvoti were excluded because area biases could not be corrected. The resulting data set of 984 observations, accumulated over 19 harvest seasons (1979 – 1997), was considered potentially suitable for statistical examination.

The distribution of these observations between the three districts is shown in Table 10. Although the largest area under sugarcane is situated at Umvoti, this district has the smallest number of fields resulting in relatively few observations (Section 3.1). The number of observations at Umvoti was further reduced because the partial harvesting of fields for seed-cane was most prevalent on the Canema estate. Of the 984 crop cycles, 412 observations could be comprehensively checked for accuracy (Section 3.8). These records were coded as a subset of data separate from the remaining 572 observations because the accuracy of these records had been confirmed (Table 10).

Table 10: Summary of all the complete crop cycle observations potentially available for statistical analysis in the three districts.

District	Number of fields	Total number of crop cycles	Checked data		Remaining data	
			Number of observations	Harvest seasons	Number of observations	Harvest seasons
Kranskop	55	417	144	1989 – 1997	273	1981 – 1993
Umvoti	31	134	73	1990 – 1997	61	1985 – 1993
Richmond	57	433	195	1989 – 1997	238	1979 – 1993
Total	**143**	**984**	**412**	**1989 – 1997**	**572**	**1979 – 1993**

In spite of extensive editing, the quality of the agronomic data in the field records is poor when compared with data from research trials. The problems, many of which are inherent in the data, have been described in detail (Chapter 3) and must be considered when the field records are used to derive management information. They are, however, not atypical of good commercial records (Hellmann, 1988; 1993; Inman-Bamber *et al.*, 1993) and the extent of the Mondi records is impressive when compared with other field record systems in the South African sugarcane industry.

4.1 PRELIMINARY INVESTIGATIONS

A preliminary investigation of the data set at the estate level using scatter plots showed none of the expected relations between cane yield and harvest age, N fertilizer or rain total accumulated for each crop cycle (Appendix 6) [21]. Cane yield had been expected to increase with increasing age, possibly at a decreasing rate as the crop became older. While Figure A6.1a (Appendix 6) shows that relatively low yields were generally achieved where sugarcane was harvested at very young ages (< 80 TCH), and relatively high yields where fairly old sugarcane was harvested (> 90 TCH), both very low and very high yields were achieved across all the dominant age classes. For most of the observations the crops were harvested between 17 and 32 months of age. Sugarcane harvested at younger ages represents frosted crops (frost damage class 3, Section 3.5.5) and immature crops destroyed by wild fire or harvested to meet guaranteed crop estimate commitments at the mills in certain seasons. Sugarcane harvested at older ages represents carry-over sugarcane from seasons where a portion of the mature crop could not be harvested for various reasons. Both the very young and very old crops were produced predominantly at Richmond (Figure A6.1a).

Similarly, cane yield was expected to increase with increasing application rates of N fertilizer, recorded for 887 of the 984 observations. For these observations cane yield was plotted against N fertilizer inputs (Figure A6.2a, Appendix 6). The large range in N fertilizer applications (36 – 225 kg ha^{-1}) was the result of applying Spencer Holley recommendations to incorrect field areas,

---

[21] These Figures would normally be included sequentially in the text but have been presented in an appendix to also allow a comparison between the "full" and "restricted" data sets (see Section 4.1.1).

and also from the omission of prescribed top-dress applications of Urea in numerous plant crop cycles. The plot reveals a random scatter of points (Figure A6.2a). This distribution of yield response to N may be expected where inputs are managed to optimally supply the crop with nutrients. It should, however, not be expected to apply in this situation because of the large range of N inputs.

The expected relation of cane yield with harvest age and N fertilizer may partly have been masked by considering drought years together with high rainfall years. This is supported by the graph of cane yield versus rain total accumulated for each crop cycle (Figure A6.3a, Appendix 6) where a weak positive relation between yield and rain is apparent. An additional complication arises from the fact that these simple two-way plots ignore all other variation inherent in the data, further masking trends that might be present in the data.

The relation between cane yield and rain total accumulated for each crop cycle was expected to be particularly strong because it forms the basis of the Smith yield model for dryland sugarcane yield predictions (Section 2.3.2). Since this was not the case (Figure A6.3a), it became clear that it would be necessary to restrict the data for this study using exclusion rules determined on an *a priori* basis, to reduce the noise inherent in the full data set. This approach is also justified because records of important sugarcane management variables are missing, e.g. fertilizer inputs for 97 observations.

4.1.1 Restricting the data

The crop cycle classification (Section 3.11) provided a useful starting point for restricting the data. Since different crop cycles have inherently different growth potentials, it is essential to consider each cycle separately when yield comparisons are made (Hellmann, 1988; 1993). It follows that crop cycles should be expected to have contributed towards masking the anticipated yield trends in Figures A6.1a, A6.2a and A6.3a (Appendix 6).

The following *a priori* exclusion rules were applied to the 984 observations:

a exclude all crop cycle types that have less than 20 observations;

b exclude all observations without fertilizer input records;

c exclude all observations without topsoil fertility analyses;

d exclude all observations for which agronomic inputs differ by more than 10% in the estate managers' and Spencer Holley records, and could not be corrected when checking the records; and

e exclude all observations for mixed sugarcane varieties, and varieties with less than 20 observations.

Each exclusion rule is discussed separately in the following paragraphs, demonstrating how the size of the database was reduced by the successive application of these five criteria. In all cases where observations were excluded from the data set, the unrepresentative groupings could be clearly identified without having to use the arbitrarily defined minimum sample size of 20 observations.

a *Exclude all crop cycle types that have less than 20 observations*

An examination of the data showed that the data set comprised 20 different crop cycle types, eight of which were for plant crops only. The distribution of plant crops between crop cycle type and district (Table 11) showed that the majority belonged to one of only two crop types.

Table 11: Summary of plant crop cycles, according to district, potentially available for statistical analysis.

Crop cycle type	District			**Total number of observations**
	Kranskop	Umvoti	Richmond	
PA.S*1S	4	0	1	5
PA.W*1S	0	0	4	4
PS.S*1S	1	0	0	1
PA.S*2S	1	0	2	3
PA.W*2S	2	1	3	6
PS.S*2S	32	10	20	62
PS.W*2S	21	3	23	47
PS.W*3S	0	0	6	6

Under rainfed conditions spring planting of sugarcane is recommended except where severe frost is expected (Mann, 1991). Since these parts of the Mondi estates are generally not planted to sugarcane, the predominance of spring plantings in the field records points towards sound management of the crop. In line with sugarcane management recommendations for the Midlands (Hellmann, 1988), the plant crops were harvested predominantly after two summers' growth (Table 11). The data set was therefore reduced by 25 observations to include only PS.S*2S and PS.W*2S plant crop cycles.

For the ratoon crops, two summers' growth cycles also dominate the field records and consequently most of these crop cycles belonged to one of only four types (Table 12). Thus a further 53 observations were removed, and only RS.S*2S, RS.W*2S, RW.S*2S and RW.W*2S crop cycles were retained in the data set (Table 12). This means that only crop cycles with harvest ages ranging in extreme from 16 months for RS.S*2S and PS.S*2S cycles to 32 months for RW.S*2S, RS.W*2S and RS.S*2S cycles (Appendix 5) are represented in the restricted database.

Table 12: Summary of ratoon crop cycles, according to district, potentially available for statistical analysis.

Crop cycle type	District			**Total number of observations**
	Kranskop	Umvoti	Richmond	
RS.S*1S	4	0	0	**4**
RS.W*1S	1	0	4	**5**
RW.S*1S	12	3	3	**18**
RW.W*1S	0	0	3	**3**
RS.S*2S	47	13	53	**113**
RS.W*2S	120	35	111	**266**
RW.S*2S	54	7	35	**96**
RW.W*2S	113	62	147	**322**
RS.S*3S	4	0	3	**7**
RS.W*3S	0	0	3	**3**
RW.S*3S	1	0	4	**5**
RW.W*3S	0	0	8	**8**

The six two-summer crop cycles for plant and ratoon crops that were retained in the data set represent the dominant crop cycles produced on the Mondi estates. Since these cycles correspond to those recommended for the Midlands, the general crop cycle management on the estates is sound even though harvest ages often exceed the recommended 18 to 24 months. The 14 crop cycle types excluded from the data set are produced infrequently, and only under exceptional circumstances. Consequently they only contribute noise in an investigation aimed at determining the driving variables of the production system and no useful information was lost as a result of their exclusion from the investigation.

b *Exclude all observations without fertilizer input records*

Fertilizer applications to sugarcane are important in terms of crop management and production costs. Crop nutrient supply, as manipulated by fertilizer inputs, defines a basis for crop management using agronomic records. Consequently little can be concluded from the yields recorded for those observations where this information is not available. Thus a further 87 observations were removed from the data set for which no fertilizer applications had been recorded.

c *Exclude all observations without topsoil fertility analyses*

When evaluating and analysing the sugarcane management system it is important to consider fertilizer applications in conjunction with topsoil fertility analyses. This is because fertilizer inputs are generally determined on the basis of soil nutrient levels (Section 3.8.1). Although Spencer Holley Agronomic Services insisted on a topsoil fertility analysis for each crop cycle from every field for which fertilizer recommendations were made from the 1986 season onwards, 28 cycles were managed without a soil analysis where exceptional circumstances made it impossible to submit a soil sample. These observations were excluded from the database. For many of the earlier Mondi field records (crop cycle starting seasons from 1977 to 1985) which Spencer Holley compiled from historic documents, no soil analyses were available. These 154 crop cycles were generally managed using "blanket" applications of fertilizer mixes, ignoring the specific nutrient requirements of individual fields. Although each of these crop cycles received very similar treatments on the basis of fertilizer inputs alone, yield relations will have been distorted by over and under fertilization of the crop caused by differences in natural soil fertility.

They were therefore also removed from the data set, reducing the total number of observations quite substantially.

d *Exclude all observations for which agronomic inputs differ by more than 10% in the estate managers' and Spencer Holley records, and could not be corrected when checking the records*

When the field records were reworked into an appropriate structure for this study, the duplicated data for each observation were compared to check for transcription errors. For most of the crop cycles the estate managers' chronological records corresponded exactly to the field history records maintained by Spencer Holley (Chapter 3). Where discrepancies were found, they were usually very small, apparently caused by rounding errors. Within this context, 29 observations were found for which the fertilizer inputs recorded in the two record subsections differed by more than 10%. An additional nine observations were found for which the fertilizer types, and consequently the quantity of plant nutrients applied, differed. Since the inconsistencies could not be resolved with the assistance of the estate managers, these 38 observations were removed from the data set.

e *Exclude all observations with mixed sugarcane varieties, and varieties with less than 20 observations*

Four main sugarcane varieties were produced on the Mondi estates during the study period. The full data set, however, comprised 12 different classes defined for seven varieties and various combinations of these (Table 13). For observations from fields growing more than one sugarcane variety, the relative proportion of each variety was not recorded and therefore the relative performance of the varieties could not be estimated. For NCo382 and N14 there were only two and four observations respectively, while for N13 there were only observations where it was grown together with another variety. All observations for mixtures of the sugarcane varieties, and those varieties with less than 20 observations, were removed from the data set. On this basis observations for only N12, NCo293, NCo376 and N16 were retained and 34 observations representing the other eight classes (Table 13) deleted.

Table 13: Summary of the sugarcane varieties and combinations of varieties included in the original 984 crop cycles from the Mondi field records.

Sugarcane variety	District			**Total number of observations**
	Kranskop	Umvoti	Richmond	
N12	88	66	61	**215**
NCo293	86	38	348	**472**
NCo376	178	18	6	**202**
N16	33	5	4	**42**
NCo382	0	2	0	**2**
N14	0	0	4	**4**
NCo293 + NCo376	12	5	0	**17**
N13 + N16	0	0	6	**6**
N13 + N12	0	0	4	**4**
N12 + NCo293	7	0	0	**7**
NCo293 + N12 + N16	7	0	0	**7**
N12 + N16	6	0	0	**6**

The application of the five exclusion rules resulted in the data set being reduced from 984 to 565 observations. Depending on the sequence in which the rules were applied, a variable number of observations were excluded because more than one exclusion condition applied to many of the observations. However, the number of observations retained in the data set remained constant, irrespective of the sequence in which the data were restricted. These data were checked to ensure that no observations were retained that had become unrepresentative of the data set as a result of the 419 exclusions. Checks were made in terms of harvest season, physical field resources and crop management strategies.

All observations for the 1979, 1980 and 1982 harvest seasons were excluded from the data set, mainly because either the fertilizer applications or topsoil fertility analyses had not been recorded (exclusion rules b and c). Only one observation for each of the 1981 and 1983 harvest seasons, and three observations for the 1984 harvest seasons were retained. These five observations were removed from the data set because the subsets of data were too small to adequately describe the seasonal variation in sugarcane yields illustrated by Hoekstra (1976). Consequently the data set spanned the period from the 1985 harvest season to the 1997 harvest season (13 seasons). The starting dates of these crop cycles spanned the period 1983 to 1995 because only crops with two summers' growth were included in the database. Thus the database included one full seasons' records originally extracted from Mondi documents – crops starting during 1983 and harvested during 1985, two seasons recorded partially in Mondi documents and partially by Spencer Holley – crops starting during 1984 and 1985, harvested after Spencer Holley Agronomic Services became responsible for the field records in 1986, and 10 seasons' data recorded only by Spencer Holley Agronomic Services.

The range of physical field resources represented in the restricted database was checked using a GIS (ARCVIEW version 2.1). All resources were well represented by these observations (>20 observations for each category) except for two soil groupings, one of which was represented by four observations from only one field (Field AS09 on Uplands at Richmond; Orthic yellow soil type), the other by 13 observations from two fields (Fields AS07 and AS08 on Uplands at Richmond; Orthic red / yellow soil type). The GIS analysis showed that the three fields were neighbours. All 17 observations were excluded from the data set on the basis of the unusual soil resources they represented in relation to the available data, and their spatial association with each other.

Crop management strategies were fairly constant across all observations retained in the data set. In terms of crop cycles, only crops growing through two summers were included, thus reducing some of the variability in crop harvest ages. Although the types of fertilizer applied varied over time and a considerable number of different herbicide active ingredients were used, all crops in the data set were fertilized according to topsoil analysis, often supported by a leaf analysis to confirm adequate nutrition, and weeds were controlled chemically in the first instance and thereafter by hoeing. It was therefore not necessary to exclude additional observations because of unrepresentative crop management practices.

The sequence of data restriction and the number of observations excluded from the data set by each condition is presented in Table 14. A total of 441 observations (45% of the data) were excluded for the reasons described above. This considerable reduction was unfortunate, but essential to ensure meaningful comparisons. Since the Mondi data are typical of good commercial records, large data sets need to be available when commercial sugarcane records are used for scientific studies because one can expect to exclude up to half the observations.

Table 14: Restriction of the original data set of 984 observations.

Data subsets			Number of observations
Original data			984
Exclusion rules	a	Unrepresentative crop cycles	-78
	b	No fertilizer records	-87
	c	No topsoil fertility analyses	-182
	d	Recording inconsistency of >10%	-38
	e	Unrepresentative varieties or variety combinations	-34
Subsequent adjustments	Unrepresentative seasons (1981, 1983, 1984)		-5
	Unrepresentative orthic yellow soil type		-4
	Unrepresentative orthic red / yellow soil type		-13
Restricted data set suitable for statistical analysis			**543**

Adjustments made subsequent to restricting the data resulted in a data set of 543 observations (Table 14). The distribution of these observations across the three districts, and between the checked and remaining data subsets, is presented in Table 15. The effect of the data restriction can be seen by comparing Table 10 with Table 15. The checked data set was reduced from 412 to 345 observations while the remaining data set was reduced from 572 to 198 observations. It follows that besides excluding observations unsuitable for the investigation, the general quality of the data was improved by the restrictions since the relative proportion of checked data increased from 42% to 64%.

Table 15: Summary of the 543 observations selected for statistical analysis in the three districts from the full data set of 984 observations.

District	Number of fields	Total number of crop cycles	Checked data		Remaining data	
			Number of observations	Harvest seasons	Number of observations	Harvest seasons
Kranskop	53	229	123	1989 – 1997	106	1985 – 1993
Umvoti	28	102	68	1990 – 1997	34	1987 – 1992
Richmond	50	212	154	1989 – 1997	58	1985 – 1989
Total	**131**	**543**	**345**	**1989 – 1997**	**198**	**1985 – 1993**

The scatter plots of cane yield versus harvest age (Figure A6.1b), cane yield versus N fertilizer (Figure A6.2b) and cane yield versus rain total accumulated for each crop cycle, using the restricted data are shown in Appendix 6. By comparing Figures A6.1a with A6.1b, A6.2a with A6.2b and A6.3a with A6.3b, the effect of the restriction on these relations is evident. Although the ranges of harvest age and N fertilizer application have been reduced considerably, no trends are apparent and the points remain essentially a random scatter. The relation between cane yield and rainfall is stronger for the restricted data set ($r = 0.408$) than for the complete data set ($r = 0.388$) – compare Figure A6.3b with A6.3a.

4.1.2 Classification of the variables

For each observation all available agricultural resource and production data, and the various indices and ratios calculated from these data (Chapter 3), were included in the database. Where a particular parameter was not available for an observation, it was coded as a missing value. For the statistical analyses the parameters were classified as being either dependent variables or independent (predictor) variables.

4.1.2.1 *Dependent variables*

The four yield indices, TCH, TSH, TCHM and TSHM (Section 2.3) were defined as dependent variables since the objective of the study is to determine key parameters strongly associated with actual sugarcane yields in field records from the Mondi estates in the Midlands.

The correlation matrix indicating strong relations between the yield indices was calculated for the 543 observations (Table 16). To confirm that the extensive data restriction (Section 4.1.1.1) had not distorted the nature of the data, these correlation coefficients (r) were also calculated for the original data set of 984 observations.

Table 16: Correlation matrix for TCH, TSH, TCHM and TSHM using the selected 543 observations. The corresponding correlation coefficients for the original data set of 984 observations are shown in brackets.

	TCH	TSH	TCHM	TSHM
TCH	1.000 (1.000)	0.946 (0.943)	0.923 (0.882)	0.900 (0.859)
TSH	0.946 (0.943)	1.000 (1.000)	0.848 (0.805)	0.927 (0.896)
TCHM	0.923 (0.882)	0.848 (0.805)	1.000 (1.000)	0.950 (0.941)
TSHM	0.900 (0.859)	0.927 (0.896)	0.950 (0.941)	1.000 (1.000)

The high correlation between TCH and TSH was anticipated since relative sucrose was recorded in the field records. The aim of the relative sucrose measure is to express sucrose yield for the season as a function of cane yield only (Section 3.9.2) and the correlation coefficient of 0.946 confirms that this is usually achieved. The restricted data set shows the same correlation characteristics as the original data (Table 16).

It may seem surprising that TCH and TCHM are highly correlated ($r = 0.923$). Cane yield per unit area (production) is not related to the cane yield per unit time (productivity), especially since yield increases at a decreasing rate for older crops, i.e. TCH increases but TCHM decreases. However, the crop cycles are managed to harvest the bulk of the sugarcane within a fairly restricted age range according to current recommendations (Hellmann, 1993), and therefore high TCH observations are associated with high TCHM observations. This tendency is strengthened by the normal management approach of harvesting high yielding fields at a younger age than low yielding fields (Inman-Bamber, 1994b). The correlation between TCH and TCHM is stronger for the restricted data set than the original data (Table18) because the range in harvest ages was reduced from 26 to 15 months by limiting the investigation to crop cycles with two summers' growth only (exclusion a, Section 4.1.1).

As for the TCH:TCHM relation, TCH and TSHM are highly correlated, and the restricted range of harvest ages in the selected data set resulted in a slightly higher correlation coefficient being calculated than for the original data set (Table 16). Had measures of absolute sucrose content been recorded in the field records, these relations would not be expected to hold because of the considerable seasonal variation in sucrose concentration which is independent of cane yield. While the high correlation between the cane and relative sucrose yields shows that the relative sucrose yield indices do not provide much additional management information, absolute sucrose yields would be useful for investigating the effects of different crop cycles and lengths of milling season on sucrose yields. These crop management issues are becoming increasingly important as sugarcane supply agreements are being negotiated with millers in the recently deregulated South African sugar industry.

Since TCH is highly correlated with TSH, TCHM and TSHM ($r > 0.9$) it was decided to only investigate the relation between TCH and the independent variables because:

a the other three yield indices provide little additional management information since they are highly correlated with TCH and with each other;

b based on the correlation coefficients, a simple linear regression equation can be developed using TCH as the only independent variable to predict TSH, TCHM and TSHM with estimated R^2 values of 0.89, 0.85 and 0.81 respectively, without considering the complex agricultural resource and production data;

c some inherent data quality problems are associated with the sucrose yield data (Section 3.9.2); and

d the TCH measure is the yield index that is most meaningful to managers.

The distribution of TCH was confirmed to be almost normal and summary statistics for TCH, and those for TSH, TCHM and TSHM, are presented in Appendix 7. Thus TCH can be used with confidence in statistical analyses.

4.1.2.2 *Independent variables*

All agricultural resource data and agronomic variables recorded for each observation were used as independent variables in analyses to describe sugarcane yields (TCH). Two broad categories of independent variables were defined, viz. categorical variables and continuous variables. Qualitative variables were defined as categorical variables, expressed in dummy variable format. In multiple regression analyses these variables generally function as intercept shifters and only influence the partial regression coefficients within the model when they interact with continuous variables. Categorical variables group the data into subsets according to qualitative conditions describing some of the variation in natural field conditions (Hellmann, 1988) and separate models are calculated for each data group. For any particular condition, e.g. location, the data set could be grouped into a number of levels, depending on the scale chosen. Thus locality could be coded as a dummy variable at the district scale (three levels) or at the estate scale (five levels). Care was exercised not to consider aliased categorical variables simultaneously. When two or more categorical variables are considered simultaneously in regression analyses, the data set is subdivided into groups in such a manner that all conditions specified by the levels are met, i.e. the number of models increase multiplicatively according to the number of levels per categorical variable. The levels defined for each categorical variable and the number of observations per level are presented in Appendix 7. The remaining variables were all defined as continuous variables and their summary statistics are presented in Appendix 7.

4.2 REGRESSION ANALYSIS

Linear regression analysis based on the method of least squares was used to establish a functional relation among the variables. Analyses were performed using the computer software package GENSTAT version 5, release 3.2. Unless specified otherwise, program default settings were used for the options exercised. In order to achieve the two main objectives of the study, viz.

(a) determine the key variables influencing commercial sugarcane yields using field records, and

(b) predict commercial sugarcane yields using field record data,

two general modelling approaches were used. To achieve objective (a), a regression model was developed to account for as much of the observed variation in TCH as possible, screening all available independent variables (Appendix 7). This model is referred to as the "explanatory model". To achieve objective (b), a second regression model was developed, also aiming to account for as much of the observed variation in TCH as possible, but screening only those independent variables known at least six months in advance of the planned harvest date. This model is referred to as the "prediction model".

Summary statistics of the variables (Appendix 7) were used to check for errors, and ranking the observations for each variable was useful for tracing outliers. Despite having restricted the data set, missing values for many of the continuous variables precluded their use as potential predictors, e.g. topsoil acid saturation analyses for only 44 of 543 observations. Scatter diagrams of TCH on the independent variables were also generated. While these simple plots are inadequate for illustrating the relation between the dependent and independent variable when more than one independent variable is considered, they were useful for showing peculiarities of the data set that had to be considered when analysing the data, e.g. very high applications of P fertilizer to plant crops and consistently lower applications to ratoon crops (Figure A6.4, Appendix 6), and for checking for possible outliers, e.g. very low yield response to high rainfall for one observation at Kranskop (Figure A6.3b, Appendix 6).

Correlation coefficients were calculated to gain some understanding of the relations among variables. Correlation matrices were calculated for subsets of the continuous data, grouped according to common missing values. Correlation coefficients between TCH and each continuous variable were also calculated. The only variables with correlation coefficients greater than 0.30 with TCH were the rain variables (rain total for cycle, and rain for 12 months' growth). Independent variables with high bivariate correlations among themselves may be restricted prior to regression analysis to avoid the unnecessary inclusion of groups of variables expressing similar influences on TCH, i.e. multicollinear variables. Rather than excluding correlated variables, they were grouped and tested as alternatives for each other in regression analyses, e.g. topsoil clay content and subsoil clay content. This was done because the simple correlation test for multicollinearity can be misleading since the problem is not as simple as collinearity between two variables and multicollinearity usually also exists among and between groups of variables, in which case detection is not simple.

4.2.1 Selection of independent variables

The most parsimonious yet relevant explanatory and prediction models were developed using a manual forward selection procedure. The use of automated selection procedures, e.g. stepwise and all subsets regression, was precluded by the large number of categorical variables (Appendix 7). Since GENSTAT excludes all observations with a missing value for any variable included in the terms statement, a data set of continually varying size was used depending on which variables were included in an analysis. Careful evaluation was therefore required when different models were compared because the final model should be based on as large a sample as possible.

4.2.2 Model building

In this section regression models are presented in tabular form, showing the regression coefficients, their standard deviations, the t-values of the regression coefficients and significance levels. The coefficient of multiple determination (R^2) and standard error of estimate (SE) for the models are also shown. The aim of model building is to achieve a high R^2 and low SE, matching predicted to measured yields as closely as possible, using the smallest number of easily measured, readily comprehendible, independent variables. A regression model should preferably contain only those variables with the highest possible levels of significance. Significance levels ranging from 10% to 25% are often used as the exclusion criterion in regression models in preference to standard significance levels (1% or 5%) (Schutz 1990). Deciding which significance level to use therefore involves a compromise between including meaningful explanatory variables and increasing the Type I error of the model. For this study the 10% level of significance was used.

The explanatory and prediction models were developed in close association with each other during the model building phase. For presentation purposes, the models are discussed separately, both to retain clarity and demonstrate the logical processes that were followed to derive the final models. A term used to describe a variable will be printed in *italics* when it refers to a variable in statistical analysis.

4.2.2.1 *The explanatory model*

Using the approach described in Section 4.2.1, variables were selected for the preliminary explanatory model (Model 1.0). The model is presented in Table 17.

Table 17: Model 1.0 Regression coefficients for the preliminary explanatory model.

Categorical variables	Levels	Regression coefficient differences	Standard deviation	t-value (507 d.f.)	Probability
Constant	Reference categories †	141.7	33.3	4.25	<0.001
Season	1985	-	(Reference category)		
	1986	2.88	6.39	0.45	0.652
	1987	8.62	5.09	1.69	0.091
	1988	-2.17	5.15	-0.42	0.674
	1989	2.51	5.11	0.49	0.624
	1990	-16.26	5.46	-2.98	0.003
	1991	-8.29	5.23	-1.59	0.113
	1992	-9.66	5.93	-1.63	0.104
	1993	-31.00	6.38	-4.86	<0.001
	1994	-17.80	5.89	-3.02	0.003
	1995	-14.34	5.60	-2.56	0.011
	1996	-29.07	5.84	-4.97	<0.001
	1997	-13.96	5.70	-2.45	0.015
Soil type	Orthic red	-	(Reference category)		
	Humic red	-6.57	3.11	-2.11	0.035
	Humic yellow	-12.17	3.48	-3.50	<0.001
	Humic red / yellow	-14.48	3.79	-3.82	<0.001
Variety	N12	-	(Reference category)		
	NCo293	-8.25	3.24	-2.55	0.011
	NCo376	-11.28	3.76	-3.00	0.003
	N16	-4.09	3.74	-1.09	0.275
Plant / Ratoon crop	Plant	-	(Reference category)		
	Ratoon 1	-5.68	3.57	-1.59	0.112
	Ratoon 2	-5.42	3.65	-1.48	0.139
	Ratoon 3	-11.09	3.93	-2.82	0.005
	Ratoon 4	-11.28	4.15	-2.72	0.007
	Ratoon 5	-13.66	4.58	-2.99	0.003
	Ratoon 6	-17.80	4.85	-3.67	<0.001
	Ratoon 7	-19.91	5.88	-3.38	<0.001
	Ratoon 8	-24.00	10.40	-2.32	0.021
	Ratoon 9	-12.60	11.80	-1.07	0.283
District	Kranskop	-	(Reference category)		
	Umvoti	-12.92	3.34	-3.87	<0.001
	Richmond	1.40	4.14	0.34	0.735
Crop cycle	Summer harvest	-	(Reference category)		
	Winter harvest	-4.28	2.03	-2.10	0.036
Continuous variables		**Regression coefficient**	**Standard deviation**	**t-value (507 d.f.)**	**Probability**
N fertilizer (kg ha^{-1})		0.255	0.051	5.06	<0.001
Rain total for cycle (mm)		0.010	0.003	3.03	0.003
Altitude (m a.s.l.)		-0.067	0.028	-2.42	0.016
Total K (kg ha^{-1})		0.014	0.005	3.05	0.002
Profile gravel (%)		-0.285	0.144	-1.97	0.049
$R^2 = 0.519$	SE = 18.1 TCH				

† The constant term refers to the base yield for the reference categories.

Refining the season – yield relation

The regression coefficient differences for *season* (Table 17) show considerable yield variations between seasons, with yields during the 1990's generally being lower than during the 1980's. It is only possible to speculate about the cause of the observed seasonal yield variations. Hoekstra (1976) attributed seasonal yield variations to the general climate. In Model 1.0 differences in total rainfall have been accounted for and other climate parameters, e.g. biannual temperature and radiation differences (for crops growing through two summers), could be incorporated in *season*. Appropriate climatic data are not available to confirm this. Using the CANEGRO model, Inman-Bamber *et al.* (1998) predicted large differences in attainable yields under rainfed conditions in South Africa. These yield differences were attributed primarily to differences in rainfall. Potential yields, limited only by radiation and temperature were predicted to be fairly constant, the worst year in 25 years differing by only 15 TCH from the best year. It is disturbing that the regression coefficient differences suggest seasonal yield differences of more than double this amount and these trends warrant further research.

The regression coefficient differences for some season levels were not statistically significantly different from each other (Table 17) and categories were combined where appropriate. In this manner it was possible to reduce the number of levels from 13 to five. The changes improved the explanatory model since the R^2 remained essentially unchanged (decreasing from 0.519 to 0.516), the degrees of freedom increased by eight and the model SE decreased slightly (from 18.1 TCH to 18.0 TCH). These statistics can be used as an indication that cane yields did not differ between the seasons that were grouped. The final *season* groupings presented in Table 18 were derived from a number of successive groupings of the 13 season levels and are presented in a ranked order, with the poorest seasons as the reference category, to avoid regression coefficient differences with a negative sign.

Table 18: Regression coefficient differences and related statistics for *season* in the revised explanatory model based on grouped seasons using five levels.

Categorical variable	Levels	Regression coefficient differences	Standard deviation	t-value (515 d.f.)	Probability
Season	1 (1993, 1996)	-	Reference category		
	2 (1990, 1994, 1995, 1997)	14.10	2.36	5.97	<0.001
	3 (1991, 1992)	20.90	2.93	7.14	<0.001
	4 (1985, 1988)	27.13	3.83	7.09	<0.001
	5 (1986, 1987, 1989)	34.21	3.29	10.39	<0.001

Refining the soil type – yield relation

Regression coefficient differences for humic yellow and humic red / yellow soils in Model 1.0 are almost equal and require closer examination. In general, it is important to bear in mind peculiarities of the data set which may influence interpretations of statistical trends when analysing the data. As shown in Table 19, the four soil types are not uniformly distributed across the study area.

Table 19: Distribution of the four soil types across the three regions.

Soil type	Number of observations			
	Kranskop	Umvoti	Richmond	**Total**
Orthic red	0	0	49	**49**
Humic red	48	102	96	**246**
Humic yellow	142	0	43	**185**
Humic red / yellow	39	0	24	**63**

When grouping soil types, care was exercised not to group soils that differ in their spatial distribution. Humic yellow and humic red / yellow soils were grouped on the basis of their similar regression coefficients (Table 17), bearing in mind that both soil types were represented at Kranskop and Richmond. Had groupings of the remaining soil types been statistically justified these would have been rejected because of their different spatial distributions. The goodness of fit of the revised model remained largely unchanged, the R^2 decreasing to 0.515 with a SE of 18.0 TCH. The revised regression coefficient differences for *soil type* are presented in ranked order in Table 20 to avoid coefficients with a negative sign.

Table 20: Regression coefficient differences and related statistics for *soil type* in the revised explanatory model based on grouped soil types using three levels.

Continuous variable	Levels	Regression coefficient differences	Standard deviation	t-value (516 d.f.)	Probability
Soil type	**1** (Humic yellow and red / yellows)	-	Reference category		
	2 (Humic red)	6.46	2.28	2.84	0.005
	3 (Orthic red)	12.69	3.30	3.85	<0.001

Refining the variety – yield relation

The regression coefficient differences and probabilities for *variety* in Model 1.0 (Table 17) suggest that the two NCo varieties are both significantly different from N12, while N16 is shown to be not significantly different from N12. Unfortunately there are only 36 observations for N16 (Appendix 7) and, as a result of the small sample size, the standard deviation of its regression coefficient is so large that no significant difference between N16 and the NCo varieties was calculated. Consequently the two NCo varieties were grouped and the effect of allocating N16 to either this group or N12 tested. The best result, based on model R^2 and SE, was obtained by grouping NCo293 with NCo376 and N12 with N16. This grouping also makes sense in terms of practical agriculture because N12 and N16 together represent the "new" generation varieties for the Midlands, while NCo293 and NCo376 together represent the "old" varieties in the field records. The grouping resulted in a smaller than expected reduction in the R^2 (0.514) and model SE remained unchanged, confirming that it was justified in terms of sugarcane yield response. The regression coefficient differences for the grouped varieties are presented in Table 21.

Table 21: Regression coefficient differences and related statistics for *variety* in the revised explanatory model based on grouped varieties using two levels.

Continuous variable	Levels	Regression coefficient differences	Standard deviation	t-value (518 d.f.)	Probability
Variety	**1** (NCo376 and NCo293)	-	Reference category		
	2 (N16 and N12)	8.60	2.79	3.08	0.002

Refining the plant / ratoon crop – yield relation

As expected, the regression coefficient differences for *plant / ratoon crop* (Table 17) show a general yield decline associated with increasing ratoon number (Section 3.6.6). Since a number of the coefficients are of similar magnitude, levels were grouped where appropriate. In this manner the original 10 levels were reduced to three (Table 22). The grouping of successive crops in each new level is fortuitous, although this situation is desirable because it allows statements to be made about practical crop management decisions – it is not possible to produce the older ratoon crops without first having produced the younger ones. For this reason also, the levels are presented in order from youngest to oldest, showing negative regression coefficient differences which indicate the yield sacrifice expected on average, when older crop classes are produced (Table 22). The R^2 for the revised explanatory model decreased to 0.509 and the SE decreased to 17.9 TCH.

Table 22: Regression coefficient differences and related statistics for *plant / ratoon crop* in the revised explanatory model based on grouped crop categories using three levels.

Categorical variable	Levels	Regression coefficient differences	Standard deviation	t-value (525 d.f.)	Probability
Plant / Ratoon crop	1 (Plant, Ratoons 1 and 2)	-	Reference category		
	2 (Ratoons 3, 4 and 5)	-7.05	2.14	-3.29	0.001
	3 (Ratoons 6, 7, 8 and 9)	-13.84	3.06	-4.52	<0.001

The distribution of observations among the plant / ratoon crop classes is shown in Appendix 7. Most of the fields are managed according to the norm for the Midlands and only relatively few fields are allowed to continue beyond the fifth ratoon in spite of the Spencer Holley philosophy regarding the replant decision (Section 3.6.6). There are only four and three observations for ratoons 8 and 9 respectively, and these were all produced at Kranskop.

The grouping of all crop cycles beyond the sixth ratoon is recommended on the basis of sample size alone (Appendix 7). Based on the regression coefficient differences presented in Table 17, one might conclude that the yield trend may be increasing beyond ratoon 8 and yields for ratoon

9 observations are statistically no different to those from plant crops. This apparent anomaly may be explained in part by the small sample size for the observations. However, poor-yielding ratoon crops are generally replanted early, creating an upward yield bias in the older ratoon groupings. Alvarez *et al.* (1982) excluded all observations beyond the fifth ratoon from their yield analyses for sugarcane in Florida for this reason. It follows that the estimates of yield decline shown by the regression coefficient differences in Table 22 may be conservative. Nonetheless, a significant yield decline is shown with increasing ratoon age and this can be used to provide useful management information for the replant decision.

The district – yield relation

While there is statistical merit for grouping the data from Kranskop and Richmond based on the regression coefficient differences presented in Table 17, this was not done because of the extreme differences in physical location. Kranskop represents the northern-most and Richmond the southern-most region of the study area. The levels were reclassified using Umvoti as the reference category to avoid reporting coefficients with an unnecessary negative sign.

The crop cycle – yield relation

In spite of the importance of crop cycles reported in the sugarcane literature (Lonsdale and Gosnell, 1975; Landrey *et al.*, 1981; Sweet and Patel, 1985; Hellmann, 1988; Inman-Bamber, 1991a; Hellmann, 1993), the crop cycles defined in Section 3.11 did not contribute significantly to the regression when *rain total for cycle* was included in the model. Some levels were, however, significantly different from each other and grouping the six crop cycles on the basis of harvest season (summer or winter harvest) resulted in a significant contribution to the regression (Table 17). The levels were reclassified using the winter harvest as reference category to avoid reporting coefficients with an unnecessary negative sign.

On average, crops harvested in summer produce a significantly higher cane yield than crops harvested in winter, irrespective of the season in which the cycles started. This seems contrary to the findings of Inman-Bamber (1994a) who reported superior radiation use efficiencies and faster rates of canopy formation for summer ratoons compared with winter ratoons, concluding that the crop starting season was more important than the harvest season. It is only possible to speculate about the reasons for this apparent contradiction because of the limitations of the empirical investigation reported in this study. A possible explanation is that the yield difference between the two harvesting seasons (4.28 TCH) is sufficiently small that seasonal variations in sugarcane moisture content could account for most of the variation, although the appropriate data were not recorded to confirm this. Nevertheless, the significant statistical relation of superior yields from summer harvests has been established using a large number of observations and therefore warrants further investigation to establish whether or not there is a physiological relation explaining the observation.

The continuous variable – yield relations

N fertilizer, *rain total for cycle*, and *total K* were found to be significantly positively related to cane yield (Table 17). These trends had been expected because N is known to strongly influence absolute biomass yields, rain (crop water supply) is positively related to yield in other sugarcane yield models (Section 2.3), and K is known to be especially important in sugarcane nutrition (Van Dillewijn, 1952).

Altitude was significantly negatively associated with cane yield (Table 17). At a local scale higher altitudes are usually associated with slightly moister, cooler micro-climatic conditions. Since cane yield increases with moisture supply and temperature (Van Dillewijn, 1952), the altitude effect on yield at the Mondi estates probably represents a temperature (radiation) gradient. Climatic data are not available at an appropriate scale to confirm this. *Profile gravel* is also significantly associated with lower cane yields (Table 17). Thus, as the soil gravel content increases, yields decrease – probably because of a reduced rooting volume.

Transformations

The dependent variable

Analyses of residuals were used to check the explanatory model for bias and curvilinear trends. In all cases the assumptions of normality and homoscedasticity were upheld and therefore no transformation of TCH was required.

The continuous variables

A number of potentially appropriate transformations, e.g. quadratic, square root, logarithmic, and inverse, of all the continuous variables were screened during model building. None of these improved the explanatory model.

Interactions

Interactions between pairs of categorical variables and between categorical and continuous variables were investigated for the explanatory model. Significant interactions were found between *ERD* and *district*, and *rain total for cycle* and *aspect*. Both *ERD* and *aspect* did not contribute significantly to the model on their own, but when they were tested in association with the relevant interaction terms significant contributions were obtained.

When *ERD* and the *ERD* × *district* interaction were included in the explanatory model, *profile gravel* was dropped because it no longer contributed significantly to the regression. Since *profile gravel* and *ERD* have a high bivariate correlation coefficient and both describe soil volume characteristics, it is desirable to include only one of these terms in the model. These changes improved model accountability using the regression coefficients presented in Table 23.

Table 23: Regression coefficients and related statistics for the *ERD* × *district* interaction in the revised explanatory model.

Categorical variable	Levels	Regression coefficient differences	Standard deviation	t-value (523 d.f.)	Probability
District	Umvoti	-	Reference category		
	Kranskop	45.3	15.8	2.86	0.004
	Richmond	67.0	14.4	4.64	<0.001
Continuous variables		Regression coefficient	Standard deviation	t-value (523 d.f.)	Probability
ERD (cm)		0.2973	0.0891	3.34	<0.001
ERD × *Umvoti*		-	Reference category		
ERD × *Kranskop*		-0.254	0.1240	-2.04	0.042
ERD × *Richmond*		-0.406	0.1070	-3.80	<0.001

Based on the regression coefficients and regression coefficient differences (Table 23) it can be concluded that *ERD* is a very important yield explanatory variable at Umvoti. At both Kranskop and Richmond *ERD* appears to be less important for explaining yield differences and does not contribute significantly to their individual regressions. It is also for this reason that the regression coefficient differences for *district* in Table 23 have changed substantially from those presented in Table 17. The revised explanatory model predicts high yields at Kranskop and Richmond, largely independently of soil rooting depth, while at Umvoti high yields are generally predicted only on the deeper soils.

The reason for the different yield responses to *ERD* between the three districts is difficult to explain, especially since *ERD* was shown to be a significant predictor in the prediction model, independent of district (see later). No other significant interactions between either *ERD* or *district* were found to help explain the statistical trend. Umvoti is the driest of the three regions based on median rain accumulated for the duration of each crop cycle and should therefore be expected to be more sensitive to soil rooting volume than the other two districts. (Umvoti median cycle rain = 1717 mm, Kranskop median cycle rain = 2348 mm, and Richmond median cycle rain = 1945 mm). It is, however, unlikely that only differences in rainfall account for the observed trends, firstly because *rain total for cycle* is included as an independent variable in the model and no significant interaction between this variable and *district* was found, and secondly because the response to *ERD* is least important at Richmond although these crops usually receive only ca. 200 mm more rain than at Umvoti. When *TAW* was substituted for *ERD* to account for possible differences in estimated soil water holding capacities between the districts the same trends were observed as for *ERD*. Although the TAW – yield trends were slightly stronger than those for ERD, *ERD* was used as a predictor in preference to *TAW* because TAW is calculated from ERD and soil clay content using a sophisticated formula according to the FSD system (Anon., 1994).

The *rain total for cycle* and *aspect* main-effects and interaction (*rain* × *aspect*) terms together contributed significantly to the model and improved its accountability and goodness of fit statistics. The regression coefficients and coefficient differences presented in Table 24 show that, in general terms, fields with southern and eastern aspects produce superior yields to those with northern and western aspects on the Mondi estates. The warmer fields with northern and western aspects are more responsive to rain than those with southern and eastern aspects, as expected (Section 3.5.1). It follows that for crop cycles growing during high-rainfall years, greater yields will usually be produced on northern and western aspects than on southern and eastern aspects, while the converse will generally apply during low-rainfall years.

Table 24: Regression coefficients and related statistics for the *rain* × *aspect* interaction in the revised explanatory model.

Categorical variable	Levels	Regression coefficient differences	Standard deviation	t-value (521 d.f.)	Probability
Aspect	North and West	-	Reference category		
	South and East	15.54	6.44	2.41	0.016
Continuous variables		**Regression coefficient**	**Standard deviation**	**t-value (521 d.f.)**	**Probability**
Rain total for cycle (mm)		0.01422	0.00243	5.86	<0.001
Rain × *North and West*		-	Reference category		
Rain × *South and East*		-0.00623	0.00299	-2.08	0.038

Observations with poor model fit

In the examination of observations with poor model fit, observations with standardized residuals >3 in absolute value and those with leverage >0.1 were investigated. Since the explanatory and prediction models were developed in close association with each other, "problem" observations in each model were investigated simultaneously. Eight observations were identified as outliers, of which four were common to both models. Careful examination of the records confirmed that these observations were at variance with previous field history and the general production trends. They were therefore excluded from the database as abnormal observations. These observations are described in Appendix 8 and each exclusion justified individually. The final models were therefore both based on 535 observations.

Muticollinearity

Multicollinearity is a common feature of natural resource data and has been identified as an important limitation in site-factor studies for timber and other agricultural commodities (Grey, 1983; Schutz, 1990), especially those based on linear regression techniques (Donkin, 1994). It results from high correlations between the independent variables, causes the correlation matrix to approach singularity, and gives rise to imprecise estimates of the regression coefficients (Grey, 1983). Although bivariate correlations give an indication of collinearity, their limitations as a diagnostic tool have been discussed. The inclusion of interaction terms in the model is normally also expected to increase the variance inflation quite considerably. No formal tests for multicollinearity other than bivariate correlations are available in GENSTAT version 5.3 release 2, although a more recent release provides a warning for models with problematic levels of variance inflation. The final model was therefore run in GENSTAT version 5.4 release 1. Since no variance inflation warnings were issued it was concluded that multicollinearity was not a problem in the model.

The improved explanatory model

The adjustments to Model 1.0 described above resulted in an improved model (Model 1.1, Table 25), having increased the R^2 from 0.519 for Model 1.0 to 0.547 for Model 1.1; R_a^2 increased from 0.486 to 0.529. Since multicollinearity is not a problem in the model, the contribution of each variable to yield can be estimated using the partial regression coefficients. Analysis of residuals showed no marked deviation from the assumptions of normality and variance homogeneity (Figures A9.1 and A9.2, Appendix 9).

Table 25: Model 1.1 Regression coefficients for the final explanatory model.

Categorical variables	Levels	Regression coefficient differences	Standard deviation	t-value (513 d.f.)	Probability
Constant	Reference categories †	-0.7	30.9	-0.02	0.982
Season	1 (1993, 1996)	-	(Reference category)		
	2 (1990, 1994, 1995, 1997)	13.77	2.16	6.36	<0.001
	3 (1991, 1992)	21.42	2.60	8.23	<0.001
	4 (1985, 1988)	29.41	3.49	8.42	<0.001
	5 (1986, 1987, 1989)	33.44	2.94	11.37	<0.001
Soil type	1 (Humic yellow and red / yellows)	-	(Reference category)		
	2 (Humic red)	5.70	2.14	2.67	0.008
	3 (Orthic red)	13.40	3.05	4.39	<0.001
Variety	1 (NCo376 and NCo293)	-	(Reference category)		
	2 (N16 and N12)	10.41	2.46	4.24	<0.001
Plant / Ratoon crop	1 (Plant, Ratoons 1 and 2)	-	(Reference category)		
	2 (Ratoons 3, 4 and 5)	-5.90	2.00	-2.95	0.003
	3 (Ratoons 6, 7, 8 and 9)	-12.75	2.87	-4.45	<0.001
District	Umvoti	-	(Reference category)		
	Kranskop	50.7	14.9	3.39	<0.001
	Richmond	67.1	13.5	4.97	<0.001
Crop cycle	Winter harvest	-	(Reference category)		
	Summer harvest	4.90	1.70	2.89	0.004
Aspect	North and West	-	(Reference category)		
	South and East	14.65	6.03	2.43	0.015
Continuous variables		Regression coefficient	Standard deviation	t-value (513 d.f.)	Probability
N fertilizer (kg ha^{-1})		0.1985	0.0391	5.08	<0.001
Rain total for cycle (mm)		0.0139	0.0027	6.13	<0.001
Rain × North and West		-	(Reference category)		
Rain × South and East		-0.0059	0.0028	-2.12	0.035
Altitude (m a.s.l.)		-0.0492	0.0248	-1.99	0.048
Total K (kg ha^{-1})		0.0147	0.0040	3.70	<0.001
ERD (cm)		0.3410	0.0833	4.09	<0.001
ERD × Umvoti		-	(Reference category)		
ERD × Kranskop		-0.2840	0.1770	-2.42	0.016
ERD × Richmond		-0.3963	0.1000	-3.96	<0.001
$R^2 = 0.547$	SE = 16.5 TCH				

† The constant term refers to the base yield for the reference categories.

Other considerations

Observations which had been checked for accuracy with the estate managers and coded as a subset separate from the remaining data during the preparation of the final database (Section 4 and Table 15) were used to define a categorical variable describing data quality (Appendix 7).

This variable was tested as a yield predictor to establish whether or not significant differences of importance to this study existed between the subsets. Since the two levels were not significantly different from each other, and the variable did not contribute significantly to the regression, the general accuracy of all the observations was acceptable and the observations could be analysed as a single data set.

The laboratory at which the soil sample for each observation was analysed was also defined as a categorical variable, *laboratory* (Appendix 7). The purpose was to differentiate between FAS and GCI analyses to establish whether or not the Triomf conversion factors (Section 3.7.1) adequately accounted for analytical differences between the laboratories for the purposes of this study. Since no significant difference exists between the levels and no significant interactions were found between *laboratory* and soil nutrient levels or fertilizers, it was not necessary to differentiate between the origins of the analyses. These findings should not be used as confirmation of the adequacy of the conversion factors since the GCI analyses comprised less than 10% of the database (44 observations) and only their importance as yield predictors was investigated.

4.2.2.2 *The prediction model*

The prediction model was developed using the same approach described for the explanatory model. *Season* and *rain total for cycle* were excluded from model building because they can be determined only after harvesting. Since both these variables account for a large portion of the yield variation described by the explanatory model (Model 1.1), alternative variables were sought to predict yield. The regression coefficients and related statistics for the prediction model are presented in Table 26.

Only some of the variables used in the explanatory model contributed significantly to the prediction model and the inclusion of additional categorical variables did not improve the model. Those that did contribute significantly were refined by grouping statistically similar levels and the resulting groups were identical to those determined for the explanatory model (Table 25). *Crop cycle* was retained in the model because estate managers usually plan their field harvesting schedule well in advance of the harvest and this information can be used to determine whether a field will be harvested during summer or winter (Section 3.11) when yield predictions are made.

Table 26: Model 2 Regression coefficients for the prediction model.

Categorical variables	Levels	Regression coefficient differences	Standard deviation	t-value (520 d.f.)	Probability
Constant	Reference categories †	46.7	31.2	1.49	0.136
Soil type	1 (Humic yellow and red / yellows)	-	(Reference category)		
	2 (Humic red)	4.43	2.36	1.88	0.061
	3 (Orthic red)	13.39	3.38	3.96	<0.001
Plant / Ratoon crop	1 (Plant, Ratoons 1 and 2)	-	(Reference category)		
	2 (Ratoons 3, 4 and 5)	-10.12	1.81	-5.59	<0.001
	3 (Ratoons 6, 7, 8 and 9)	-20.71	2.56	-8.10	<0.001
District	Umvoti	-	(Reference category)		
	Kranskop	11.04	3.07	3.60	<0.001
	Richmond	12.50	4.22	2.96	0.003
Crop cycle	Winter harvest	-	(Reference category)		
	Summer harvest	7.17	1.83	3.91	<0.001
Continuous variables		Regression coefficient	Standard deviation	t-value (520 d.f.)	Probability
N fertilizer (kg ha^{-1})		0.2190	0.0414	5.29	<0.001
Rain for 12 months' growth (mm)		0.0756	0.0138	5.49	<0.001
(*Rain for 12 months' growth*)^2 (mm^2)		-0.00002114	0.00000598	-3.54	<0.001
Altitude (m a.s.l.)		-0.0595	0.0269	-2.22	0.027
Total K (kg ha^{-1})		0.0184	0.0044	4.18	<0.001
Topsoil Ca:Mg		-0.7840	0.2170	-3.62	<0.001
ERD (cm)		0.0903	0.0437	2.07	0.039
$R^2 = 0.430$	SE = 18.4 TCH				

† The constant term refers to the base yield for the reference categories.

The regression coefficients for the continuous variables in the prediction model (Table 26) show that the effect of *N fertilizer*, *altitude*, *total K* and *ERD* on yield is similar to that established in the explanatory model, although *ERD* makes a significant positive contribution to yield irrespective of district in Model 2. *Rain for 12 months' growth* replaces *rain total for cycle* because this variable is known in advance of harvesting and makes a highly significant contribution to the regression.

Topsoil Ca:Mg (ratio calculated using the kg ha^{-1} analytical values; Appendix 7) makes a highly significant contribution to the prediction model (Table 26). It is disturbing that, in spite of its apparent importance in the prediction model, this variable does not feature in the explanatory model. Clearly variables in the explanatory model that were excluded from the prediction model describe most of the yield variation accounted for by the nutrient ratio. When the data were examined, an association was found between *season* and an increasing topsoil Ca:Mg ratio. This is probably the result of continued heavy applications of lime and gypsum recommended by Spencer Holley. A widening of the soil Ca:Mg ratio should be expected even when dolomitic lime is used because the South African sources of lime are not Mg-rich. In addition, gypsum applications increase the mobility of Mg in the soil profile, increasing the losses of the cation from the topsoil, thereby widening the Ca:Mg ratio. Efforts to maintain topsoil Mg levels by applying dolomitic lime after gypsum applications (Section 3.8.2) appear to have been inadequate in general.

Some information about the source variables is always lost when using ratios. The negative yield response to a widening topsoil Ca:Mg ratio has not been documented in SASEX research although negative yield responses to heavy lime applications, and no yield response to gypsum, have been reported (Meyer *et al.*, 1991; Schroeder *et al.*, 1993; 1995). It is possible that at least a portion of the negative regression coefficient (Table 26) can be attributed to the peculiarities of the data set. This is because *season* (excluded from Model 2) showed that, in general, yields during the 1990's were lower than during the 1980's, while the Ca:Mg ratio during the 1990's tended to be wider than during the 1980's. However, *topsoil Ca* on its own contributed negatively to the regression, although not significantly, while *topsoil Mg* was associated with a significant positive yield response. Since the yield response to *topsoil Mg* was statistically weaker that for *topsoil Ca:Mg*, the ratio was used in the model. It follows that the Mondi data show some evidence of a yield response to increasing Mg nutrition, especially in relation to topsoil Ca. While it is likely that the importance of the relation is overstated because of the yield-season characteristics of the available data, it may provide important crop management information and warrants further research.

Appropriate transformations of the continuous variables were examined for the prediction model and the inclusion of a quadratic transformation for *rain after 12 months' growth* made a highly significant contribution to the regression (Table 26). Thus for Model 2, yield increases with increasing rainfall during the first 12 months' growth, but at a decreasing rate with increasing rainfall. This indicates that water is less yield-limiting than other important growth factors during the first years' growth in high rainfall years on the Mondi estates. No other transformed variables contributed significantly to the regression and the analysis of residuals (Figures A9.3 and A9.4, Appendix 9) confirmed that no transformation of TCH was required for Model 2.

No significant interactions were found for the prediction model.

Model 2 was checked for multicollieարity using GENSTAT version 5.4 release 1. Considerable variance inflation is expected when using both the linear and quadratic terms in a model. Multicollinearity is, however, within acceptable limits since no inflation warning was issued.

Other potential variables

During model building, *Zn fertilizer* (Appendix 7) showed potential as an additional explanatory variable in Model 2 while *topsoil PDI*, and *TAW* (Appendix 7) could be used as a alternative predictors to *ERD*.

Zinc fertilizer

Zinc fertilizer appeared to make a significantly negative yield contribution in Model 2. The negative regression coefficient was investigated because management could easily reduce inputs if yield was reduced by Zn. It seemed unusual that *Zn fertilizer* was not included in the explanatory model (Model 1.1) but appeared to be very important in the prediction model. When the data were examined it was obvious that zincated fertilizers had only been used in recent years, although there was some evidence that zincated fertilizers had been used previously without recording the inputs (Section 3.8.1). Since Zn applications during the 1980's were generally negligible and considerable inputs had been made during the 1990's, it appeared that the association between generally high yields during the 1980's and generally low yields during the 1990's had been inadvertently "accommodated" in *Zn fertilizer*. This was tested by adding *season* to the regression and confirmed that the Zn fertilizer – yield relation was based on peculiarities of the data. *Zn fertilizer* was therefore not used as a predictor for Model 2.

Topsoil phosphorus desorption index (PDI)

When *PDI* was tested as a predictor the model was based on 478 instead of 535 observations because a number of PDI analyses were missing in the database (Appendix 7). A significant negative relation between *PDI* and yield was found and the inclusion of *PDI* required *ERD* to be dropped from the model because of a non-significant t-value. Revised models, including *PDI* as a predictor, are presented in Appendix 10. Differences in *PDI* between Umvoti and Richmond appeared to account adequately for the yield differences between the two districts because their regression coefficient differences were non-significant when *PDI* was included in the model (cf. Table 26 for Model 2). Using *PDI* in preference to *ERD* resulted in a model with similar goodness of fit statistics to Model 2 although it was based on fewer observations. For this reason either *ERD* or *PDI* could be included in the prediction model. The primary consideration for not using *PDI* as a predictor in Model 2 is that it is not determined by the GCI laboratory which now conducts Mondi's routine soil analyses, i.e. it is pointless including a variable in the prediction model which is not available to model users.

In the explanatory model *PDI* was also negatively associated with yield (Appendix 10). The term was not included in the model because the t-value was only marginally significant (approximately 10%) and the model would have been based on fewer observations. These models are therefore not discussed any further. However, the practise of basing P fertilizer recommendations on PDI values from previous FAS analyses when the GCI laboratory is used (Section 3.7.1.1) is questionable because PDI varies considerably between analyses for individual fields, and this variation was shown to be associated with observed yields. This is an issue worth pursuing especially because P is the most expensive fertilizer element per unit commonly applied in sugarcane agriculture.

Total available water (TAW)

A possible substitution of *TAW* for *ERD* had been investigated for the explanatory model. During model building a similar situation arose for the prediction model. For Model 2, *ERD* was statistically stronger than *TAW* and since the former variable is also more easily determined, it was selected as the more appropriate predictor for the model.

Location-specific models

It is unlikely that the explanatory model (Model 1.1) could be refined further to fit the data better, although the R^2 indicates that some 45% of the total variation in observed yields remains unaccounted for. This could be due to the omission of some unrecorded parameters, e.g. temperatures for each crop cycle, or to unquantifiable sources of variation such as animal damage, lodging, etc. The restrictions placed on the variables screened for the prediction model (Section 4.2.2.2) resulted in a model with even poorer accountability (Model 2, $R^2 = 43\%$; 67% of the total variation in observed yields is unaccounted for). Consequently models were developed for each district individually in an attempt to improve model accountability. The most promising candidate models had goodness of fit statistics very similar to those for Model 1.1 and no additional predictor variables were statistically significant. It was concluded that nothing was to be gained by using location-specific models in preference to Models 1.1 and 2 and this modelling approach was abandoned.

4.3 MODEL VALIDATION

The R^2 and SE statistics refer to model fit alone and do not indicate how the model will perform when applied to new data sets, even from the same resource base. Models are truly valid only for the data from which they are calculated or from which the data are a random sample (Bell, 1981). For this reason validation is essential before a model can be used to predict yield with confidence. Although it is preferable to use a data set excluded from model construction (Bell, 1981) sugarcane yield models for which validation procedures have been reported all used cross validation techniques. It is also important to remember that before empirical models can be used on different data sets or are applied in a different area to the one in which they were developed, further validation is required to ensure that the same statistical relations between variables still apply (Schutz, 1990).

Harvest data for the 1998 season were collected from the Mondi estates to validate the explanatory model (Model 1.1) and the prediction model (Model 2). These data were not included in the dataset used for model building (Section 4.2.2) although they were obtained from the same fields. Of the fields that were harvested during 1998 a total of 47 fields, 21 from Kranskop, 9 from Umvoti and 17 from Richmond, met the conditions specified for the models (Section 4.1.1). The sugarcane yields (TCH) for these observations are presented in Appendix 11 together with the independent variables required to run the two models. The yields of the validation data are generally high and range from 69 TCH to 151 TCH. This was expected since the size of the 1998 sugarcane crop was a new national record. It follows that the validation data do not represent the full yield range of the data used to derive the models (Appendix 7) and the lowest yields, in particular, are poorly represented.

4.3.1 Validating the explanatory model

Model 1.1 was run in GENSTAT to predict sugarcane yield (TCH) using the fixed conditions for each observation from the 1998 harvest season (Appendix11). A difficulty arose when classifying the 1998 harvest season in terms of the 5-level season variable because no data from this season had been used to develop the model. Various preliminary groupings showed that the 1998 harvest was most similar to season level 4 and therefore this level was specified in the predictions, i.e. the 1998 harvest was predicted according to the 1985 and 1988 crop performance (Table 25). The observed yields were compared to predicted yields (Figure 6).

Model 1.1 fits the data well since 75% of the predicted sugarcane yields lie within 20% of the observed yields; 47% of the predicted sugarcane yields lie within 10% of the observed yields. Satisfactory predictive performance was also achieved at the individual district level (Figure 6). The distribution of observed versus predicted TCH yields (Figure A11.1, Appendix 11) reveals no systematic pattern and the majority of observations lie well within the 95% confidence interval for individual predictions. This confirms that the predictor variables identified for the explanatory model were also important yield determining factors during the 1998 season.

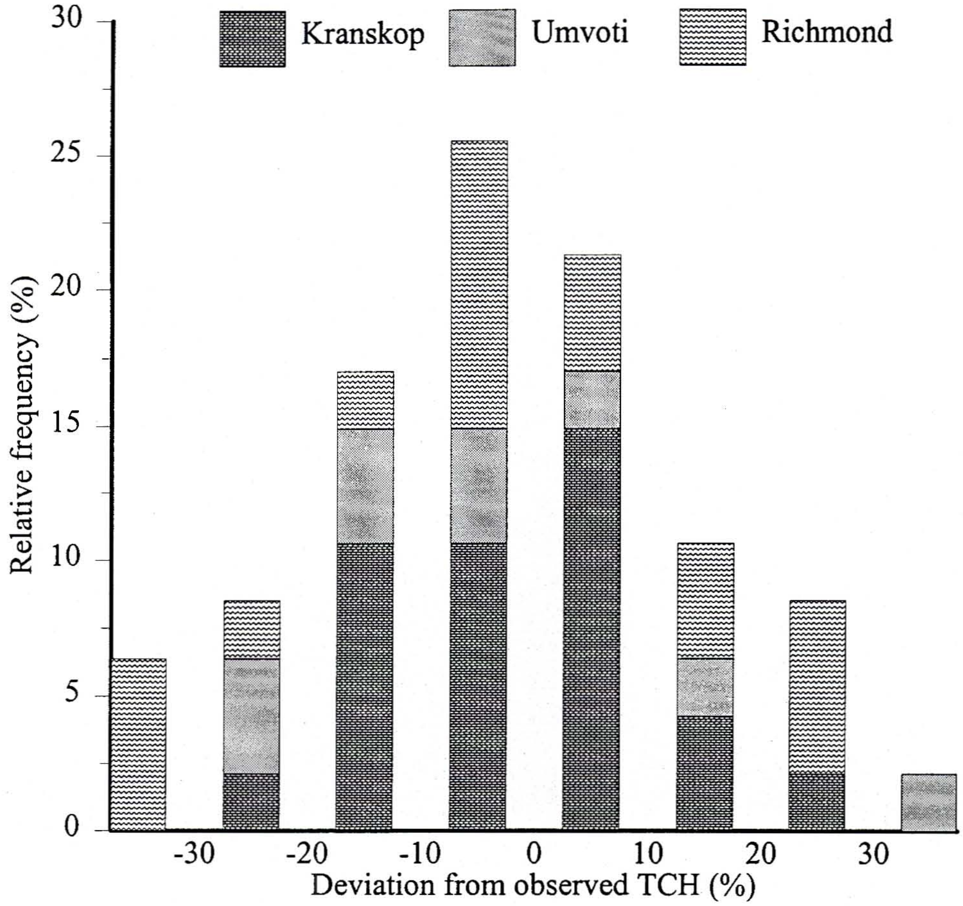


Figure 6: Graph showing the relative frequency of deviations of predicted sugarcane yields from observed yields (TCH) for the 1998 harvest season using the explanatory model (Model 1.1).

Although the model appears to predict observed sugarcane yields (TCH) particularly well, the broad confidence interval for individual predictions (mean confidence limit ±33.6 TCH) reduces the meaning and usefulness of these yield predictions for management. Yield predictions for the population mean are, however, considerably more reliable (95% mean confidence limit ±8.4 TCH) and can be used to categorize mean yield responses to the independent variables.

4.3.2 Validating the prediction model

Model 2 was validated using the conditions for the variables needed to run the model for each observation from the 1998 harvest season (Appendix11). The prediction model was developed specifically to predict future sugarcane yields and therefore all predictors are known at least six months in advance of the scheduled harvest date (Section 4.2.2.2). Since the prediction model

R² is lower and the SE is larger than for the explanatory model, poorer yield predictions are expected from this model. The validation results are shown in Figure 7.

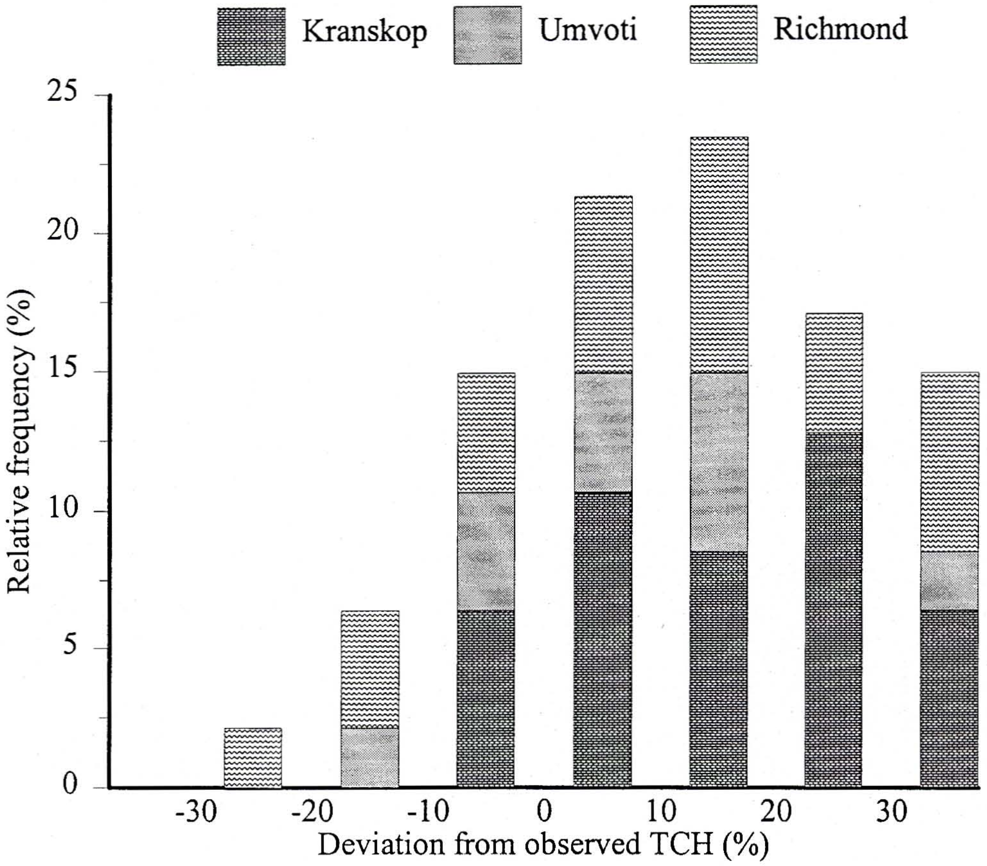


Figure 7: Graph showing the relative frequency of deviations of predicted sugarcane yields from observed yields (TCH) for the 1998 harvest season using the prediction model (Model 2).

Although the model under-predicts the sugarcane yields that were achieved during the 1998 season (Figure 7), 66% of predicted sugarcane yields lie within 20% of the observed yields; 36% of predicted sugarcane yields lie within 10% of the observed yields. The prediction model predicts sugarcane yields for the "average" season, based on the 13 years' data used to derive the model because *season* was deliberately excluded as a predictor. Since the 1998 harvest was particularly large across the entire industry, one might anticipate that this model would underestimate actual yields. This is shown in the graph of observed versus predicted TCH yields (Figure A11.2a, Appendix 11). Similarly, one should expect this model to over-predict yields in poor seasons.

It is possible to modify the predictions as more information becomes available during the harvest season to improve the accuracy of the absolute yield predictions using Model 2. To demonstrate this the actual yields for the first month's harvest for the 1998 season (six observations for March / April) were compared to the predicted yields. The mean difference between the observed and predicted yields was 21 TCH. The intercepts of the predicted yields and confidence intervals of individual predictions were therefore adjusted upwards by 21 TCH. This calibration resulted in yield predictions that are better distributed about the zero-mean deviation from observed sugarcane yields (Figure 8) than for the un-calibrated predictions (compare Figure 8 with Figure 7). Most of the observed yields lie well within the 95% confidence limits for these calibrated individual predictions (Figure A11.2b, Appendix 11) and the effect of the intercept adjustments can be seen by comparing Figure A11.2a with Figure A11.2b (Appendix 11). In spite of the apparent improvement in the accuracy of absolute yield predictions achieved by increasing predicted yields by 21 TCH (Figure 8 and Figure A11.2b, Appendix 11), no real improvement was achieved in the distribution frequency of the deviations of predicted yields from observed yields: 66% of predicted sugarcane yields lie within 20% of the observed yields; 36% of predicted sugarcane yields lie within 10% of the observed yields (Figure 8).

It is fortuitous that the available validation data result in very similar distributions of the absolute deviation classes from observed yields for predictions using Model 2, and the calibrated predictions. Based on the available evidence, however, there is little to be gained from calibrating yield predictions from Model 2 using current yield data.

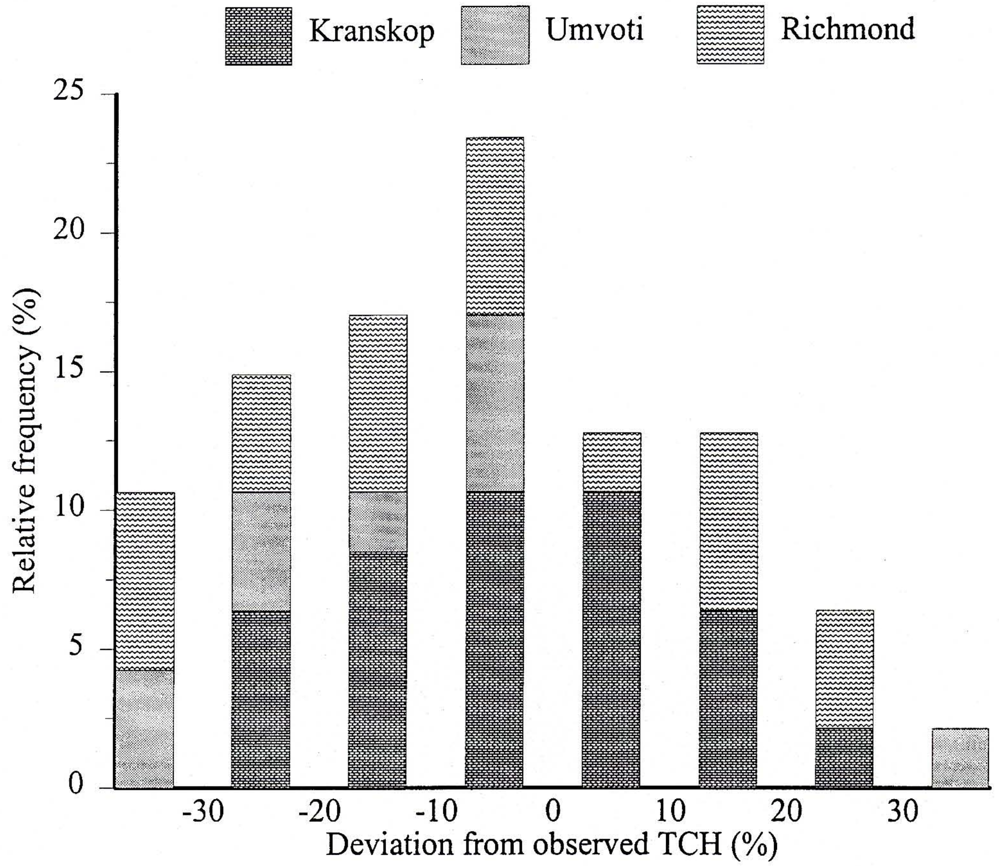


Figure 8: Graph showing the relative frequency of deviations of predicted sugarcane yields from observed yields (TCH) for the 1998 harvest season using the prediction model (Model 2) and additional information available after the first month's harvest.

Ideally Model 2 should be capable of predicting individual field yields reliably. Since field yields for 66% of the observations in the validation study were predicted to within 80% of observed yields model users may be tempted to apply the model at this scale. However, the 95% confidence interval for individual field predictions is extremely large (mean confidence limit ±36.8 TCH). It follows that the predicted yield for a field is expected to vary by almost as much as the total yield variation between fields for the 1998 season (73.6 TCH versus 82 TCH). Thus the model should preferably not be applied at this scale. Since the mean 95% confidence limit standard deviation for the population mean is ±6.1 TCH managers should expect to use the model with considerable success for yield predictions at a coarse resolution, i.e. broad-based mean yield responses to the various predictor categories.

The validation studies for both the explanatory model and the prediction model have shown conclusively that, based on data for the 1998 season, the models are valid for the sugarcane enterprise on the Mondi Midlands estates. The models can therefore be used with confidence by the estate managers to guide decisions that require an estimate of expected future yields. Model users must be aware of the implication of the large confidence interval for individual predictions which result from the low R^2 and high SE of both models. Predictions for the population mean are considerably more reliable than for individual fields primarily because of the large number of observations on which the models are based. These predictions may be particularly useful to management because the 95% confidence interval of both models is less than 10% of the median yield of 88 TCH over 13 years, across the three districts (Appendix 7).

The explanatory model should be used in preference to the prediction model when model users are not constrained by the availability of input data because the model has a larger R^2 (55% versus 43%), a lower SE (16.5 TCH versus 18.4 TCH) and the validation study showed predictions from Model 1.1 to be more reliable than those from Model 2 (compare Figure 6 with Figures 7 and 8).

Although the prediction model (Model 2) was built for the specific purpose of predicting yields at least six months in advance of the scheduled harvest, the validation study clearly demonstrated that, while the model predicted relative yields well, the model did not predict actual yields satisfactorily (Figure A11.2a, Appendix 11). Less biassed predictions can be obtained from Model 2 after commencement of the harvest season when predicted yields can be calibrated using a sample of actual yields for the season (Figure A11.2b, Appendix 11). Such calibrations can be repeated periodically during the harvest season as more data become available to improve the reliability and accuracy of yield predictions for the remainder of the season.

4.4 MAPPING EXPECTED YIELD

The yield models developed in Section 4.2 may be used to predict expected yields for specified levels of the predictor variables on the Mondi estates. Coupled with GIS, these yield predictions give managers an overview of the spatial distribution of mean expected yields for the specified conditions, and may be useful for planning, for example, the evaluation of current production levels, and scenarios involving the allocation resources when budget allocations are made.

In this section yield predictions were made according to actual soil type, district, aspect, altitude and ERD combinations using the explanatory model (Model 1.1). To avoid the danger of extrapolative predictions, yield mapping was restricted to the existing sugarcane areas on the estates that provided data for model building (Appendix 12). The model was linked to ARCINFO on an *ad hoc* basis (Section 2.5) and the GIS used to provide spatial inputs of soil type, district, aspect, altitude, and ERD. The non-spatial predictors were specified at levels for the "average" situation, based on crop management practises recommended by SASEX and the following assumptions:

a no season level was specified and the mean season effect over the 13 years' data was estimated using marginal weights, held constant over the levels of the other categorical variables;

b the variety was one of the new-generation varieties, either N12 or N16;

c the crop belonged to group 1, i.e. plant crop, first or second ratoon;

d the crop was harvested during the summer months, i.e. March, April, October, November or December;

e N fertilizer applications were made according to SASEX recommendations (Anon., 1996), viz. 120 kg ha^{-1} on all soils, except the humic soils at Richmond which received 100 kg ha^{-1};

f median rainfall, accumulated for the duration of each crop cycle, was used on a district-specific basis, i.e. 2348 mm at Kranskop, 1717 mm at Umvoti and 1945 mm at Richmond; and

g total K was 440 kg ha^{-1} across the entire sugarcane area, based on SASEX soil fertility recommendations for single crop cycles (Anon., 1996).

Thematic maps of predicted TCH yields, generated by the GIS using yield groupings based on model outputs, are presented in Appendix 12. The predicted yields for Model 1.1 represent the expected "average" yields for the 13 seasons to which the records apply. The partial regression coefficients (Table 25) show that considerable yield differences occur between seasons and this must be borne in mind when evaluating the predicted yields (Appendix 12). Although there is no significant statistical evidence of an interaction between the rain and season variables, and the poorest season group does not coincide only with the crop cycles receiving the least rain (Table 27), model users should be cautious when defining the season level and rain total because it is unlikely that these variables are truly independent. For this reason median rainfall, accumulated for the duration of a crop cycle for each district, was used to reduce the danger of making unrealistic predictions.

Table 27: Harvest seasons ranked according to mean rain total for cycle per season. The grouped season levels for Model 1.1 are also shown; level 1 represents the poorest season and level 5 the best season.

Harvest season	Mean rain total for cycle (mm)	Season level
1993	1238	1
1992	1537	3
1994	1613	2
1986	1761	5
1995	1908	2
1990	2117	2
1991	2119	3
1987	2239	5
1996	2312	1
1997	2408	2
1989	2410	5
1985	2523	4
1988	2779	4

The sugarcane variety, plant / ratoon status and crop cycle are entirely dependent on crop management decisions. The highest-yielding conditions identified during model building (Section 4.2.2) were specified for these variables across the entire sugarcane production area to ensure that fair comparisons of yield potential can be made using the predicted yields (Appendix 12). The realities of sugarcane agriculture prohibit the achievement of these assumptions in any season on an entire estate because it is economical to produce sugarcane crops older than the second ratoon and harvesting operations continue throughout winter when the sucrose concentration is high. Yield predictions should clearly not be expected to apply where crop conditions are different to the specified assumptions.

Lower N applications are made on the humic soils at Richmond than on those at both Kranskop and Umvoti because they have a higher organic matter content and are generally classified as mineralizing category 4 soils while those at Kranskop and Umvoti are generally category 3 soils. Since there is no evidence of significant interactions between *soil type*, *district* and *N fertilizer*, these different fertilizer application rates are justified. The assumed N application rates are also very similar to the mean N applications over the 13 years' records for the districts (121.6 kg ha^{-1} at Kranskop, 126.4 kg ha^{-1} at Umvoti and 100.4 kg ha^{-1} at Richmond). The yield predictions therefore lie well within the predictive range of the model for N fertilizer inputs.

The total K of 440 kg ha^{-1} is realistic for yield predictions because the mean total K for observations used to build the model is 425 kg ha^{-1} (Appendix 7). The SASEX K nutrient threshold varies according to soil clay content and since most of the soils planted to sugarcane on the Mondi estates are clayey, the recommendations for soils with more than 30% clay were used. These recommendations suggest a target total K of 440 kg ha^{-1} because the application of 100 kg ha^{-1} K fertilizer is required when the soil K level is 340 kg ha^{-1} (Anon., 1996).

Actual confidence limits for predicted yields based on the above assumptions could not be calculated because no season level was specified. In general, confidence limits for the model predictions should, however, be of similar magnitude to those reported in Section 4.3. The field layout has been shown on the yield maps (Appendix 12) to allow an interpretation of the relief of the sugarcane field. Field numbers have not been included because the yield predictions should not be applied to individual fields since the large model SE does not allow reliable estimates for individual predictions. The yield maps (Appendix 12) should therefore be interpreted as the mean population response for the specific resource combinations (soil type, district, aspect, altitude and ERD) given the assumed levels of the other predictors; actual yields from individual fields will vary considerably about the mean yields.

At the district-level the model predicts differences in expected yields in response to soil type, aspect, altitude and ERD differences only because the levels of the other predictors are held constant. At Umvoti the sugarcane enterprise has been established on only one soil type and therefore this variable makes a constant contribution to predicted yields. *Soil type* does, however, account for yield variations of between 4 TCH and 14 TCH at Kranskop and Richmond. Different

N fertilizer rates for humic and orthic soils at Richmond widen the expected yield differences from these soils (Table 25).

The superior yields expected on southern and eastern aspects compared with northern and western aspects (Table 25) are modified by *rain* × *aspect* to the extent that when a crop cycle receives 2483 mm rain, TCH yields on southern and eastern and northern and western aspects will, on average, be equal across constant levels of the other variables. The median rainfall used for the yield predictions (2348 mm at Kranskop, 1717 mm at Umvoti and 1945 mm at Richmond) greatly reduces the expected yield differences between southern and eastern and northern and western aspects, the differences being largest at Umvoti and smallest at Kranskop. This illustrates how differences in the model assumptions can drastically alter the expected crop responses and this must be borne in mind when interpreting the predicted yield maps (Appendix 12). Areas identified as high-yielding classes on these maps will not necessarily produce the best yields (on a relative scale) during droughts or exceptionally wet seasons. The model can be re-run to predict expected yields for these scenarios but users must guard against making extrapolative predictions.

Expected sugarcane yields decrease, on average, by almost 5 TCH for every 100 m rise in altitude, based on partial regression coefficients for the explanatory model (Table 25). The altitude range for the Mondi sugarcane enterprise is 205 m although this is considerably less at the district level – 147 m at Kranskop, 67 m at Umvoti and 123 m at Richmond. The general distribution of altitude across the study area implies that yields at Richmond should be slightly greater than those at Umvoti on the basis of altitude alone. These yield differences are in addition to those defined for *district* in the explanatory model (Table 25).

The large yield differences between districts suggested by the partial regression coefficients (Table 25) are reduced by the different district-level yield responses to ERD. Yield is most responsive to increases in soil depth at Umvoti where the model predicts a yield increase of almost 3.5 TCH, on average, for every 10 cm increase in ERD. At both Kranskop and Richmond the contribution of ERD to yield is statistically non-significant and similar yields are achieved largely independently of ERD.

To summarize, for selected crop conditions (variety, crop group and crop cycle) and agronomic inputs (N fertilizer and total K), soil type, district, aspect, altitude and ERD are the key variables accounting for expected yield differences in any particular season in terms of the explanatory model. The natural resources responsible for most of the variation in expected yields vary depending on the conditions that are measured or assumed, e.g. the importance of yield differences between aspect classes in relation to rainfall, and the variability of a resource at a particular location.

The maps (Figures) referred to in this Section may all be found in Appendix 12. In the example discussed above, differences in predicted yields at Kranskop are determined primarily by soil type and altitude. The predicted yield classes on Salem (Figure A12.1) correspond almost exactly with the distribution of the two soil types on the estate (Figure A12.11) and yield variations contributed by altitude (Figure A12.13) are masked within the 10 TCH yield classes. At Sutherlands the predicted yield classes (Figure A12.3) correspond closely with soil type (Figure A12.15) on the western section but seem to reflect mainly aspect classes (Figure A12.16) on the eastern section of the sugarcane area. This highlights one of the difficulties experienced when presenting spatial data in areas of fairly homogenous production potential (Johnson, 1994). All predicted yields are rounded to the nearest 10 TCH and allocated to a yield class for presentation purposes. Thus a yield of 104 TCH would be allocated to the 100 TCH yield class and yield of 105 TCH would be allocated to the 110 TCH yield class. Most of the predicted yield variation in the eastern section of sugarcane on Sutherlands is actually explained by soil type (Figure A12.15) and altitude (Figure A12.17), although the predicted values are such that the contribution of aspect is sufficient to cause areas on a southern or eastern aspect to be classified in higher yield class than areas on a northern or western aspect. (Southern and eastern aspects are only 0.7 TCH superior to northern and western aspects, on average, because of the high rainfall that was assumed for yield predictions at Kranskop). It follows that although sugarcane yields appear to be explained primarily by aspect differences in Figure A12.3 this is incorrect. Managers should therefore not rely on a pictorial representation of GIS analyses only since certain groupings of the data may sometimes be misleading. While not detracting from the value of these maps as a summary of model output, it is essential to evaluate all available data critically when making management decisions.

At Umvoti predicted yield differences are determined primarily by ERD and aspect. Yield classes for Canema (Figure A12.5) correspond closely with ERD (Figure A12.22), modified slightly by aspect (Figure A12.24) and therefore provide a convenient summary of the yield dynamics identified on the estate. As already mentioned, the numeric values of the predicted yields do not always allow such meaningful summaries to be displayed.

Soil type is primarily responsible for the expected yield differences at Richmond. Predicted yield classes for Uplands (Figure A12.7) and Greenhill (Figure A12.9) reflect the general distribution of the three soil types on the estates (Figures A12.23 and A12.27 respectively). Differences in N fertilization increase the predicted yields on orthic soils by an additional 4 TCH, on average, compared with the humic soils. Consequently the highest predicted yields are for soil type 3 (orthic red soils). Altitude (Figures A12.25 and A12.29) is also important for explaining predicted yield differences at Richmond although its influence cannot be seen in Figures A12.7 and A12.9 because it is absorbed in the 10 TCH yield classes.

Model 2 was also used to predict yields for the "average" situation on the Mondi estates linked to the GIS as for Model 1.1. The same assumptions for crop management (variety, plant / ratoon status and crop cycle) and nutrition (N fertilizer and total K) were used to ensure that yield predictions comparable to those of Model 1.1 were made. Additional assumptions were;

a median rainfall, accumulated for 12 months growth, was used instead of median rainfall accumulated for the duration of each crop cycle, i.e. 1124 mm at Kranskop, 844 mm at Umvoti and 947 mm at Richmond; and

b the topsoil Ca:Mg ratio was 3 (based on nutrient concentrations in kg ha^{-1}) across the entire sugarcane area, derived from SASEX soil fertility recommendations for optimal soil test levels of Ca and Mg (Anon., 1996).

The topsoil Ca:Mg ratio ranges between 1.4 and 34.9 for the observations used to derive the model, but the ratio is strongly skewed and the median ratio is 4 (Appendix 7). The target Ca:Mg ratio assumed for the yield predictions using Model 2 is therefore realistic even though large differences presently exist between individual fields. Managers should endeavour to reduce the Ca:Mg ratio where fields have a wide ratio provided further research confirms that there is some physiological merit in the statistical relation between the ratio and yield.

Thematic maps of the expected yields predicted by Model 2 are presented in Appendix 12. Yield predictions using the prediction model (Model 2) are simpler to interpret than those described for the explanatory model (Model 1.1) because only four spatial predictors are used (viz. soil type, district, altitude and ERD, and yield relations are not complicated by interaction terms (Table 26). Although the prediction model is statistically inferior to the explanatory model, the model is valuable when limited input data are available to decision makers. The contribution of ERD is significant in all three districts (Table 26), unlike for the explanatory model.

All maps (Figures) referred to in this Section are included in Appendix 12. Predicted yield classes for Salem and Sutherlands at Kranskop (Figures A12.2 and A12.4) reflect the general distribution of soil types (Figures A12.11 and A12.15) and ERD (Figures A12.14 and A12.18). Predicted yields are remarkably similar to those obtained using Model 1.1 (Figures A12.1 and A12.3), especially since Model 2 excludes *aspect*. The exclusion of this predictor also explains the different yield classes predicted by the two models for the eastern section of Sutherlands. Although altitude (Figures A12.13 and A12.17) individually accounts for the largest proportion of yield variation predicted for Kranskop using Model 2, this is not apparent from the yield maps (Figures A12.2 and A12.4) because the higher altitude areas tend to be associated with the weaker, shallower soils and otherwise the contribution of altitude is absorbed in the 10 TCH yield classes.

Effective rooting depth differences individually account for most of the yield variation predicted at Umvoti. It is fortuitous, but appropriate, that the predicted yields and yield classes (Figure A12.6) are such that their association with ERD (Figure A12.22) is emphasised.

At Richmond soil type differences are the most important yield determining factor for predictions based on the assumptions specified for Model 2. Predicted yields for Uplands and Greenhill (Figures A12.8 and A12.10) reflect the general distribution of soil types (Figures A12.23 and A12.27). Effective rooting depth is of greater importance as a yield predictor at Richmond for Model 2 than Model 1.1, and in some areas the predicted yield classes are modified by ERD

(Figures A12.26 and A12.30).

It is encouraging that both models predict very similar yields for the assumed "average" cropping conditions (compare Figure A12.1 with Figure A12.2; Figure A12.3 with Figure A12.4; Figure A12.5 with Figure A12.6; Figure A12.7 with Figure A12.8; Figure A12.9 with Figure A12.10). Although individual areas are sometimes allocated to a different yield class depending on whether Model 1.1 or Model 2 was used, both models generally identify the same areas as being of high, medium or low potential for sugarcane production. Such a classification on its own may be useful to management for monitoring performance, allocating resources and selecting areas for sugarcane production.

Confidence limits for the yields predicted by Model 2 were calculated for selected observations. The largest 95% confidence limit calculated (±6.6 TCH) lies outside the predicted yield class limits (±5 TCH) on which the yield maps (Appendix 12) are based. It is, however, difficult to associate confidence limits with grouped yield classes because the confidence limits will always extend beyond the displayed yield class boundaries. For this reason it is necessary to consider actual predicted yields when determining whether or not predictions are significantly different from each other and therefore whether or not different management approaches, both in terms of agronomic input levels and target yields, are justified on different sections of the estate.

Although the mean predicted yields are shown to be fairly uniform across the estates (Figures A12.1 – A12.10), this information could be useful for reviewing Mondi's management philosophy where significant yield differences are expected. In general, sugarcane fields are currently managed according to a "target" yield of 100 TCH. Since the median yield for the estates is only 88 TCH over 13 years, this target is often not achieved. Numerous fields only produce the target yield in favourable seasons because they have an inherently low production potential while other fields should be expected to produce considerably higher yields in most years. The present management approach focuses on low-yielding fields, largely ignoring the high-yielding fields, and is ignorant of the fact that the low-yielding fields may be producing well according to their potential while production from some of the high-yielding fields may be easily improved. Model 2 could be used to identify areas of high, medium and low production potential based on predicted yields for current crop cycles. Model users should, however, be cautious not to use the predicted yields blindly when making decisions regarding resource allocations which favour parts of the sugarcane enterprise which are expected to produce larger yields because the model accounts for less than half of the observed variation in yield and the explanatory model shows conclusively that interactions between resources and climate exist even though there was no statistical evidence of these relations for Model 2. It is also critical to remember that the predictions (and the confidence limits) refer to mean yield responses of the population and that the predicted performance of individual fields is expected to vary to the extent that a meaningful grouping of predicted yields is not possible.

5 DISCUSSION

The models developed in Section 4.2 serve the dual purpose of identifying key parameters related to actual commercial sugarcane yields and for yield prediction, i.e. meeting both objectives of the study. When interpreting the models it is necessary to examine the role of the selected variables as well as possible reasons for the exclusion of others. Two important considerations should be borne in mind. Firstly, the predictor variables are not necessarily the cause of sugarcane yield variations, but are related only in a strictly statistical sense. Secondly, it is difficult to determine the importance of the variables independently of one another in multiple regression because they are seldom truly independent.

5.1 THE ROLE OF THE VARIABLES

Variables tested during model building represent natural resource and agronomic parameters of the sugarcane production system (Chapter 3 and Appendix 7). The natural resource data comprise variables defining physical field attributes which remain constant for crop cycles from the same field, e.g. soil type, and climate variables which are often highly variable, e.g. rain total accumulated for each crop cycle. The agronomic data are also variable for crop cycles, even from the same field, and relate broadly to crop nutrient supply and crop management decisions. In terms of accounting for observed differences in sugarcane yields, physical field attributes are useful only for between-field comparisons. Yield differences of up to 100 TCH for successive crops from the same field could therefore be explained using only the climate, crop nutrient and crop management variables in this study.

5.1.1 Physical field attributes

Key physical field attributes associated with sugarcane yield were locality, aspect, altitude, soil type and ERD. Locality was investigated at the district level because field yields were not significantly different from each other at the estate level within districts. Field yields differed significantly between Umvoti and Richmond / Kranskop for both the explanatory and prediction models. As mentioned in Section 4.2, the reasons for these yield differences are not easily explained and any statements based on the yield models are purely speculative. Nevertheless, in terms of general climate, growing conditions for sugarcane within the study area are more favourable at Richmond and Kranskop than at Umvoti (Appendix 1 and 2). Since the high-altitude fields of the study area are concentrated at Umvoti, lower temperatures should generally be expected than at both Richmond and Kranskop (Section 3.5.3) in addition to weaker maritime influences as determined by the distance inland from the Indian Ocean coastline (Section 3.1). Both of these parameters probably contribute to the lower yields at Umvoti compared with Richmond and Kranskop. The role of management differences between the districts, possibly related to yields, is also difficult to quantify. The magnitude of the partial regression coefficient differences for *district in* the explanatory model (Table 25) should be considered in association with *ERD* × *district* which reduces, on average, the expected yield differences between Umvoti and the other districts. Based on coefficients for the prediction model (Table 26), yields at Kranskop and Richmond should, on average, be expected to be between 11 TCH and 13 TCH superior to those at Umvoti. The danger of attaching too much meaning to these (statistical) differences is demonstrated by the models presented in Appendix 10 where the inclusion of PDI in the regression results in field yield differences between Umvoti and Richmond being reduced to non-significant levels for the prediction model.

Definitive answers on yield differences between the districts can best be determined in carefully designed agronomic trials – if this information is important to management. It is interesting to note, however, that the field records showed no evidence of interactions between *district* and *season*, suggesting that good and poor seasons are experienced with similar severity across the study area, within BRG 5. The inclusion of *district* in the models precluded the use of longitude and latitude as yield predictors because these parameters were recorded at the estate scale and were therefore aliased with the locality variables (Appendix 7).

Field aspect was identified as an important yield predictor in the explanatory model. Significant interactions with *rain total for cycle* were also found (Table 25). Initial statistical tests using *aspect* with four levels (north, south, east and west) produced unsatisfactory results, probably because of the problems experienced defining pure classes for each field (Section 3.5.1). The two-level aspect variable was shown to be a statistically superior indicator of field yielding capacity, besides being soundly based in terms of the expected aspect effects on crop growth in the southern hemisphere. It was particularly satisfying to show significant interactions between rain (crop water supply) and aspect which favoured the hotter, northern and western aspects with increasing water supply, both between seasons of differing rainfall and within seasons for crop cycles receiving different amounts of rain as a result of different ratoon and harvest months. It is likely that *aspect* did not contribute significantly to the prediction model because *rain total for cycle* was deliberately excluded as a yield predictor and therefore *rain* × *aspect* could not be used to include the *aspect* main effect, as in the case of the explanatory model (Section 4.2.2).

Altitude was included as a significant yield predictor in both models, field yields being reduced by between 5 TCH and 6 TCH, on average, for every 100 m increase in altitude. Mean field altitude was used with the rationale that it functions as a proxy for general temperature differences, averaged across all seasons because it remains constant for all crop cycles. Decreasing sugarcane yields are expected in response to decreasing temperature and therefore the negative *altitude* regression coefficients for Models 1.1 and 2 (Tables 25 and 26 respectively) were anticipated. During model building the substitution of terrain and frost damage (Appendix 7) for altitude was investigated, but resulted in statistically inferior models. There was some evidence that *terrain* was statistically a better yield predictor than *altitude* in multi-phase landscapes at Richmond. This was not pursued because firstly, the relation was restricted to Richmond, and secondly, *altitude* was an alternative significant yield predictor based on a sound physical concept, viz. reducing temperatures with increasing altitude.

Considerable difficulty was experienced defining pure field soil type classes because field boundaries seldom correspond with soil changes (Section 3.3). The target of using classes with a type-purity of 70% could be achieved only by incorporating the undifferentiated soil zones (Anon., 1994) with the dominant soil zones (on an area weighted basis) for each field. Two categories of mixed soil types were also required for all fields to be classified. It follows that statistical yield responses to *soil type* will have been distorted by the relatively impure groupings incorporated in the soil classes. Since yields had been recorded at a field scale only, it was impossible to refine the scale of the investigation to allow for more meaningful soil groupings. It is encouraging that in spite of these data limitations significant yield differences were found between most of the *soil type* levels. Soil groupings based on colour and humus content could not be included in the model because these were aliased with *soil type* by definition (Section 3.3). Soil nitrogen, soil carbon, lithology and texture (Appendix 7) were tested as alternative yield predictors to soil type, but always resulted in poorer models.

Based on statistical groupings of the soil types, the presence of yellow subsoils is associated with lower yields than on red subsoils. In terms of pedogenesis, yellow and red subsoils are formed under remarkably similar conditions, the colour being determined by the extent of soil iron hydration (Soil Classification Working Group, 1991). Since yellow subsoils are often indicative of slightly moister localised soil conditions than red subsoils, it is surprising that the yield relation is not reversed to favour the yellow soil types in the models. An alternative explanation of the formation of yellow subsoils, especially those with humic topsoils and overlying red B2-horizons, i.e. Kranskop soil form, is that organic acids leaching into the B1-horizon from the topsoil either prevent the formation of haematite (red) or convert haematite to goethite (yellow) (Soil Classification Working Group, 1991). In this situation one might be able to attribute lower yields from yellow subsoils to more acidic conditions than for red ones and therefore expect to find some evidence of an association between subsoil acidity and subsoil colour. No such statistical relation was found for the 127 subsoil analyses (Appendix 7) although this may have been distorted by the data from Umvoti which represents very acidic, but exclusively red, subsoils. The yield relation therefore remains poorly understood. In terms of this study it is only possible to state that based on 246 observations for the humic red soil type and 248 observations for the humic yellow and red / yellow soil type, there is significant evidence of superior yields being produced on humic red soils over 13 seasons. Partial regression coefficient differences for the models suggest that humic red soils produce about 4 TCH to 6 TCH more sugarcane than the yellow-type equivalents. The interpretation of these results is tenuous both because of the regression techniques used and the unquantified error introduced by using impure soil type groups. Carefully designed scientific experiments would therefore be essential to quantify the absolute yield differences that could be expected between soil types.

Managers have been puzzled by the superior yields achieved on orthic red soils, on average, compared with the humic soils (Table 25 and Table 26). Conventional agricultural wisdom teaches that humic topsoils are generally superior to orthic topsoils for cropping. In general terms, humic topsoils form under cooler and moister conditions than orthic topsoils where the rate of organic matter mineralization is slower than its rate of accumulation. Cultivation of these soils stimulates mineralization which is associated with the release of important plant nutrients, a factor incorporated into the FAS plant nutrient recommendations (Anon., 1991; 1996). The more favourable moisture regime in environments where humic soils are common may therefore be an advantage in rainfed cropping systems. Such climatic differences should not be expected to be as pronounced when humic and orthic topsoils are found in adjacent fields as in Richmond. Numerous questions arise in trying to explain the observed yield trends, all of which lie beyond the scope of this study. Are factors which retard organic mineralizing rates – the chemical activity of the soil – perhaps partly responsible for the lower sugarcane yields on humic soils? Speculation that FAS nutrient recommendations for orthic topsoils have been better refined than those for humic topsoils appear unfounded on the basis of this study because no evidence of interactions between soil fertility, soil type and fertilizer application rates were found. The nature of this study and the quality of the available data are not appropriate to make conclusive statements about such an important agronomic issue. It must be remembered that the database has only 63 observations for orthic topsoils, all from Richmond. The yield relations reported for orthic red soils may therefore be distorted by sample-size problems, although the yields are highly significantly different. In this context it is encouraging that extension staff in the Richmond area have, over the years, also concluded that the orthic red soils generally outperform their humic equivalents (Hellmann, 1998, personal communication [22]). This aspect of sugarcane production may therefore warrant focussed research in future, especially since humic topsoils are common in the BRG 5 region of the Midlands sugarcane belt.

The variables related to soil rooting volume (ERD, profile gravel, profile rocks) and water holding characteristics (TAW, topsoil and subsoil clay) were all useful yield predictors with the exception of profile rocks (Appendix 7). Most of these variables have high bivariate correlations among themselves, e.g. topsoil and subsoil clay, and some were calculated as composite variables of the others, e.g. TAW calculated as a function of ERD, topsoil and subsoil clay (Anon., 1994).

---

[22] Mr DB Hellmann, Midlands South Extension Officer,
South African Sugar Association Experiment Station,
Private Bag X02, Mount Edgecombe, 4300

Great care was consequently required when modelling these parameters because multicollinearity problems could give rise to misleading statistical relations. In general, yields increased in response to increasing rooting volume and the (associated) increase in water holding capacity. The final explanatory and prediction models both used *ERD* as a yield predictor at the expense of the other variables. A strong interaction between *district* and *ERD* was established for the explanatory model with *ERD* being statistically important only at Umvoti. At Kranskop and Richmond *ERD* did not contribute significantly to yield and the partial regression coefficient for *ERD* at Richmond is negative, suggesting that yields decrease with increasing soil depth. It must be kept in mind however, that this relation is non-significant, i.e. the coefficient could be zero, and the prediction model shows a significant yield response to increasing ERD, irrespective of district.

Of the remaining physical field attributes that were tested, none contributed significantly to the regression models. It is likely that the effect of trees, which are restricted to the field edges, did not feature in the models because yield data were recorded only for whole fields. Should yields ever be recorded at the sub-field level it is plausible that tree effects could feature in models for field edges and perhaps even field boundary-blocks along the contour. Visual differences in sugarcane growth, favouring the absence of trees, can be observed on the estates throughout the year.

Field slope also proved unimportant for explaining cane yields, suggesting that in the short-term, steep slopes can be cultivated without sacrificing yields. This relation should, however, be evaluated against the range of the data used. Mondi are a conservation conscious company and strive to be responsible custodians of their land holdings. This management approach is reflected in mean field slopes planted to sugarcane ranging between 3% and 19%, which are moderate by industry standards. Within this range of slopes it is perhaps unrealistic to anticipate negative yield responses because the present land-use is sound. An equally plausible explanation for the findings is that the mean field slope is a relatively meaningless variable, and therefore a poor yield predictor, because of the large within-field slope class variations (Section 3.5.2).

In general it is encouraging that the models have isolated physical field attributes that have historically been heralded as agronomically important factors and that it was possible to demonstrate this statistically using commercial field records.

5.1.2 Climate variables

The climate variables *season*, *rain total for cycle* and *rain for 12 months' growth* were important yield predictors. The role of *season* in the explanatory model is not clear. The 13 harvest seasons could be grouped using five classes, all significantly different from each other (Table 24). Partial regression coefficients suggest yield differences of up to 33 TCH between the season classes and it would be useful to understand the underlying relations defined by this variable. Annual rainfall patterns, as opposed to rain total accumulated for each crop cycle, may contribute towards some of the yield variation captured by the season variable, although Table 27 shows no consistent trends. It is realistic to expect seasonal temperature and radiation differences to be incorporated into the variable since *altitude*, the temperature and radiation proxy in the models, remains constant for a field. No appropriate data were available to separate these important effects. Under rainfed conditions agriculturalists expect considerable yield variation between seasons – traditionally ascribed mainly to differences in annual rainfall. Since *rain total for cycle* is included in the explanatory model, the *season* yield differences are in addition to this effect. It is not possible to explain the dynamics of *season* because of limitations of the regression techniques and the field data. The CANEGRO model could be used to simulate expected seasonal yield differences with appropriate historic data for conditions similar to those on the Mondi estates. This may provide some insight into the yield variations one should attribute to climatic factors for the 13 seasons and the meaning of *season*. Such a specialist investigation lies beyond the scope of this study.

The full crop cycle classification (Section 3.11) could be used as an alternative yield predictor to the rainfall variables in both models. The rainfall variables were used because these yield relations were statistically stronger than those for the crop cycles and continuous variables should be preferred to categorical variables. The yield relations suggest that the crop cycle types function mainly as a proxy for broad crop – rainfall relations and account for little additional yield variation. The crop age variables (harvest age and summer months growth) were also dropped from the models when the rainfall variables were used as yield predictors. Although cane yield increases with increasing age, rainfall is probably a better yield predictor because water is usually a yield limiting factor under rainfed conditions (Thompson, 1976; Inman-Bamber, 1995). The rain total accumulated for each crop cycle also tends to increase with increasing harvest age, implying that the variables are not truly independent. Rainfall may also be a superior yield predictor to crop age when yields from different seasons are compared because the yield – age relation is variably affected by years of high and low rainfall. The relative importance of crop age and rainfall variables for within-season yield comparisons was not investigated.

The yield – rainfall relation for *rain total for cycle* is statistically stronger than *rain for 12 months' growth*. When *rain total for cycle* was excluded from the prediction model, *rain for 12 months' growth* became a statistically significant yield predictor (Model 2). The interaction between *aspect* and *rain total for cycle* has been discussed in Section 5.1.1. A significant non-linear relation between yield and *rain for 12 months' growth* was found, indicating that yield responses to high rainfall were proportionally smaller than to low rainfall during the first year's growth. This relation may be partially attributed to the period of incomplete canopy when soil evaporation comprises a large proportion of total evaporation. Since the period of incomplete canopy comprises a relatively small proportion of the total growing period compared with the proportion for 12 months' growth, the absence of a similar curvilinear relation for yield – *rain total for cycle* is reasonable.

The rain accumulated for two months' before crop cycle starts and one month before harvest variables were tested as yield predictors because of their importance in other yield-factor studies (Panol, 1974). They were shown to be of no statistical value as yield predictors in this study.

The exclusion of *season* and *rain total for cycle* from the prediction model caused a large decrease in model accountability. The variables that were subsequently included in the prediction model but which did not feature in the explanatory model can be used as statistically weaker alternatives for these predictors. Variables included in the explanatory model but which were dropped from the prediction model also need to be considered. The substitution of *rain for 12 months' growth* for *rain total for cycle* has already been discussed. The other variables affected by these substitution effects will be considered at appropriate stages during the discussion.

5.1.3 Agronomic variables

More than 40 crop nutrient parameters comprising topsoil and subsoil fertility analyses, leaf analyses and agronomic inputs were included in the field database and represent the largest component of the Spencer Holley field record system (Appendix 7). It is disappointing that only two of these variables were significant yield predictors in both models, given their prominence in the field records.

N fertilizer was a highly significant yield predictor in both the explanatory and prediction models. It was not surprising that this yield response was linear over the range of N applications (43 – 184 kg ha^{-1} per crop cycle) included in the database (Appendix 7) since these rates are moderate in terms of general agriculture. (Up to 300 kg ha^{-1} an^{-1} N is recommended for rainfed tropical pastures in BRG 5 (Manson *et al.*, 1993)). The partial regression coefficient for N fertilizer is also particularly large, suggesting a 0.2 TCH increase for every 1 kg ha^{-1} applied N. In real terms the magnitude of this responses is unlikely and the danger of interpreting partial regression coefficients independently of the other variables must be borne in mind. The FAS fertilizer recommendations on which the N applications were based, have been determined using a large database of agronomic trial results. These recommendations formed the basis of fertilizer applications (Section 3.8) although actual application rates were often distorted by incorrect field areas. Nevertheless, 50% of the observations lie within the FAS N recommendations of between 100 kg ha^{-1} and 130 kg ha^{-1} for the Mondi fields (Appendix 7).

Soil N levels are not routinely determined in South Africa and residual soil N reserves are generally small. Where N mineralization has been shown to contribute to crop nutrition FAS fertilizer recommendations have been adjusted to accommodate this source of nutrients (Anon., 1991; 1996). It is perhaps for this reason that yields could be related directly to N fertilizer inputs. The soil fertility levels and fertilizer applications for nutrients other than N were generally poorly related to yield and the question arises whether or not the soil samples were representative of the average field conditions. It may be worth while reviewing the soil sampling procedure used on the Mondi estates to ensure that representative samples are taken – especially since samples were taken from fields before the mature crop was harvested (Section 3.7) under extremely difficult working conditions. The mean fertility reported for a composite soil sample often overestimates field fertility considerably because of a skewed "fertility distribution" of soil cores.

As stated elsewhere, fertilizer applications are generally large for nutrients deficient in the soil and small for those present in sufficiency. For this reason it was more convenient to express the crop nutrient supply as the sum of soil fertility and fertilizer input where these were known. In this manner the total K, total P and total Zn variables were defined (Appendix 7). Similar variables could not be calculated for Ca, Mg and S because the total input of these nutrients from the various fertilizer, lime, gypsum and calmag sources could not be determined.

Total K was a highly significant yield predictor in both models. Sugarcane is known to have a high K requirement relative to other crops and luxury uptake of K is common (Van Dillewijn, 1952). Total K levels vary considerably across Mondi's sugarcane fields and the range for the observations used to develop the models was 1212 kg ha^{-1} K (Appendix 7). The FAS specify no threshold value for total K because this parameter is not used in their recommendations. For this study a value of 440 kg ha^{-1} K was derived from the various FAS thresholds (Section 4.4) although a more appropriate threshold may be calculated from agronomic trial results. Since the median total K is 402 kg ha^{-1} (Appendix 7), more than half the observations used in the statistical analyses might have had sub-optimal K nutrition – although leaf analyses generally showed no K deficiencies. Spencer Holley frequently reduced recommended K fertilizer inputs where leaf analyses from previous crop cycles had indicated high-to-excessive levels of K. It is difficult to determine whether or not this manipulation of K nutrition was prudent because the management of K nutrition is complicated by the phenomenon of luxury uptake in sugarcane.

Given the highly significant yield response to *total K* in both the explanatory and prediction models and the relatively low cost of K, it is anticipated that management will be eager to correct total K "deficiencies" pending the results of more definitive research. Partial regression coefficients suggest only a small yield response of approximately 0.02 TCH per 1 kg ha^{-1} applied K. It is, however, difficult to determine the absolute magnitude of the yield response to K and the limitations of the regression coefficients in this regard have been discussed. Based on the results from this study there appears to be sufficient evidence to support increased levels of K nutrition – at least up to the suggested level of 440 kg ha^{-1} total K on the clayey soils.

Total P was not a significant yield predictor even when the different application rates for plant and ratoon crops were considered. The importance of *PDI* as a yield predictor has already been discussed and its exclusion from the final explanatory and prediction models justified (Section 4.2). It is likely that the different P fixing capacities of soils reduced the meaning of P nutrition in relation to observed yields. Various interactions between soil P, PDI and soil types did not improve the yield – nutrient relation. In spite of the failure to demonstrate a significant relation between yield and P, the importance of P to crops, especially in Midlands soils, has been adequately demonstrated in research trials. Perhaps the relations are too complex to be expressed in the simple mathematical forms tested in this investigation.

The anomalous yield – total Zn relation and the reasons for its exclusion from the prediction model have been discussed (Section 4.2).

Sugarcane yields were not related significantly to lime, gypsum or calmag applications. The rates of agricultural lime used to correct soil acidity problems have been a contentious issue in the Midlands during the last decade and many growers apply up to double the amount of lime recommended by FAS in addition to gypsum. The ameliorative effect of lime is known to occur over a number of seasons and therefore any yield benefit should not be expected to be immediate. This makes it difficult to quantify the effect of lime on crop performance, especially because theory suggests that the main benefit from liming is realized during droughts. Since lime, gypsum and calmag inputs were only recorded for the crop cycles to which they were applied, the effect of the ameliorants could not be traced in subsequent crop cycles. It follows that the nature of the investigation was perhaps inappropriate for establishing a yield response to liming.

One might expect soil acidity and measures of Aℓ toxicity to be more directly related to observed yields since the progressive ameliorative effect of lime in combination with gypsum increases soil pH (lime) and decreases $A\ell^{3+}$ activity (lime and gypsum). No significant yield response to topsoil acidity, EAI or ASI could be demonstrated using the field records. These findings are supported by the general insensitivity of sugarcane to soil acidity and Aℓ toxicity reported by SASEX (Anon, 1996) and other sugarcane research institutions internationally (Van Dillewijn, 1952).

The GCI parameters for soil acidity and Aℓ toxicity could not be tested as meaningful yield predictors because only 44 observations were included in the database (Appendix 7). As more data accumulate in the field records it would be possible to investigate these relations further. Research trials are, however, necessary to quantify the magnitude of yield responses to lime and gypsum. Many such trials have been conducted and the results published (Moberly and Meyer, 1975; Meyer *et al.*, 1991; Schroeder *et al.*, 1993; 1994; 1995; Schroeder, 1997). It is not acceptable scientific process to simply reject the findings of trials because the expected yield responses to soil amelioration could not be shown. Results from carefully designed and executed trials are certainly much more reliable than general trends shown in field records that may or may not be causally related to the predictor variable. An additional complication in the interpretation of results is that low pH inhibits mineralization and observed yield responses to lime, particularly of acid tolerant crops such as sugarcane, may result from an improved N supply rather than from the effect of lime *per se*. Should growers wish to pursue the amelioration of soil acidity and Aℓ toxicity it would be prudent to support additional SASEX research initiatives to gain further insight into these relations. The approach of justifying considerable expenditure on lime and gypsum, based largely on intuition, yield increases supposedly documented in unprocessed field records (these could not be demonstrated statistically using the available field records) and otherwise on research results for crops other than sugarcane, should be questioned. A case may or may not be made for lime and gypsum applications for "soil health" reasons, rather than for crop requirements, depending on the soil acidity and Aℓ thresholds used.

The significance of the topsoil Ca:Mg ratio in the prediction model suggests some evidence of a nutrient imbalance, possibly created by heavy gypsum applications to some fields with the consequent migration of Mg^{2+} (and K^{+}) with the SO_4^{2-} down the soil profile. This has already been discussed in Section 4.2. Since *topsoil Ca:Mg* was included as a yield predictor only when *season* was dropped from the prediction model, it is not possible to determine to what extent the ratio merely acts as a proxy for *season*, and to what extent it is directly related to cane yields. Based on the available evidence, it is recommended that further research on the importance of the Ca:Mg ratio be conducted for the Mondi estates. The FAS provide no threshold Ca:Mg ratio for sugarcane, although a ratio of 3:1 (on a kg ha^{-1} basis) can be calculated using the individual Ca and Mg nutrient thresholds (Section 4.4). The topsoil K:Mg ratio is sometimes evaluated by agronomists to further investigate cation imbalances which may affect the plants' ability to absorb adequate nutrients. This ratio was of no value for yield predictions.

None of the subsoil fertility analyses could be meaningfully related to observed sugarcane yields. An additional disadvantage of these data was the large number of missing observations, candidate models being based on only 127 observations (Appendix 7).

Difficulties were experienced in attempting to define meaningful relations between observed yields and herbicide applications. It was not possible to summarize these field records and, as a compromise, only the three most commonly used herbicide ingredients were tested as yield predictors (Section 3.8.3). None of these were significantly related to yield. When the data were reviewed, it was clear that all crop cycles had received chemical weed control treatments even though some observations had received none of the three active ingredients tested. For this reason any relations would be distorted and the herbicide data are not appropriate for incorporation in general yield equations. Field workers have, however, often associated weeds with reduced yields and extension officers frequently rate management ability in terms of weed populations in sugarcane fields. Since all fields receive chemical weed control treatments it may be more meaningful to differentiate between fields where adequate weed control was achieved and those where weed populations remained a problem. One method of collecting these data would be to count weed populations in representative sample plots for each sugarcane field. This method represents the likely scientific approach appropriate for research trials, requiring the time of a trained worker. Accuracy of the technique would be subject to obtaining a truly representative sample for each field. Where inadequate weed control is achieved with chemicals, fields are usually hand-hoed during the season. Another more practical method of obtaining an index of the weed problem in fields would be to calculate the number of man days used per hectare for hoeing each field, possibly modified by a productivity factor to account for different tasks set by contractors. This approach has not been tested in other studies although, on a conceptual basis, it promises to provide more meaningful weed-data possibly related to yields. Minimal additional administration would be required to include the necessary data in the field records because the information is captured to calculate summaries of labour activities on the Mondi estates. The inclusion of this entry in the field records is therefore recommended.

Foliar nutrient concentrations determined from leaf analyses were expressed as a percentage of the FAS thresholds to overcome the problem of different optimum nutrient contents for various crop classes and varieties (Section 3.7.2). No foliar nutrients were included in the final explanatory and prediction models. This is disappointing because Meyer (1998, personal communication [23]) used foliar nutrients to successfully predict cane yields in SASEX variety-release trials and the question arises why their contribution in this study was non-significant. A possible reason is that samples were taken from crops of very different ages within seasons, making comparisons between samples meaningless. Even though the FAS thresholds are specified independently of crop age at sampling for all nutrients except N, Meyer *et al.* (1989) emphasise the need to consider foliar nutrient concentrations and sampling age concomitantly. These relations may have been further distorted because the crop age at sampling could not be determined accurately (Section 3.10.2). Another confounding factor is that leaf analyses for samples taken from stressed crops were sometimes included in the records. Leaf samples should be taken only from vigorously growing crops – but during the unfavourable growing conditions which occur periodically under rainfed conditions, managers are often required to decide whether or not a crop is adequately stressed to justify sacrificing the information from the leaf analyses. In these situations the value of the information to management is questionable because where serious nutrient deficiencies were revealed, these were dispelled on the basis of stressed crop conditions although satisfactory results were incorporated in the field records. Selective data recording of this nature may cause inconsistent and misleading trends to be observed in the data and consequently these foliar nutrient records are not appropriate for a scientific investigation. In addition to these complications, leaf samples were frequently taken from only selected field sections which exhibited poor growth. This was done to determine whether or not fertility problems were responsible for the poor crop performance and these samples were not intended to be representative of the average field conditions.

It is necessary to question the benefit that Mondi derive from the leaf analyses because the records appear to be meaningless in relation to cane yields. The express purpose of analysing leaf

---

[23] Mr JH Meyer, Head, Crop Nutrition and Soils Department,
South African Sugar Association Experiment Station
Private Bag X02, Mount Edgecombe, 4300

samples is to confirm that sufficient fertilizer was applied to supply the crop with adequate nutrition. Spencer Holley also use historic leaf analyses to adjust recommended fertilizer inputs of especially K, based on the ability of previous crops to extract sufficient nutrients in relation to soil analytical results (Section 3.8.1). Where serious nutrient deficiencies were revealed by leaf analyses and Spencer Holley recommended supplementary fertilizers, these were never applied during the 13 seasons' records. When these recommendations are not implemented for practical reasons (the application of fertilizer to a one year old crop of sugarcane is unpleasant, labour-intensive work) the requirement of a leaf sample from each crop cycle should, perhaps, be re-assessed. The FAS recommend a soil analysis for plant crops only, and base fertilizer recommendations for subsequent ratoon crops on this soil analysis, using leaf analyses to confirm adequate nutrition for each ratoon. Alternatively, fertilizer recommendations are based on soil samples for single crop cycles and no leaf analyses are used. Perhaps Spencer Holley's inclusion of both a soil and leaf sample for each crop cycle is excessive, especially considering managements reluctance to implement corrective fertilizer recommendations based on leaf analyses and the poor association found between cane yields and foliar nutrient concentrations in this study.

Ratios of foliar macro-nutrients were calculated (Appendix 7) and tested as yield predictors because they remain fairly stable irrespective of crop age at the time of leaf sampling. This was done in an attempt to overcome some of the data distortions expected because of the inability to calculate sampling ages accurately and the wide range of crop ages at leaf sampling. None of the nutrient ratios were significant yield predictors, suggesting that factors other than sampling age problems detract from the anticipated association between cane yields and foliar nutrient concentrations.

The sugarcane variety planted, its plant / ratoon status and harvest season were the crop management variables significantly associated with cane yields. A highly significant yield benefit from using the improved generation of sugarcane varieties bred by SASEX in preference to older varieties was identified in the field records. The partial regression coefficient differences show that N12 and N16 yield approximately 10 TCH more than NCo376 and NCo293 on Mondi's Midlands estates. As stated previously, the magnitude of these partial coefficients should be used with the necessary discretion even though research trials have also shown N12 and N16 to be higher yielding than NCo376 and NCo293. It is encouraging that it was possible to show significant benefits from adopting new technology developed by specialized agricultural research using commercial records. Managers should be encouraged to replant all fields still producing NCo varieties. A decision would need to be made whether to plant N12 or N31, a promising new variety which is expected to supercede N12 in parts of the Midlands.

Variety did not contribute significantly to the prediction model and was dropped as a predictor variable. Most of the observations for NCo376 and NCo293 were from the 1980's while most of the N12 and N16 observations were from the 1990's. As discussed previously, yields during the 1980's were generally superior to those during the 1990's. Some of this yield trend was incorporated in *variety* when *season* was excluded from the prediction model (Section 4.2). This reduced the apparent yield benefit from the new varieties because most observations were for relatively poor seasons, and lead to non-significant differences between the two variety groups. The importance of the season-effect in the prediction model was demonstrated in Section 4.3 where the model was shown to predict absolute cane yields unsatisfactorily unless a calibration coefficient for "season" was used (Figure A11.2, Appendix 11). Such a calibration coefficient could be calculated to adjust yield predictions for each harvest season. Once the season-effect has been accounted for varietal yield differences could become apparent although these should not be "added on" to the model without validating the revised equations.

Plant / ratoon status was a highly significant yield predictor, confirming that higher yields are generally obtained from younger ratoons than from older ones. The crop groupings should be of particular interest to management for guiding the replant decision. Based on the yield trends investigated for the 535 observations from 13 seasons, there were no significant differences within the crop groups. Plant crops, and first and second ratoons yielded equally well on average, and under normal circumstances there would be no reason to replant sugarcane until after the second ratoon. After the second ratoon a significant yield decline was found in the field records. Based on partial regression coefficient differences this yield decline varies between 6 TCH and 10 TCH for crop group 2 in the explanatory and prediction models respectively. Once a crop is allowed to continue beyond the second ratoon there would usually be no reason to replant the sugarcane until after the fifth ratoon because, on average, third, fourth and fifth ratoons were shown to yield equally well. Sixth and seventh ratoon sugarcane produced significantly lower yields than crop group 2, on average. The database included too few observations beyond the seventh ratoon to calculate meaningful yield trends for these crop classes. It follows that, for the average situation, a replant should only be considered after the second ratoon, fifth or seventh ratoon crops. This decision would be influenced by the economics of sugarcane production, especially anticipated sucrose prices. These relations strictly apply only to the data from which they were derived although they should be sufficiently robust and reliable for the average situation, because of the large sample size, to provide management with a useful planning tool. The relations are specific to the Mondi estates in the Midlands and will not necessarily apply to other sugarcane enterprises.

The duration of the fallow period before replanting sugarcane was of no significance as a yield predictor. No relation was expected since the fallow period applies only to plant crops, and first and second ratoons were shown to produce yields equal to those from plant crops. The duration of the fallow period was generally short (Appendix 7), being determined primarily by practical considerations such as completing the necessary field operations and killing the old sugarcane roots. No data were available to investigate the yield effects of long-duration fallows or rotational cropping.

The crop cycles defined in Section 3.11 according to the methods of Hellmann (1988; 1993) were restricted so that only six types were included in the database (Section 4.1.1). The importance of *crop cycle* as a yield predictor was reduced by including rainfall variables in the models. Statistical groupings of similar crop cycles resulted in a variable defining the crop harvest season. The significance of this variable has been discussed (Section 4.2). Harvest season is a relatively more important yield predictor in the prediction model than in the explanatory model. It is not possible to provide conclusive reasons for this difference between the two models because of the limitations of regression analysis. However, the prediction model uses rain for 12 months' growth as a yield predictor and it is likely that any estimate of additional rainfall until harvesting would improve yield predictions, given the importance of rain as a yield predictor. Sugarcane is usually harvested between 21 and 25 months of age on the Mondi estates (lower and upper quartiles,

Appendix 7). It follows that for most crops no rainfall for the final nine to 13 months' growth is accounted for in the prediction model. The proportion of this period that lies in summer months will strongly influence the additional rainfall which crops would normally receive. Since crops harvested in summer will usually grow through more summer months, on average, than those harvested in winter (Appendix 5) and therefore usually receive more rain, this may explain a portion of the larger regression coefficient favouring summer harvests in the prediction model.

Possible interactions between *soil types*, *ERD* and *crop cycle* were investigated because the recommended estate harvesting schedules supposedly favoured fields with shallow soils during the early (summer) part of the harvesting season, those with deep humic-type soils during winter, and those with orthic-type soils during the later (summer) part of the harvesting season (Section 3.6.4). No significant interactions were found. Closer examination of the data showed that these recommendations were not followed rigidly, and that the sequence of field harvesting appeared to be influenced more strongly by other considerations related to practical aspects of the harvesting operation.

Absolute TSH yields should be added to the field records. These yields will probably be less strongly correlated with TCH than relative TSH and therefore different key predictors may be important in the sucrose yield relations. These are of potentially greater important than TCH yield relations because growers should strive to maximise sucrose yields from their estates.

5.2 THE MODELS

Cane production is fairly stable at an industry scale when general wet and dry climatic cycles are accounted for. As the scale of investigation is increased, so the variability of cane yields increases. For individual fields yield differences of up to 100 TCH were recorded for successive crops from the Mondi estates. It is perhaps unrealistic to account for this amount of variation using the available records when the median yield over 13 years is only 88 TCH (Appendix 7). Notwithstanding the limitations of the explanatory model (Model 1.1) caused by the relatively low R^2 and high SE, it is robust because the yield relations are based on a large sample size (513 d.f.) and statistically sound because the regression assumptions were upheld and multicollinearity was not a problem (Section 4.2). The R^2 alone is an inadequate gauge of model prediction, but model validation results using independent data confirmed a highly satisfactory model performance (Section 4.3). The prediction model (Model 2) is statistically inferior to the explanatory model because important predictors were deliberately excluded to make the model more useful for yield predictions. Although validation results for the prediction model were satisfactory in the sense that predicted yields were significantly correlated with observed yields, the model under-predicted yields for an above-average season. Appropriate calibrations improved the accuracy of absolute yield predictions (Section 4.2).

The large SE for both the explanatory and prediction models cause 95% confidence limits of yield predictions for individual fields to be too large for practical planning and management applications. Consequently the models are not suitable for applications at a field scale. This is a serious limitation because the sugar industry requires growers to submit yield estimates for individual fields and estate managers deal with the sugarcane enterprise at this scale for crop husbandry and harvesting operations. The prediction model may, however, provide a useful check of the total cane production expected from an estate when the field yield estimates are updated periodically during the harvesting season when appropriate data are already available for model calibration (Section 4.3).

The 95% confidence interval for yield predictions for the population mean lie within 10% of mean long-term yields for both the explanatory and prediction models. In addition the validation study (Section 4.3) showed 75% of yield predictions using the explanatory model, and 66% of yield predictions using the prediction model, to be within 20% of observed yields. Both models can therefore be used to make reliable yield predictions for broad resource combinations on the estates within the normal range of tolerance for most management and planning applications. Consequently the models should be of interest to planners and decision makers who deal with land-use issues at a more general scale than individual fields. A potential application of the models at this scale has been presented and discussed in Section 4.4.

The linear regression techniques used in this study were appropriate for analysing the data since the assumptions of normality and homoscedasticity were upheld and multicollinearity was not a problem in the models. Principle component analysis and other multivariate analyses have been used as an alternative to linear regression to derive prediction equations. These techniques were not used because the objective of the study was to determine key variables related to yield and it is difficult to determine what parameters are represented by the composite factors that are calculated from the original variables. Concerns were raised about the strongly linear yield response established for most predictor variables. Diminishing yield responses to increases in soil fertility and rainfall might have been expected at the high levels of inputs and the large range of most variables included in the database (Appendix 7). The standard data transformations tested showed significant evidence of a quadratic relation only for *rain for 12 months' growth*. The possibility of other curvilinear relations in the data were investigated using artificial neural networks analysis (NN). Preliminary investigations using TRAJAN NEURAL NETWORKS version 3 software in association with Warren [24] produced extremely poor models (R^2 <20%), probably confirming the findings of regression analysis, viz. that the yield trends were best described by linear functions, and NN was abandoned.

It is not possible to determine the relative success of this investigation because no studies involving the use of field records to develop predictive equations have been published in the South African sugar industry. The analytical approach adopted (Chapter 4) is conceptually different to that used in the field record analyses of Hellmann (1988; 1993) and Hellmann *et al.* (1995) because no *a prior* assumptions were made regarding the importance of yield determining factors or their proxies. In their studies records were stratified using a variable of perceived importance, e.g. crop cycle, and yields compared between these classes, as discussed in Section 2.1. Cane yield equations derived from field records in other countries all used similar analytical techniques to those applied in this investigation. The best models derived by Bouldin (1969, cited by Early, 1980) for the Philippines and Phillips *et al.* (1989) for Hawaii are statistically inferior to the explanatory model (Model 1.1). Only the study by Alvarez *et al.* (1982) for Florida sugarcane produced models with a larger coefficient of multiple determination (0.65) than the

---

[24] Prof PR Warren, School of Mathematics, Statistics, Computer Science and Information Systems
Faculty of Science and Agriculture
University of Natal,
Private Bag X01, Scottsville, 3209

explanatory model (0.55), but this model was not validated and the adequacy of the model is therefore unknown. Direct comparisons between the models are meaningless because they are based on empirical relations and refer to extremely different cropping conditions. Nevertheless, the Alvarez *et al.* (1982) study was based on a similar sample size to this investigation and the goodness of fit statistics give some indication of the amount of yield variation one might hope to explain in commercial field records. The Alvarez *et al.* (1982) model probably accounted for a greater proportion of yield variation than Model 1.1 because predictor variables included detailed temperature and radiation measures, and yields were generally less variable than in the Midlands because crops had access to a perched water table and were harvested annually.

5.3 GENERAL

The accuracy of the data was limited by a number of factors which precluded the development of models with greater accountability. General problems in the field records relating to field areas, the recording of rainfall and harvest ages, the mis-allocation of fertilizer and cane consignments, etc., contribute to unexplained error in the data (Chapter 3). The importance of accurate field areas has been discussed (Section 3.4) and while the New GIS areas are the most reliable field areas available, it is likely that actual areas for certain fields will differ from those defined in Appendix 3. All sugarcane fields should be GPS-surveyed to provide accurate field areas which can be used with confidence for the planning and control of enterprise inputs and outputs. The GPS technique is recommended because it is highly accurate (Gillespie, 1998), individual fields can easily be re-surveyed and areas updated if field layouts are changed when replanting, and Mondi have access to the technology in-house. Alternatively, a more different approach to the management-by-row-length system used for applying fertilizer and tasking cane cutters can be used to improve the accuracy and reliability of the field records. The total row length of sugarcane in each field can be accurately determined from the row lengths harvested by cane cutters and recorded in labour records. This information can in turn be used to control fertilizer applications which are made according to row length using ropes and 500 mL oil cans (Section 3.8.1). The planted field area can then be calculated by multiplying the total row length by the interrow spacing, and used to calculate TCH and TSH yields accurately.

The importance of rainfall as a yield predictor has been demonstrated and it is essential that accurate records are available. More disciplined recording of daily rainfall is essential on the Mondi estates to provide useable data. When records are kept the costs of data collection are incurred irrespective of the accuracy of data, and consequently the implementation of appropriate data quality controls would enhance the value of the Mondi rainfall records. These controls are urgently required at Kranskop where serious inaccuracies and discrepancies were found in recent estate rainfall records. The use of inaccurate or unreliable data in the explanatory and prediction models should be avoided because predicted yields are calculated as a function of the levels specified for the predictor variables.

Crop ages at leaf sampling and harvest cannot be calculated accurately when only the month of planting and harvesting are recorded (Section 3.10). Computers are able to process calendar dates efficiently, allowing ages to be calculated as days. Where appropriate, the crop age in days can be converted to ages expressed in other units of time. For this reason the specific dates of field- and harvesting operations should in future be recorded in the field records.

The Spencer Holley record system has numerous internal checks and balances to ensure general data accuracy but the accuracy of source information can be checked only for discrepancies relative to the field history, e.g. where the harvest age or ratoon number of a crop cycle has been incorrectly recorded. There are no controls to ensure that all the inputs drawn from estate stocks are applied to the fields specified by estate managers. If fertilizers were applied accurately according to row lengths, managers would have detected the problems caused by the field area changes according to the revised estate maps, and the under- and over-application of fertilizers to certain fields would not have occurred (Section 3.4). Similarly, no controls are in place to ensure that the correct field numbers are entered on cane consignment notes. Changes in the field numbering system and the use of identical field numbers on different estates within districts controlled by single managers increase the potential for mis-allocations.

The impure soil type and aspect classes defined to describe some of the natural variation of physical resources between fields will have reduced the effectiveness of these variables as yield predictors by increasing the unexplained error within classes. The use of area weighted means for field altitude and ERD is also not ideal because sugarcane plants respond individually to specific resource combinations within fields. This variability inherent in the natural resources poses a major difficulty to crop modellers irrespective of whether yield models or growth models are used. The solution to this problem is not simple and is constrained mainly by the scale of the investigation (Section 2.4).

The explanatory and prediction models are derived from data for which the individual sugarcane field represents the lowest level of spatial aggregation and consequently mean field conditions must be specified, ignoring all within-field variability of resources, inputs and outputs (yield). In the long-term field boundaries could be repositioned to correspond better with broad resource boundaries. These are generally of a gradual nature and field layouts are influenced by numerous other practical considerations which take precedence over natural resource boundaries. It is therefore unlikely that this approach would be feasible. At the other extreme, growth models such as CANEGRO simulate physiological processes at the individual plant level. When field yields are modelled this is done by multiplying the predicted individual plant yield by the expected plant population of the field. This simple "scaling-up" procedure extrapolates the natural resource conditions specified of an individual prediction uniformly across the field and also assumes uniformly-optimum plant nutrition. These assumptions are unreasonable, and it follows that growth models are also inadequate for predicting crop yields at a field level.

Precision farming and site specific agricultural techniques discussed in Section 2.5 provide an alternative for actively managing crops at the sub-field level in response to natural resource, soil fertility and weed population differences. Automated data collection techniques provide detailed site specific data and computer software is available to process and analyse these records to determine detailed management recommendations. The site specific management approach may in future allow the determination of yield responses attributable to specific resource combinations with greater reliability. This may improve our understanding of the sugarcane production system and possibly also the accountability of yield models based on site-specific records. Precision farming promises more efficient use of resources by strategic placement of fertilizers and herbicides within fields. The extent to which this technology will be adopted in the sugar industry will depend largely on its cost effectiveness.

The yield models presently used in the South African sugar industry predict potential and attainable cane yields, mainly in response to estimated crop water use. Factors are used to reduce yields to levels that are thought to be achievable under commercial cropping conditions. The explanatory and prediction models aim to predict actual yields on the Mondi estates and therefore do not contribute directly to, or improve, the existing models. They should, however, be seen to provide an additional approach in the crop modelling process which has not previously been available in the South African sugar industry. After all, the aim should not be to replace functional models and the model ultimately chosen for any particular application is usually constrained by the availability and suitability of input data.

The explanatory and prediction models were derived from actual commercial yield and production data. Predicted yields consequently apply to actual expected yields, specific to Mondi's Midlands estates. Since differences in management ability are known to have a large effect on yields, different statistical yield relations may apply outside the study area. The contribution of management ability to cane yield variation could not be quantified in this investigation because different managers were responsible for each district and therefore the management contribution to cane yields will have been incorporated in *district*. In this context it is also appropriate to re-emphasise that the explanatory and prediction models are based on empirical yield – factor relations and the statistical relations identified in this study are not necessarily evidence of a cause – effect relation between a variable included in the model and yield (Section 2.2.1).

Key parameters strongly associated with actual sugarcane yields were successfully identified for Mondi's Midlands sugarcane estates based on the data contained in field records and other secondary sources of information. It is possible that other parameters which were not tested in this investigation are critical determinants of cane yield – temperature variables in particular are an obvious omission from the data available for model building, but there is no objective method of identifying these parameters or quantifying their importance in yield relations. Yield models were developed using the relations identified between cane yield and the key parameters associated with actual yields (Section 4.2) and were confirmed to be appropriate for predicting cane yields on the Mondi estates (Section 4.3). Ideally yield predictions should be made for individual fields but both the explanatory and prediction models are not suited to applications at this scale. Yield predictions for broad resource combinations can, however be made reliably using the models and this information can be used to identify real differences in production potential on different areas of the Mondi estates. For this reason the models may also provide a convenient tool for site selection, either for the expansion or reduction of the Mondi sugarcane enterprise. The explanatory model is functionally sounder than the prediction model (Section 4.2) and its yield predictions are also more reliable (Section 4.3). Consequently the explanatory model should be used in preference to the prediction model when users are not constrained by the availability of input data. Important predictors are, however, usually unavailable when yield predictions are required, i.e. before harvesting, and then the prediction model may provide a useful tool, especially when calibrated. For this reason the similar yield classes predicted by both models in Section 4.4 is a reassuring result.

It is disappointing that the majority of the agronomic parameters recorded in the field records were not consistently associated with cane yield. These factors should only be recorded for purposes other than relating them to crop performance and some rationalization of the number of variables included in the field records may be possible. A number of possible reasons for the poor association between cane yield and agronomic records have already been discussed. These factors may or may not be relevant. Field records in their unprocessed form are, however, useful for compiling an exception report to management to focus attention on outstanding crop husbandry operations. They are also invaluable for diagnostic purposes in fields with disappointing crop performance. However, based on the finding of this study, the field records should be used for little more than detailing the inputs and husbandry practises a problem field received because no direct association between these factors and yield could be shown. When field records are kept in a database program (such as the revised Spencer Holley record system and CANEPRO), generalized summaries of the data can be generated with ease. Managers should be cautious when using these summaries for decision making because the information can be extremely misleading since the field records may comprise a substantial number of crop cycles that are unrepresentative of the "normal" crop production system on an estate for various reasons (Section 4.1.1).

As more data accumulate in the field records for the Mondi estates the yield models may be improved and refined. The aim should be to develop models that can be applied at a field scale and provide reliable yield predictions. Management efforts should be directed towards establishing correct field areas and improving the general accuracy of the field records. It is, however, unlikely that these improvements alone will allow better yield models to be developed. In order to improve model accountability it will also be necessary to better quantify the natural within-field variation. Precision farming and site-specific agricultural technology holds the greatest promise of achieving these objectives.

6 CONCLUSIONS

1 Key variables associated with sugarcane yields (TCH) were identified from field records and other secondary sources of information. These were locality, aspect, altitude, soil type, effective rooting depth, season, rainfall, sugarcane variety, plant / ratoon status, crop cycle, N and K nutrition, and topsoil Ca:Mg ratio. The variables accounted for up to 55% of the observed variation in commercial sugarcane yields (Model 1.1).

2 A yield model suitable for predicting actual commercial sugarcane yields (TCH) six months in advance of the harvesting season was developed (Model 2).

3 Normal sugarcane yields are approximately 88 TCH, on average, on Mondi's Midlands estates. The 95% confidence limits ranged between ±34 TCH and ±37 TCH for individual field predictions, and between ±8.4 TCH and ±6.1 TCH for the population mean, depending on the model used.

4 Model validation confirmed the satisfactory performance of the models. The accuracy of absolute cane yield predictions could be improved by calibrating the models for each harvesting season.

5 Extensive editing of the commercial field records was required to provide data suitable for research purposes. Unrepresentative observations were excluded from investigations using reasonable *a prior* criteria. The availability of large data sets was essential to ensure sufficient appropriate data for modelling.

6 Linear regression analysis was an appropriate technique for analysing the field records.

7 Models may be improved by including temperature and radiation data, using more accurate agronomic records and by better describing within-field variation of the natural resources and site-specific cropping conditions in general.

6.1 FURTHER RESEARCH

1 The potential of using the models outside the study area, but within the Moist Midlands Mistbelt bioresource group, should be tested. Similar models should be developed to establish absolute sucrose yield relations.

2 The CANEGRO model should be used to further investigate the importance and meaning of *season*.

3 The poor performance of humic soils compared with orthic soils, identified statistically in this investigation and by extension officer's field observations, needs to be studied.

4 The importance of total K nutrition (available soil K + fertilizer inputs) needs to be confirmed. Appropriate nutrient thresholds should be established if the parameter is meaningfully related to yield.

5 The reason for greater yields generally being achieved from summer harvested crops compared with those harvested during winter on the Mondi estates needs to be established.

6 Continued experiments to investigate the importance of lime and gypsum in sugarcane are needed to resolve the contention relating to appropriate levels of these inputs. Existing experiments should be expanded to consider soil "health" issues in addition to the potential sugarcane yield response to soil amelioration. The importance of the topsoil Ca:Mg ratio as a cane yield determining factor should also be investigated in these experiments.

7 Improved methods of quantifying within-field variation of natural resources need to be developed.

6.2 MANAGEMENT AND FIELD RECORDS

1 Accurate sugarcane field areas should be determined as a matter of priority.

2 The partial harvesting of fields should be avoided as far as possible.

3 All NCo-varieties should be replaced with suitable new varieties.

4 The size of the residual weed population in sugarcane fields after chemical treatments should be quantified.

5 The quantities of lime and gypsum applied to fields should be reviewed.

6 The use of the GCI / FERTREC soil laboratory should be reviewed since fertilizer recommendations are based on FAS nutrient threshold.

7 The importance of leaf analyses to the sugarcane enterprise should be reassessed.

8 The interrow spacing of sugarcane should be specified for plant crops.

9 The tillage and cultivation practices used in plant and ratoon crops should be recorded.

10 The general accuracy of all records should be controlled to ensure the collection of useable data.

11 The specific dates of field operations should be recorded.

12 The extent of crop damage caused by frost, animals, lodging, etc. should be quantified and the age of the crop recorded.

13 The absolute sucrose content of sugarcane should be included in the field records.

7 SUMMARY

1 The study area comprises 146 sugarcane fields from five selected KwaZulu-Natal Midlands sugarcane estates, between 29.01 °S and 29.87 °S latitude, and 30.29 °E and 30.94 °E longitude. The estates lie in the Kranskop, Umvoti and Richmond districts, all within the Moist Midlands Mistbelt bioresource group. The climate on the estates is broadly representative of that for the higher-potential rainfed sugarcane production region of the Midlands sugarcane belt.

2 Field data were obtained for 19 seasons and comprehensive records of soil fertility, foliar nutrient concentrations, agronomic inputs and yields were available for 13 seasons. A detailed natural resource inventory in a GIS format was available for the estates. Reliable local climatic data could not be obtained and only estate rainfall records were used.

3 Extensive editing and cleaning of the field records was essential to provide a data set suitable for electronic manipulation and statistical analysis. The accuracy of all records were confirmed as far as possible, deleting observations for which problems were identified that could not be corrected. The data set was restricted using five exclusion criteria to remove unrepresentative observations from the database. These criteria were:

a exclude all crop cycle types that have less than 20 observations;
b exclude all observations without fertilizer input records;
c exclude all observations without topsoil fertility analyses;
d exclude all observations for which agronomic inputs differ by more than 10% in the estate managers' and Spencer Holley records, and could not be corrected when checking the records; and
e exclude all observations for mixed sugarcane varieties, and varieties with less than 20 observations.

4 The agricultural resource and production data were analysed using GENSTAT version 5, release 3.2 software. Multiple linear regression models were developed to establish variables strongly associated with actual sugarcane yields (TCH). Key physical field

attributes associated with sugarcane yield included locality, aspect, altitude, soil type and effective rooting depth. Season and rainfall were important climatic variables for yield predictions. Of the factors influenced by management, sugarcane variety, plant / ratoon status, crop cycle, N and K nutrition and the topsoil Ca:Mg ratio were important yield predictors. These variables accounted for up to 55% of the observed variation in commercial sugarcane yields. A yield model suitable for predicting actual commercial sugarcane yields (TCH) six months in advance of the harvesting season was developed.

5 Multiple regression was confirmed to be an appropriate analytical technique for analysing the field records. The regression assumptions normality and homoscedasticity were upheld and multicollinearity was confirmed not to be a problem in the models. The models are robust because they have 513 or more degrees of freedom.

6 The models were validated using an independent data set and model performance was confirmed to be acceptable. The accuracy of absolute cane yield predictions could be improved by calibrating the models for each harvesting season. The 95% confidence limits ranged between ±34 TCH and ±37 TCH for individual field predictions, and between ±8.4 TCH and ±6.1 TCH for the population mean, depending on the model used. This confidence interval is too large to use the models for applications at a field level because normal sugarcane yields on the Mondi estates are approximately 88 TCH. The models may be useful for managers and planners who deal with land-use issues at a more general scale than the individual field, where yield predictions for the population mean often apply, since the 95% confidence intervals lie within 10% of normal yields – which is well within the tolerance for most management and planning applications.

7 The models were used to generate thematic maps of predicted sugarcane yields in response to various natural resources combinations under "average" cropping conditions, linked to a GIS on an *ad hoc* basis for data input and display purposes. This demonstrates how the models can be used, for example to identify cropping areas of similar yield potential on the estates.

8 REFERENCES

Alvarez, J, Crane, DR, Spreen, TH and Kidder, G (1982). A yield production model for Florida sugarcane. *Agric Syst* **9:** 161-179.

Anonymous (1977). Laboratory manual for South African sugar factories. South African Sugar Technologists' Association. South Africa.

Anonymous (1991). Preparation of whole cycle and single crop fertilizer recommendations using soil analyses (effective from 1 April 1991). South African Sugar Association Experiment Station. South Africa.

Anonymous (1994). Forest Industry Soils Database (FSD) Co-operative: Soil survey standards for consultants. Unpubl. Document.

Anonymous (1996). Preparation of whole cycle and single crop fertilizer recommendations using soil analyses (effective from 1 October 1996). South African Sugar Association Experiment Station. South Africa.

Bell, CJ (1981). The testing and validation of models. In: Rose, DA and Charles-Edwards, DA Eds. (1981). *Mathematics and plant physiology*. Academic Press, London.

Blackmore, S (1994). Precision farming: An introduction. *Outl Agric* **23:** 275-280.

Blume, H (1983). Environment and cane sugar yield. *Proc Int Soc Sug Cane Technol* **18:** 277-288.

Boote, KJ and Loomis, RS (1991). The prediction of canopy assimilation, 1991. In: *Modelling crop photosynthesis from biochemistry to canopy.* CSSA Special Publication. **No. 19:** 109-140.

Bouldin, DR (1969). Yields of sugar in Victorias district as influenced by fertilizer additions, soil properties, varieties and climatic factors. Mimeo, univ. of Phil. Los Banos, Phil., 27.

Boyce, JP (1968). Plant crop results of a row spacing experiment at Pongola. *Proc S Afr Sug Technol Ass* **42:** 136-141.

Burrough, PA, Van Deusen, W and Heuvalink, G (1988). Linking spatial process models and GIS: A marriage of convenience or a blossoming partnership? *Proc GIS / LIS '88* Falls Church ASPRS / ACSM **2:** 598-607.

Buyeye, SM, Fey, MV and Mott, CJ (1985). Maize growth and acid soil responses to treatment with gypsum and lime. *Proc S Afr Soc Crop Prod Cong* **15:** 366-384.

Camp, KGT (1999). A bioresource classification for KwaZulu-Natal, South Africa. Unpubl. M.Sc (Agric.) thesis. Univ. Natal, Pietermaritzburg, South Africa.

Carmean, WH (1975). Forest site quality evaluation in the United States. *Adv Agron* **27:** 209-269.

Chang, YT, Wang, PL and Yao, TP (1983). The characteristics of water consumption in sugarcane, V. Models of sugarcane yield and evapotranspiration in Taiwan. *Taiwan Sugar Res. Inst. Report* **100:** 15-23.

Cluverwell, TL (1984). Field records as an aid to the management of sugarcane crops. *Proc S Afr Sug Technol Ass* **58:** 179-181.

De Haas, AO (1981). Observations of the effects of frost in sugarcane varieties. *Proc S Afr Sug Technol Ass* **55:** 146-148.

De Jager, JM, Van Zyl, WH, Kelbe, BE and Singels, A (1987). Research on a weather service for scheduling the irrigation of winter wheat in the OFS. Report to the Water Research

Commission. WRC Report No. **117/1/87**.

Donkin, MJ (1994). A multivariate reappraisal of a site evaluation study in *Acacia mearnsii* De Wild. *Jl S Afr For Ass* **169:** 1-10.

Doorenbos, J and Kassam, AH (1979). Yield response to water. FAO Irrigation and Drainage Paper No. **33**.

Doorenbos, J and Pruitt, WO (1977). Crop water requirements. FAO Irrigation and Drainage Paper No. **24**.

Early, AC (1980). Farm production data analysis with respect to controllable factors, agronomic inputs and environmental conditions for synchronization of sugarcane field and factory operations. *Proc Int Soc Sug Cane Technol* **17:** 377-395.

Farina, MPW and Channon, P (1988). Acid-subsoil amelioration: II. Gypsum effects on growth and subsoil chemical properties. *Soil Sci Soc Am J* **52:** 175-180.

Farina, MPW and Channon, P (1991). A field comparison of lime requirement indices for maize. *Pl Soil* **134:** 127-135.

Farina, MPW, Sumner, ME, Plank, CO and Letzsch, WS (1980). Exchangeable aluminium and pH as indicators of lime requirement for corn. *Soil Sci Soc Am J* **44:** 1036-1041.

France, J and Thornley, JHM (1984). Mathematical models in agriculture. Butterworths, London.

Gillespie, GL (1998). Developing a sugarcane information system for the Mpumalanga Lowveld. Proceedings of the South African Sugar Industry Agronomist's Association, 26-11-1998.

Glover, J (1972). Practical and theoretical assessments of sugarcane yeild potential in Natal. *Proc S Afr Sug Technol Ass* **46:** 138-141.

Grey, DC (1983). The evaluation of site factor studies. *Jl S Afr For Ass* **127:** 19-22.

Hauser, GF (1973). The callibration of soil tests for fertilizer recommendations. FAO Soils Bulletin No. **18**.

Hellmann, DB (1988). Using FRS to provide advice to growers on the optimum cane age at harvest in the Midlands South area. *Proc S Afr Sug Technol Ass* **62:** 175-179.

Hellmann, DB (1993). The use of FRS data to interpret the effect of different growth cycles on the yield of variety N12. *Proc S Afr Sug Technol Ass* **67:** 88-93.

Hellmann, DB, Wallace, MG and Platford, GG (1995). Interpreting farm sugarcane yields using a Geographic Information System (GIS). *Proc S Afr Sug Technol Ass* **69:** 87-93.

Helwig, OJ (1991). Functions of crop yield from applied water. *Agron J* **83:** 769-773.

Heskith, JD, Baker, DN and Duncan, WG (1971). Simulation and growth and yield in cotton: respiration and the carbon balance. *Crop Sci* **11:** 394-398.

Hoekstra, RG (1976). Analysis of when to plough out a sugarcane field. *Proc S Afr Sug Technol Ass* **50:** 103-113.

Hughes, AD (1992). Sugarcane yield simulation with the ACRU modelling system. Unpubl. M.Sc (Agric.) thesis. Univ. Natal, Pietermaritzburg, South Africa.

Hulbert, EO and Harding, RL (1980). The computer analysis of farm records from an extension area. *Proc S Afr Sug Technol Ass* **54:** 121-126.

Inman-Bamber, NG (1991a). Some physiological factors affecting the optimum age and season for harvesting sugarcane. *Proc S Afr Sug Technol Ass* **65:** 103-108.

Inman-Bamber, NG (1991b). A growth model for sugar-cane based on a simple carbon balance and the CERES-Maize water balance. *S Afr J Plant Soil* **8:** 93-99.

Inman-Bamber, NG (1994a). Temperature and seasonal effects on canopy development and light interception of sugarcane. *Field Crop Res* **36:** 41-51.

Inman-Bamber, NG (1994b). CANEGRO: Its history, conceptual basis, present and future uses. Workshop on research and modelling approaches to examine sugarcane production opportunities and constraints. St Lucia, Queensland, Australia. pp31-34.

Inman-Bamber, NG (1995). Climate and water as constraints to production in the South African Sugar Industry. *Proc S Afr Sug Technol Ass* **69:** 55-59.

Inman-Bamber, NG (1996). Opportunities for increased radiation utilization by sugarcane in South Africa. In: *Research towards efficient and sustainable production.* Wilson, JR, Hogart, DM, Cambell, JA and Garside, AL (Eds). CSIRO Division of Tropical Crops and Pastures, Brisbane, pp162-165.

Inman-Bamber, NG, Culverwell, TL and McGlinchey, MG (1993). Predicting yield responses to irrigation of sugarcane from a growth model and field records. *Proc S Afr Sug Technol Ass* **67:** 66-72.

Inman-Bamber, NG and Stead, BA (1990). A method for choosing the most profitable commercial sugarcane variety. *Proc S Afr Sug Technol Ass* **64:** 1-4.

Inman-Bamber, NG, Singels, A and Muchow, RC (1998). A systems approach to benchmarking for sugarcane production in Australia and South Africa. *Proc S Afr Sug Technol Ass* **71:** 117-122.

Inman-Bamber, NG and Thompson, GD (1989). Models of dry matter accumulation by sugarcane. *Proc S Afr Sug Technol Ass* **63:** 212-216.

Jenny, H (1961). Reflections on the soil acidity merri-go-round. *Soil Sci Soc Am Proc* **25:** 428-432.

Johnson, AKL (1994). The use of models to analyse performance of sugarcane production systems: Use in spatial analysis. *Proc of Workshop on Research and Modelling Approaches to Examine Sugarcane Production Opportunities and Constraints* St. Lucia, Queensland.

Jones, CA and Kiniry, JR (1986). CERES - Maize: A simulation model of maize growth and development. Texas A & M University Press, 194 pp.

Kassam, AH (1977). Report of the agro-ecological zone project 1: Africa. FAO World Soil Resources Report No. **48**.

Keig, G, McAlpine, JR and Pointing, EM (1991). Cane-land productivity monitoring and enhancement - appropriate use of information technology. *Proc Aust Soc Sug Cane Technol* **13:** 138-145.

Kingston, G (1994). Benchmarking yield of sugarcane from estimates of crop water use. *Proc Aust Soc Sug Cane Technol* **16:** 201-209.

Land Type Survey Staff (1988). Land types of the map 2830 Richards Bay. *Mem. Agric. Nat. Resour. S. Afr.* **No. 11.**

Landrey, OP, Schorn, DPK and Truen, RD (1981). Use of a computer as a management aid to crop prediction during time of restriction. *Proc S Afr Sug Technol Ass* **55:** 131-134.

Lonsdale, JE and Gosnell, JM (1975). Effects of age and harvest season on the yield and quality of sugarcane. *Proc S Afr Sug Technol Ass* **49:** 177-181.

Lorber, MN, Fluck, RC and Mishoe, JW (1984). A method for analysis of sugarcane (*Saccharum* sp.) Biomass production systems. *Trans ASAE* **27:** 146-158.

Mann, QV (1991). Some observations of frost damage in sugarcane in the Natal Midlands. *Proc S Afr Sug Technol Ass* **65:** 36-40.

Manson, AD, Milborrow, DJ, Miles, N, Farina, MPW and Johnston, MA (1993). Explanation of the Cedara computerised fertilizer advisory service. Cedara Agricultural Development Institute, Department of Agricultural Development. South Africa.

McCree, KJ (1970). An equation for the rate of respiration of white clover plants grown under controlled conditions. In: *Prediction and measurement of photosynthetic productivity.* Purdoc, Wageningen. pp 221-229.

McGlinchey, MG, Inman-Bamber, NG, Culverwell, TL and Els, M (1995). An irrigation scheduling method based on a crop model and an automatic weather station. *Proc S Afr Sug Technol Ass* **69:** 69-73.

McGlinchey, MG and Inman-Bamber, NG (1996). Effect of irrigation scheduling on water use efficiency and yield. *Proc S Afr Sug Technol Ass* **70:** 55-56.

Meyer, JH (1983). Rapid determination of nitrogen in cane leaves. *Proc S Afr Sug Technol Ass* **57:** 109-112.

Meyer, JH, Tucker, AB and Wood, RA (1997). The SASEX fertilizer advisory service: Over 40 years service to the South African sugar industry. *Proc S Afr Sug Technol Ass* **71:** 42-49.

Meyer, JH, Turner, PET and Fey, MV (1991). Interim evaluation of phosphogypsum as an ameliorant for soil acidity in sugarcane. *Proc S Afr Sug Technol Ass* **65:** 41-46.

Meyer, JH, Wood, RA and Harding, RL (1989). Fertility trends in the South African sugar industry. *Proc S Afr Sug Technol Ass* **63:** 159-163.

Meyer, JH, Wood, RA, McIntyre, RK and Liebbrandt, NB (1983). Classifying soils of the South African Sugar industry on the basis of their nitrogen mineralizing capacities and organic matter contents. *Proc S Afr Sug Technol Ass* **57:** 151-158.

Moberly, PK (1974). The decline in rate of evapotranspiration of fully canopied sugarcane during a winter stress period. *Proc Int Soc Sug Cane Technol* **15:** 694-700.

Moberly PK and Meyer JH (1975). The amelioration of acid soils in the South African sugar industry. *Fert Soc S Afr J* **2:** 57-66.

Monteith, JL (1965). Evaporation and the environment. In: *The state and movement of water in living organisms*. Ed. AE Frogg. Academic Press, New york. pp 205-234.

Monteith, JL and Unsworth, MH (1990). Principles of environmental physics. 2cd Edition. Ed. E Arnold. Hodder and Stoughton, London. 291pp.

Neen, JL, Mann, QV, Domleo, KE and McArthur, DG (1994). A computer based extension tool for an assessment of district resources. *Proc S Afr Sug Technol Ass* **68:** 31-35.

Ometto, JC (1977). An equation for the estimation of agro-industrial sugarcane yield in the Piracicaba region. *Proc Int Soc Sug Cane Technol* **16:** 899-905.

Panol, FY (1974). Sugarcane yield models for production simulation. *Proc Int Soc Sug Cane Technol* **15:** 871-883.

Papajorgji, P, Jones, JW, Hoogenboom, G and Calixte, JP (1994). Exploring concepts for linking GIS and crop models. *Proc Computers in Agriculture* **5:** 598-602.

Penning de Vries, FWT and Van Laar, HH (1982) (Eds). Simulation of plant growth and crop production. In: *Simulation Monographs* **No 29.** Purdoc, Wageningen. 271 pp.

Phillips, VD, Sing, D, Khan, MA, Takahashi, PK (1989). Geographical information system for site selection of biomass energy crops. *Proc Hawaii Sug Technol* **48:** FE9-FE14.

Prins, A, Bornman, JJ and Meyer, JH (1997). Economic fertilizer recommendations for sugarcane in KwaZulu-Natal, incorporating risk quantification using the KYNO-CANE computer program. *Proc S Afr Sug Technol Ass* **71:** 38-41.

Reeve, NG and Sumner, ME (1970). Effects of aluminium toxicity and phosphorus fixation on crop growth of Oxisols in Natal. *Soil Sci Soc Am Proc* **34:** 263-267.

Roth, G (1966). Studies of the effects on sugarcane of damage caused by frost and associated micro-organisms. *Proc S Afr Sug Technol Ass* **40:** 343-350.

Schönau, APG (1987). Problems in using vegetation or soil classification in determining site quality. *Jl S Afr For Ass* **141:** 13-18.

Schönau, APG and Aldworth, WJK (1991). Site evaluation in Black Wattle with special reference to soil factors. *Jl S Afr For Ass* **156:** 35-43.

Schroeder, BL (1997). Advances in dealing with soil acidity. SASEX Annual Report 1996 / 97: pg 23. South African Sugar Association. South Africa.

Schroeder, BL, Meyer, JH, Wood, RA and Turner, PET (1993). Modifying lime requirements for sandy, to sandy clay loam soils in the Natal Midlands. *Proc S Afr Sug Technol Ass* **67:** 49-52.

Schroeder, BL, Robinson, JB, Wallace, MG and Turner, PET (1994). Soil acidification: Occurrence and effects in the South African sugar industry. *Proc S Afr Sug Technol Ass* **68:** 70-74.

Schroeder, BL, Turner, PET and Meyer, JH (1995). Evaluation of a soil aluminium saturation index for use in the South African sugar belt. *Proc S Afr Sug Technol Ass* **69:** 46-49.

Schulze, RE (1982). Mapping mean monthly temperature distributions for Natal by trend surface analysis. *S Afr J Sci* **78:** 246-248

Schulze, RE, Domleo, FB, Furniss, PW and Lecler, NL (1995) Crop Yield Estimation. In: Schulze, RE (1995). *Hydrology and Agrohydrology: A text to accompany the ACRU 3.00 agrohydrological modelling system.* Water Research Commission, Pretoria, Report **TT 69/95**. South Africa.

Schulze, RE and Maharaj, M (1991). Mapping A-pan equivalent potential evaporation over southern Africa. In: *Proceedings, 5th South African National Hydrological Symposium*, Stellenbosch. NSI, Cape Town. 4B-4-1 to 4B-4-9.

Schutz, CJ (1990). Site relations for *Pinus patula* in the Eastern Transvaal escarpment. Unpubl. PhD (Agric.) thesis. Univ. Natal, Pietermaritzburg, South Africa.

Shainberg, I, Sumner, ME, Miller, WP, Farina, MPW, Pavan, MA and Fey, MV (1989). Use of gypsum on soils: A review. *Adv Soil Sci* **9:** 1-111.

Shih, SF (1988). Sugarcane yield, biomass and water use efficiency. *Trans ASAE* **31(1):** 142-148.

Shih, SF and Gascho, GJ (1980). Water requirement for sugarcane production. *Trans ASAE* **23(41):** 934-937.

Singels, A, Kennedy, AJ and Bezuidenhout, CN (1998). Irricane: A simple irrigation scheduling procedure for sugarcane using automatic weather stations. *Proc S Afr Sug Technol Ass* **71:** 117-122.

Smith, JMB (1993). Land capability classification. Farm Business Planning in Natal. Resource and Information Section, Cedara Agricultural Development Institute. Department of Agricultural Development. South Africa.

Smith, JMB (1994). Crop, pasture and timber yield estimate index. Cedara Report No. **N/A/94/4**. Department of Agricultural Development: Natal Region. South Africa.

Soil Classification Working Group (1977). Soil classification: A binomial system for South Africa. Department of Agricultural Technical Services, Pretoria, South Africa.

Soil Classification Working Group (1991). Soil classification: A taxonomic system for South Africa. Memoirs on the Agricultural Natural Resources of South Africa, No. 15, SIRI, Department of Agricultural Development, Pretoria, South Africa.

Steel, RGD and Torrie, JH (1981). Principles and procedures of statistics - a bionometrical approach. McGraw-Hill Book Company, Singapore.

Sweet CPM and Patel, R (1985). Changes in cane yield of irrigated variety NCo376 due to season and their implications when evaluating field performance. *Proc S Afr Sug Technol Ass* **59:** 210-214.

Thomas, JR and Oether, GF (1976). Estimation of crop conditions and sugarcane yields using aerial photography. *Proc Am Soc Sug Cane Technol* **6:** 93-99.

Thompson, GD (1965). The effects of trash conservation on soil moisture and the sugarcane crop in Natal. *Proc S Afr Sug Technol Ass* **39:** 143-157.

Thompson, GD (1976). Water use by sugarcane. *S Afr Sug J* **60:** 593-600 and 627-635.

Thompson, GD (1977). Irrigation of sugar cane. *S Afr Sug J* **61:** 126-131 and 161-174.

Thompson, GD (1986). Agrometeorological and crop measurements in a field of irrigated sugarcane. Mount Edgecombe Research Report No. **5.** South African Sugar Association Experiment Station, South Africa.

Thompson, GD (1988). Comparisons of the growth of plant and first ratoon crops of sugarcane at Pongola. *Proc S Afr Sug Technol Ass* **62:** 180-184.

Thompson, GD and Boyce, JP (1968). The plant crop results of two irrigation experiments at Pongola. *Proc S Afr Sug Technol Ass* **42:** 143-153.

Thompson, GD and Boyce, JP (1971a). Estimating water use by sugarcane from meteorological and crop parameters. *Proc Int Soc Sug Cane Technol* **14:** 813-826.

Thompson, GD and Boyce, JP (1971b). Comparisons of measured evapotranspiration of sugarcane from large and small lysimeters. *Proc S Afr Sug Technol Ass* **45:** 169-177.

Thompson, GD and De Robillard, PJM (1968). Water duty experiments with sugarcane on two soils in Natal. *Exp Agric* **4:** 295-310.

Thompson, GD and Du Toit, JL (1967). The effects of row spacing on sugarcane crops in Natal. *Proc Int Soc Sug Cane Technol* **12:** 103-112.

Thompson, GD and Harding, RL (1986). A computerized model for evaluating irrigation schemes for sugarcane. *Proc S Afr Sug Technol Ass* **60:** 177-182.

Thornley, JHM (1976). Mathematical models in plant physiology. Academic Press, New York.

Tim, US, Milner, M and Majure, J (1992). Geographic Information Systems / simulation model linkage: Processes, problems and opportunities. *Proc ASAE* Nashville, Tennessee. No. **92:** 3610.

Tobin, PD and Ellis, RD (1988). The use of microcomputers and spreadsheet programmes to aid replant decision making. *Proc S Afr Sug Technol Ass* **62:** 169-174.

Tovey, DA and Bull, TA (1977). SUCROWS I - A simulation model of sugar-cane growth. *Proc Qd Soc Sug Cane Technol* **44:** 149-155.

Tsuji, GY, Uehara, G and Balas S (eds.). (1994). DSSAT v3. University of Hawaii, Honolulu, Hawaii.

Tucker, AB (1975). The use of crop data from comparable farm groups as an extension aid. *Proc S Afr Sug Technol Ass* **49:** 174-176.

Turner, PET, Meyer, JH and Wood, RA (1992). Some results of field experiments to test lime and phosphogypsum for aluminium amelioration. *Proc S Afr Sug Technol Ass* **66:** 44-49.

Van Antwerpen, R, Meyer, JH and Inman-Bamber, NG (1993). Cane root development and crop modelling. *Proc S Afr Sug Technol Ass* **67:** 73-77.

Van Antwerpen, R, McGlinchey, MG, Inman-Bamber, NG and Bennie, ATP (1996). Estimating root water use in sugarcane. *Proc S Afr Sug Technol Ass* **70:** 57-58.

Van Dillewijn, C (1952). Botany of Sugarcane. The Chronica Botanica Co., Waltham, Mass.

Vazquez, R (1970). Water requirements of sugarcane under irrigation in Lajas Valley, Puerto Rico. *Univ. Puerto Rico Agric. Expt Sta. Bull.* **224:** 5-36.

Wallace, MG (1996). Geographical information system (GIS): Potential applications in the South African sugar industry. *Proc S Afr Sug Technol Ass* **70:** 73-76.

Wilson, J (1960). Some observations on the effects of frost on sugarcane in the cane belt during 1960. *S Afr Sug J* **44:** 837-839.

Wood, RA (1990). The roles of nitrogen, phosphorus and potassium in the production of sugarcane in South Africa. *Fertilizer Research* **26:** 89-98.

Wood, RA, Meyer, JH and Govender, M (1985). A rapid system of cane leaf analysis using x-ray spectrometry and infra-red reflectance. *Proc S Afr Sug Technol Ass* **59:** 195-201.

Yost, R and Pandutama, M (1990). Application of geographical information systems (GIS) to sugarcane production. *Proc Hawaii Sug Technol* **49:** A11-A17.

APPENDIX 1

BIORESOURCE UNITS FOR MONDI'S MIDLANDS SUGARCANE ENTERPRISE

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Page

A1.1 BRU 212 **Kranskop** (representative BRU for sugarcane on Salem and Sutherlands)

Table A1.1: Characteristics of BRU 212.

FEATURE	CHARACTERISTICS	
BRU Code	Yc8	
Terrain		
Terrain type	Rolling, Broken	
Altitude range	819 - 1299 m ASL	
Slope	Moderate, Steep	
Extent of cultivation	Widespread	
Vegetation		
Bioresource Group	Moist Midlands Mistbelt	
BRG Subgroup	2	
All Crop Ecotopes (ha)	20853	71%
Potential Cropping Soil (ha)	17802	61%
High Potential Soil (ha)	15213	52%
Vegetation pattern	Grassland	
Indicator species (Appendix 2, Camp, 1995)	4, 56, 60	
Number of units in this BRU	1	

Table A1.2: Climate for BRU 212.

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Annual
Rainfall													
Median Rainfall (mm)	134	114	100	46	15	7	6	16	43	83	115	128	
Mean Rainfall (mm)	135	124	97	51	34	19	18	28	59	90	119	135	909
Temperature													
Mean (°C)	20.2	20.3	19.4	17.4	15.2	12.8	12.8	14.3	16.0	16.9	17.9	19.6	16.9
Maximum (°C)	25.3	25.4	24.7	23.1	21.3	19.1	19.3	20.8	22.2	22.6	23.2	25.0	22.7
Minimum (°C)	15.2	15.2	14.2	11.8	9.2	6.5	6.5	7.8	10.0	11.2	12.7	14.2	11.2
Heat Units													
Base 10 °C	317	290	292	223	162	83	88	132	181	213	238	298	
Base 4.4 °C	490	448	466	391	335	251	262	306	349	387	406	471	
Base 5 °C	472	431	447	373	317	233	243	287	331	368	388	453	
Evaporation													
Apan (mm)	170	148	144	124	109	98	110	130	147	159	159	180	1679
Sunshine													
Sunshine (hours d^{-1})	6	6	6							6	6	6	6.4

A1.2 BRU 427 Wartburg (Representative BRU for sugarcane on a portion of Canema)

Table A1.3: Characteristics of BRU 427.

FEATURE	CHARACTERISTICS	
BRU Code	Wc33	
Terrain		
Terrain type	Rolling /Broken	
Altitude range	740 - 1134 m ASL	
Slope	Moderate, Steep / Gentle	
Extent of cultivation	Widespread	
Vegetation		
Bioresource Group	Moist Midlands Mistbelt	
BRG Subgroup	2	
All Crop Ecotopes (ha)	13319	87%
Potential Cropping Soil (ha)	11572	75%
High Potential Soil (ha)	3908	25%
Vegetation pattern	Grassland	
Indicator species (Appendix 2, Camp, 1995)	4, 56, 60	
Number of units in this BRU	2	

Table A1.4: Climate for BRU 427.

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Annual
Rainfall													
Median Rainfall (mm)	121	102	94	43	16	5	5	17	42	76	97	104	
Mean Rainfall (mm)	121	109	104	52	33	19	17	27	48	77	96	104	807
Temperature													
Mean (°C)	20.5	20.5	19.7	17.5	15.2	12.7	12.7	14.2	16.1	17.1	18.2	19.9	17
Maximum (°C)	25.6	25.7	25	23.4	21.6	19.4	19.6	21	22.4	22.9	23.5	25.3	22.9
Minimum (°C)	15.4	15.4	14.4	11.8	8.9	6	6	7.5	9.9	11.3	12.9	14.5	11.2
Heat Units													
Base 10 °C	326	298	301	226	161	81	85	131	182	219	245	306	
Base 4.4 °C	499	456	475	394	335	249	259	305	350	392	413	480	
Base 5 °C	481	439	456	376	316	231	240	286	332	374	395	461	
Evaporation													
A-pan (mm)	169	149	144	122	107	97	107	128	145	157	158	180	1662
Sunshine													
Sunshine (hours d^{-1})	6.7	6.7	6.7							6.7	6.7	6.7	6.9

A1.3 BRU 428 **Harden Heights** (Representative BRU for sugarcane on a portion of Canema)

Table A1.5: Characteristics of BRU 428.

FEATURE	CHARACTERISTICS	
BRU Code	Xc16	
Terrain		
Terrain type	Rolling	
Altitude range	740 - 1275 m ASL	
Slope	Moderate / Steep	
Extent of cultivation	Widespread	
Vegetation		
Bioresource Group	Moist Midlands Mistbelt	
BRG Subgroup	2	
All Crop Ecotopes (ha)	18271	83%
Potential Cropping Soil (ha)	15714	72%
High Potential Soil (ha)	13947	64%
Vegetation pattern	Grassland	
Indicator species (Appendix 2, Camp, 1995)	4, 56, 60	
Number of units in this BRU	1	

Table A1.6: Climate for BRU 428.

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Annual
Rainfall													
Median Rainfall (mm)	139	116	100	47	17	4	5	17	41	83	97	121	
Mean Rainfall (mm)	147	124	107	49	28	11	11	30	65	87	105	124	889
Temperature													
Mean (°C)	20.3	20.3	19.5	17.2	14.8	12.2	12.3	14	15.9	16.9	18	19.7	16.8
Maximum (°C)	25.6	25.7	25	23.3	21.4	19.1	19.4	20.9	22.4	22.8	23.5	25.4	22.9
Minimum (°C)	15.0	15	14	11.2	8.2	5.3	5.3	7.1	9.5	11	12.5	14.1	10.7
Heat Units													
Base 10 °C	320	292	293	217	148	67	72	124	178	214	240	301	
Base 4.4 °C	493	450	467	385	322	235	245	297	346	388	408	475	
Base 5 °C	475	433	448	367	303	217	227	279	328	369	390	456	
Evaporation													
A-pan (mm)	172	149	144	124	107	97	108	132	149	158	159	181	1682
Sunshine													
Sunshine (hours d^{-1})	6.8	6.8	6.8							6.8	6.8	6.8	7.2

A1.4 **BRU 330** **Byrne** (Representative BRU for sugarcane on Uplands and Greenhill)

Table A1.7: Characteristics of BRU 330.

FEATURE	CHARACTERISTICS	
BRU Code	Yc14	
Terrain		
Terrain type	Rolling	
Altitude range	800 - 1676 m ASL	
Slope	Steep, Moderate	
Extent of cultivation	Widespread	
Vegetation		
Bioresource Group	Moist Midlands Mistbelt	
BRG Subgroup	3	
All Crop Ecotopes (ha)	308	1%
Potential Cropping Soil (ha)	283	0%
High Potential Soil (ha)	262	0%
Vegetation pattern	Grassland	
Indicator species (Appendix 2, Camp, 1995)	4, 56, 60	
Number of units in this BRU	4	

Table A1.8: Climate for BRU 330.

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Annual
Rainfall													
Median Rainfall (mm)	149	133	124	54	19	5	7	19	46	83	118	136	
Mean Rainfall (mm)	157	156	142	55	31	12	14	23	43	79	116	151	979
Temperature													
Mean (°C)	19.5	19.6	18.7	16.4	14.0	11.6	11.6	13.2	15.1	16.1	17.2	18.9	16.0
Maximum (°C)	24.8	24.9	24.1	22.4	20.4	18.2	18.4	19.9	21.5	22.0	22.8	24.5	22.0
Minimum (°C)	14.3	14.3	13.2	10.5	7.7	5.0	4.8	6.5	8.7	10.2	11.7	13.3	10
Heat Units													
Base 10 °C	295	270	268	192	125	47	49	99	153	189	216	276	
Base 4.4 °C	469	428	442	360	299	215	223	273	321	363	384	449	
Base 5 °C	450	411	423	342	280	197	204	254	303	344	366	431	
Evaporation													
A-pan (mm)	170	147	141	120	102	96	106	129	145	154	153	177	1638
Sunshine													
Sunshine (hours d^{-1})	6.4	6.4	6.4							6.4	6.4	6.4	6.7

A1.5 BRU 400 Baynesfield (BRU for sugarcane on a portion of fields AS09 and AS19 on Uplands)

Table A1.9: Characteristics of BRU 400.

FEATURE	CHARACTERISTICS	
BRU Code	Wb13	
Terrain		
Terrain type	Rolling / Broken	
Altitude range	609 - 1102 m ASL	
Slope	Steep / Moderate	
Extent of cultivation	Widespread	
Vegetation		
Bioresource Group	Moist Coast Hinterland Ngongoni Veld	
BRG Subgroup	7	
All Crop Ecotopes (ha)	14469	72%
Potential Cropping Soil (ha)	13272	66%
High Potential Soil (ha)	8628	43%
Vegetation pattern	Grassland /-/ Bush Clumped Grassland	
Indicator species (Appendix 2, Camp, 1995)	4, 19, 25, 37, 47, 51, 56, 59, 60	
Number of units in this BRU	2	

Table A1.10: Climate for BRU 400.

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Annual
Rainfall													
Median rainfall (mm)	116	100	101	45	19	6	8	19	43	72	98	104	
Mean Rainfall (mm)	121	107	106	52	26	15	14	27	48	79	110	110	816
Temperature													
Mean (°C)	21.0	21.1	20.3	18	15.5	13	13	14.5	16.4	17.5	18.7	20.4	17.5
Maximum (°C)	26.1	26.3	25.7	24	22.1	20	20.1	21.4	22.7	23.2	24	25.8	23.5
Minimum (°C)	16.0	16	15	12.1	9	6.1	6	7.8	10.2	11.8	13.3	15.1	11.5
Heat Units													
Base 10 °C	342	314	320	241	171	90	93	140	193	234	260	323	
Base 4.4 °C	515	472	493	409	344	258	267	314	361	407	428	496	
Base 5 °C	497	455	475	391	326	240	248	295	343	389	410	478	
Evaporation													
A-pan (mm)	171	152	144	121	105	94	104	127	143	155	158	180	
Sunshine													
Sunshine (hours d^{-1})	6.1	6.1	6.1							6.1	6.1	6.1	6.6

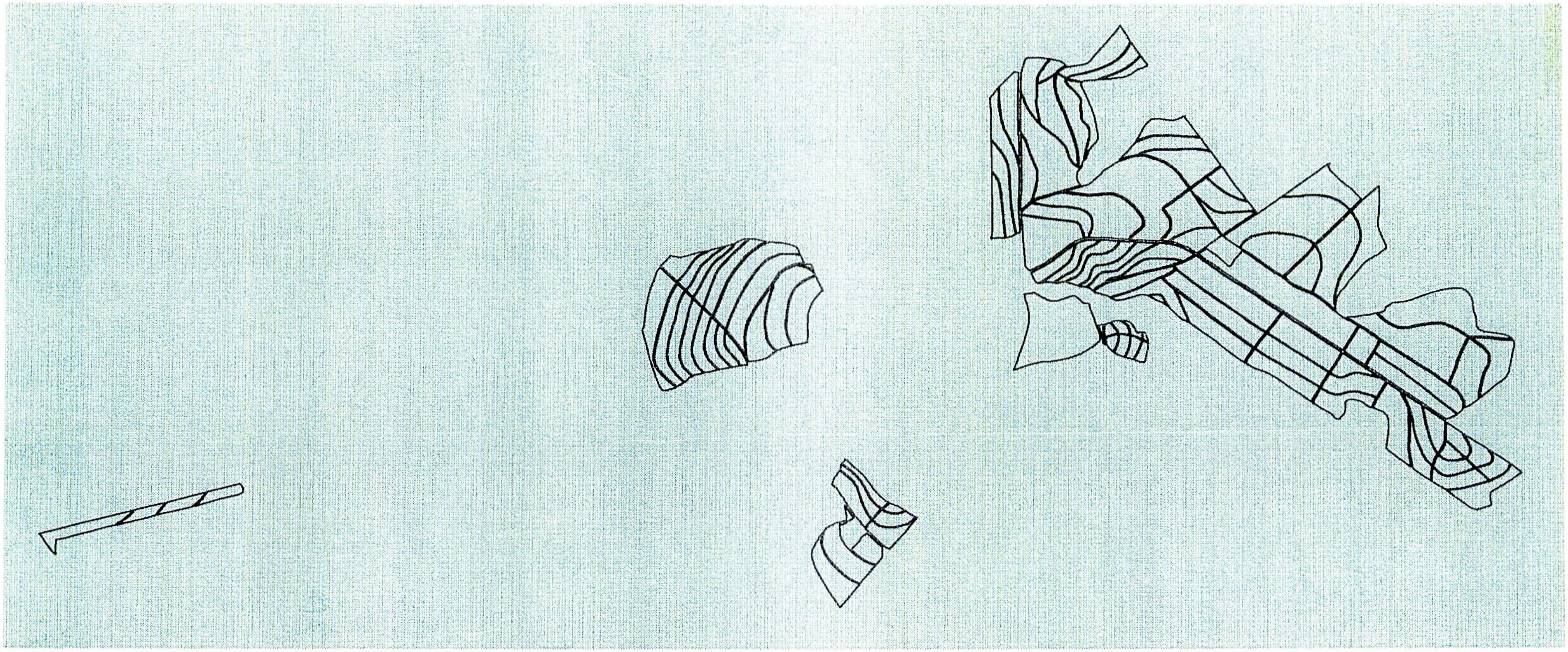


Figure A1.1: Map of bioresource units representative of the sugarcane fields on Salem at Kranskop.

☐ Sugarcane field layout

▨ BRU 212

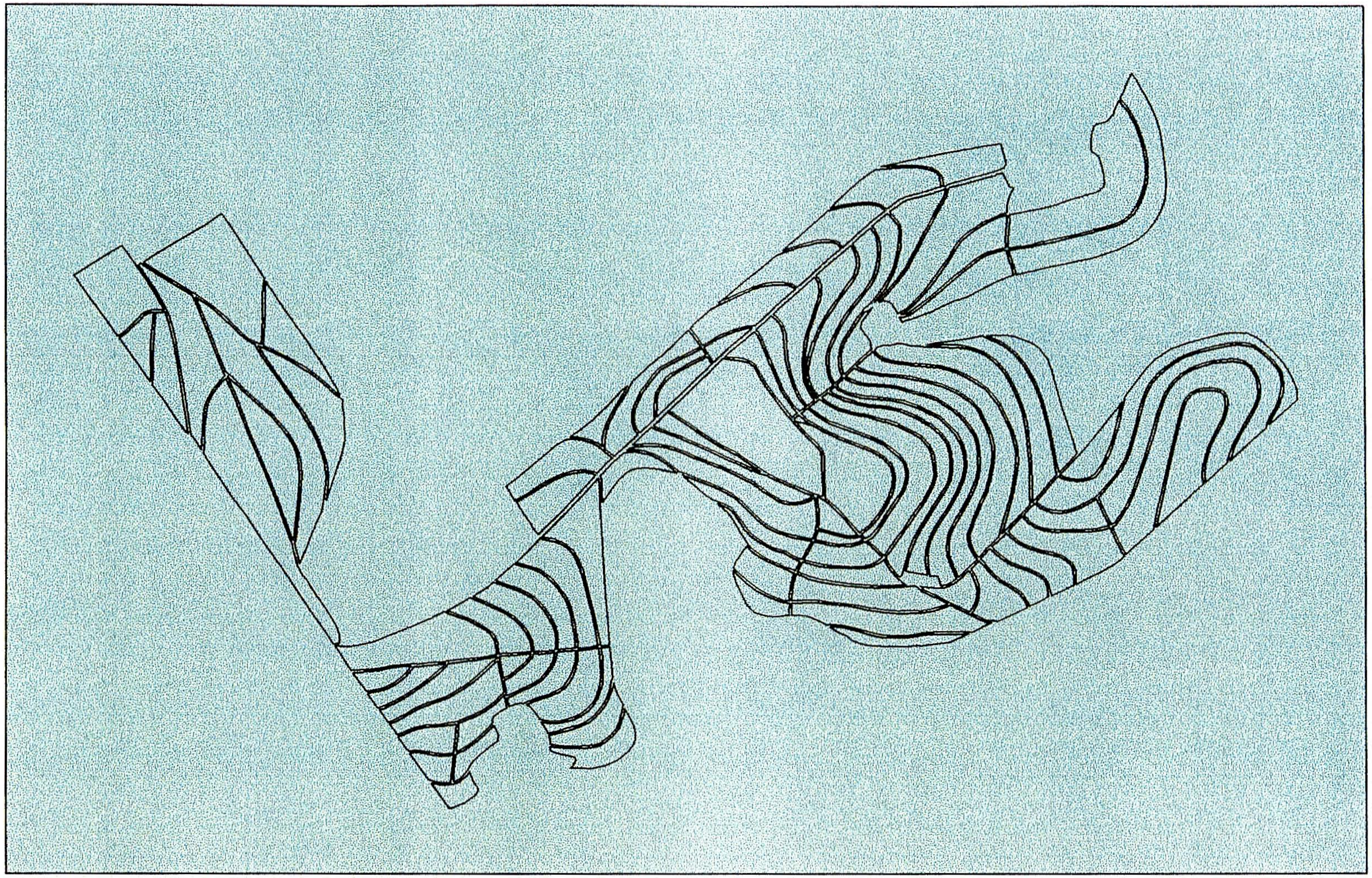
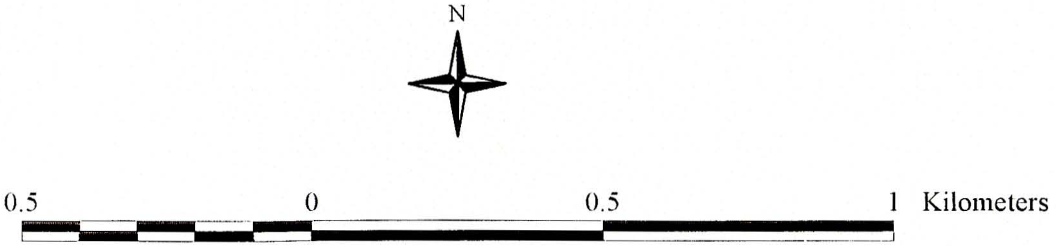


Figure A1.2: Map of bioresource units representative of the sugarcane fields on Sutherlands at Kranskop.

☐ Sugarcane field layout

▨ BRU 212



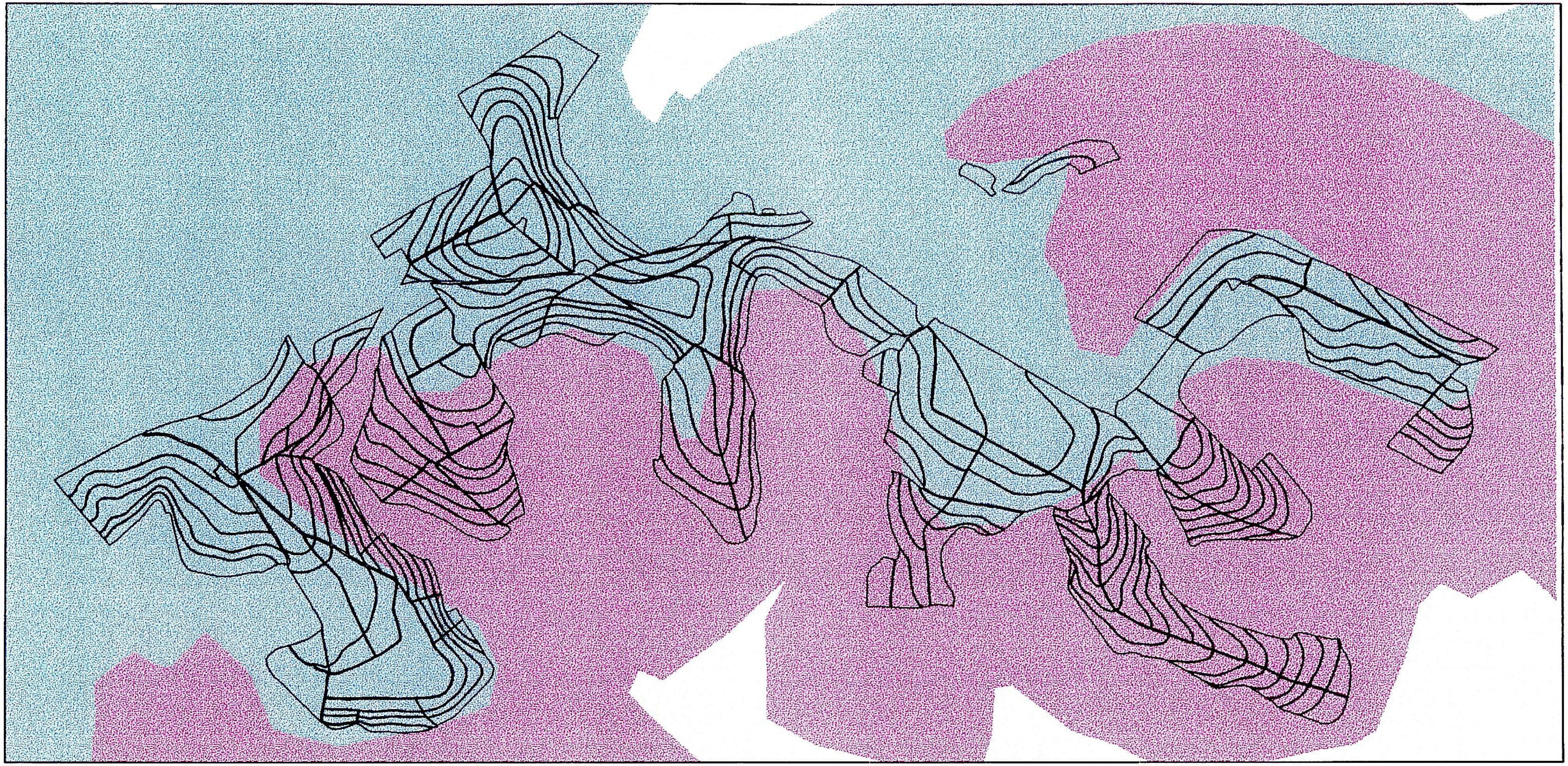


Figure A1.3: Map of bioresource units representative of the sugarcane fields on Canema at Umvoti.

Figure A1.4: Map of bioresource units representative of the sugarcane fields on Uplands at Richmond.

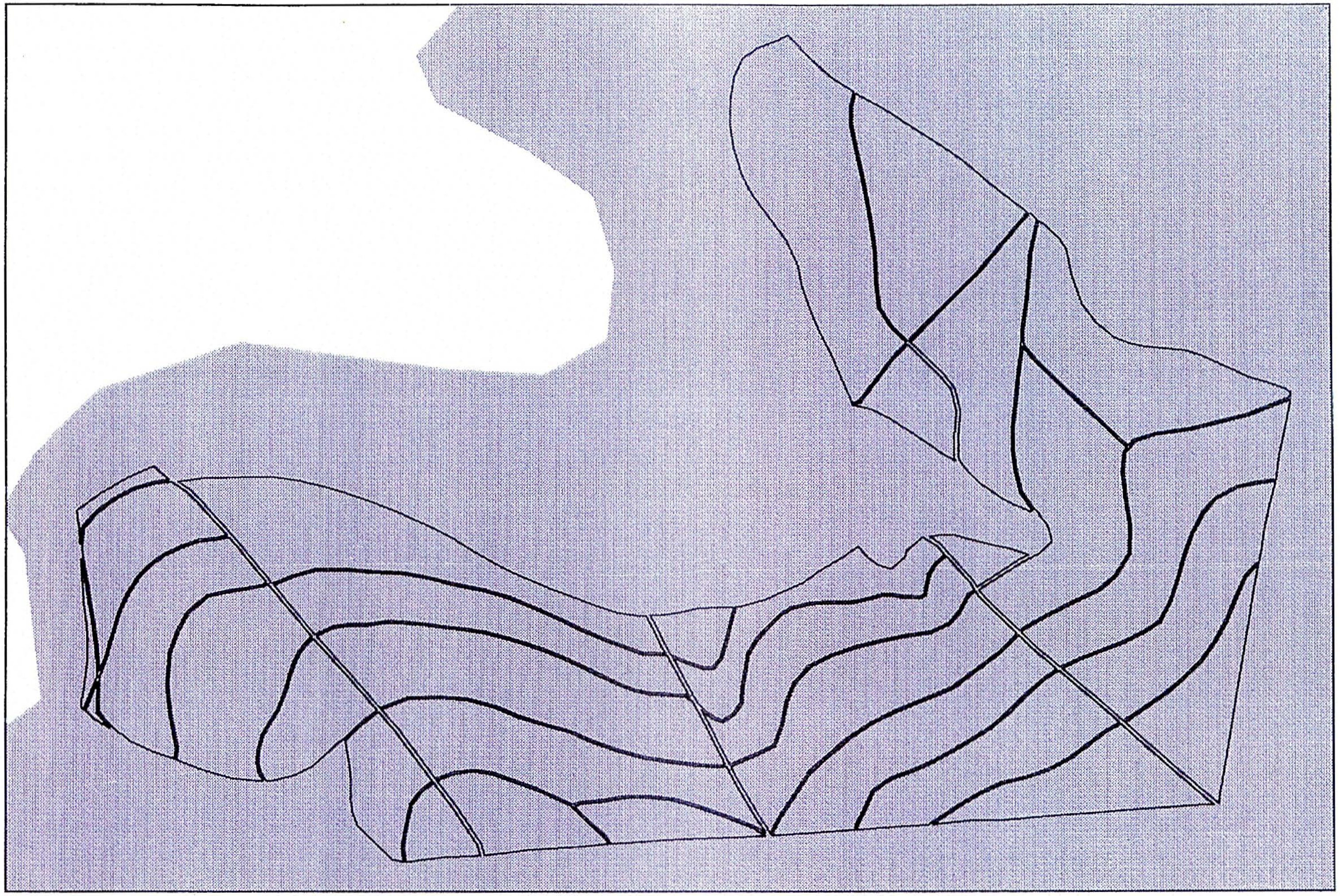


Figure A1.5: Map of bioresource units representative of the sugarcane fields on Greenhill at Richmond.

Appendix 2

MONTHLY RAINFALL DATA (1976 – 1997) FOR MONDI'S MIDLANDS SUGARCANE ENTERPRISE

TABLE OF CONTENTS

LIST OF TABLES

A2.1 Kranskop

Table A2.1: Rainfall data – Salem estate office.

Year	Jan (mm)	Feb (mm)	Mar (mm)	Apr (mm)	May (mm)	Jun (mm)	Jul (mm)	Aug (mm)	Sep (mm)	Oct (mm)	Nov (mm)	Dec (mm)	Annual (mm)
1976	86	160	268	50	34	3	9	21	50	96	103	95	975
1977	165	155	125	36	14	8	0	38	61	154	134	96	986
1978	90	139	96	51	0	0	5	59	60	148	45	165	858
1979	112	100	80	34	17	0	17	43	61	56	51	89	660
1980	76	36	48	35	12	0	3	34	126	57	134	80	641
1981	198	164	25	28	64	30	10	60	75	84	237	97	1072
1982	204	98	166	75	16	3	7	6	40	229	68	76	988
1983	148	43	76	37	14	4	29	45	7	112	191	161	867
1984	311	200	123	100	34	25	73	60	31	122	134	87	1300
1985	179	301	67	5	4	9	6	9	19	219	112	218	1148
1986	147	88	169	41	1	40	10	23	29	142	104	299	1093
1987	326	136	192	38	22	77	15	88	550	88	170	131	1833
1988	204	234	154	26	25	50	55	36	58	118	180	261	1401
1989	40	288	50	40	46	19	15	15	21	125	327	175	1161
1990	81	144	153	84	18	2	1	10	17	135	117	200	962
1991	183	251	128	7	52	27	25	23	79	140	34	95	1044
1992	108	88	27	72	0	0	0	24	47	75	140	90	671
1993	60	131	90	39	16	4	0	34	59	124	97	143	797
1994	180	31	182	15	4	8	22	80	7	76	55	104	764
1995	122	24	135	81	104	71	5	9	11	120	183	289	1154
1996	127	416	118	33	32	0	118	20	14	155	122	97	1252
1997	137	149	88	78	23	82	46	32	79	128	199	92	1133
Statistics: 1976 - 1997 (22 years' data)													
Mean	149	153	116	46	25	21	21	39	68	123	134	143	1039
CV%	49	64	51	56	97	127	135	68	164	36	52	49	26
Minimum value	40	24	25	5	0	0	0	6	7	56	34	76	641
25th percentile	95	91	77	33	13	2	5	22	20	90	99	93	860
Median	142	142	121	39	18	8	10	34	49	123	128	101	1052
75th percentile	182	191	154	67	34	29	24	56	61	142	178	173	1153
Maximum value	326	416	268	100	104	82	118	108	550	229	327	299	1833

Table A2.2: Rainfall data – Sutherlands estate office.

Year	Jan (mm)	Feb (mm)	Mar (mm)	Apr (mm)	May (mm)	Jun (mm)	Jul (mm)	Aug (mm)	Sep (mm)	Oct (mm)	Nov (mm)	Dec (mm)	Annual (mm)
1976	111	129	269	68	43	0	0	30	54	103	110	98	1015
1977	157	216	129	21	20	9	0	58	62	136	134	96	1038
1978	129	145	97	64	4	0	5	61	58	159	57	197	976
1979	164	101	82	49	19	0	18	46	78	48	54	96	755
1980	128	40	66	46	23	0	4	56	140	71	143	119	836
1981	235	113	27	33	70	28	12	69	80	67	240	87	1061
1982	186	103	207	71	10	5	9	6	39	210	66	75	987
1983	148	44	70	37	18	4	41	43	8	102	208	202	925
1984	339	245	123	103	28	26	66	61	32	124	145	96	1388
1985	190	349	52	10	4	13	4	8	26	203	110	210	1179
1986	135	97	228	40	1	41	9	24	26	140	115	299	1155
1987	318	140	155	28	20	79	18	88	560	96	163	154	1819
1988	208	261	168	21	56	56	43	35	70	128	154	262	1462
1989	59	309	77	43	49	18	17	18	23	111	341	262	1327
1990	98	190	190	137	20	4	0	67	25	129	92	224	1176
1991	163	209	133	5	60	26	30	20	78	154	34	69	981
1992	115	85	33	71	0	0	3	25	28	34	79	86	559
1993	75	51	102	55	20	3	0	34	64	252	116	153	925
1994	169	42	173	17	0	7	36	82	9	124	59	160	878
1995	159	36	119	96	97	75	7	13	28	179	270	197	1276
1996	179	368	114	45	23	3	111	17	9	133	104	82	1188
1997	112	126	73	70	21	121	21	32	101	91	213	92	1073
Statistics: 1976 - 1997 (22 years' data)													
Mean	163	155	122	51	28	24	21	41	73	127	137	151	1090
CV%	42	65	52	63	91	137	129	59	157	42	57	47	24
Minimum value	59	36	27	5	0	0	0	6	8	34	34	69	559
25th percentile	118	88	74	29	12	3	4	21	26	98	82	93	938
Median	158	128	117	46	20	8	11	35	47	126	116	136	1050
75th percentile	184	214	165	70	39	28	28	60	76	151	161	201	1186
Maximum value	339	368	269	137	97	121	111	88	560	252	341	299	1819

A2.2 Umvoti

Table A2.3: Rainfall data – Mistley estate office.

Year	Jan (mm)	Feb (mm)	Mar (mm)	Apr (mm)	May (mm)	Jun (mm)	Jul (mm)	Aug (mm)	Sep (mm)	Oct (mm)	Nov (mm)	Dec (mm)	Annual (mm)
1976	229	236	231	58	25	0	0	3	57	167	81	71	1158
1977	144	53	121	49	13	8	0	23	49	97	140	143	840
1978	167	99	96	49	3	13	10	48	62	142	107	107	903
1979	111	91	86	25	27	1	30	43	39	45	54	87	639
1980	76	47	74	27	0	0	0	13	106	53	120	73	589
1981	89	89	41	26	47	16	5	50	74	58	146	47	688
1982	133	124	136	22	11	0	3	4	28	166	82	98	807
1983	74	33	146	65	15	5	25	24	18	108	169	121	803
1984	220	105	97	80	4	15	38	42	21	111	109	120	962
1985	177	194	54	0	4	1	6	2	28	187	88	125	866
1986	72	106	114	37	0	25	1	29	25	127	66	190	792
1987	185	108	159	24	8	68	5	88	246	60	91	220	1262
1988	147	153	147	21	22	37	23	21	27	52	186	162	998
1989	82	170	33	20	32	19	14	3	29	78	193	157	830
1990	67	82	167	45	10	0	0	63	13	77	120	187	831
1991	122	106	112	9	33	9	5	16	57	116	108	132	825
1992	98	82	119	47	0	0	3	21	22	61	64	73	590
1993	54	104	113	20	12	0	0	26	43	170	69	143	754
1994	124	82	173	29	1	9	28	53	12	96	26	96	729
1995	64	34	200	52	15	48	6	17	16	85	244	214	995
1996	288	215	151	14	32	4	112	134	9	118	95	119	1291
1997	142	95	174	66	16	84	32	20	54	72	152	86	993
Statistics: 1976 - 1997 (22 years' data)													
Mean	130	109	125	36	15	16	16	34	47	102	114	126	870
CV%	47	50	40	58	87	141	157	92	107	42	45	38	22
Minimum value	54	33	33	0	0	0	0	2	9	45	26	47	589
25th percentile	78	82	96	21	4	0	2	16	21	64	81	89	764
Median	123	102	120	28	13	9	6	24	29	97	108	121	831
75th percentile	162	120	157	49	24	18	25	47	56	125	145	154	985
Maximum value	288	236	231	80	47	84	112	134	246	187	244	220	1291

A2:3 Richmond

Table A2.4: Rainfall data – Uplands estate office.

Year	Jan (mm)	Feb (mm)	Mar (mm)	Apr (mm)	May (mm)	Jun (mm)	Jul (mm)	Aug (mm)	Sep (mm)	Oct (mm)	Nov (mm)	Dec (mm)	Annual (mm)
1976	243	83	273	93	44	0	4	56	59	166	59	77	1157
1977	234	53	117	19	17	11	0	22	47	121	80	76	797
1978	131	78	134	93	8	4	4	50	63	107	126	81	879
1979	77	87	97	36	27	0	20	35	54	59	48	81	621
1980	81	36	40	29	14	0	0	16	162	59	111	69	617
1981	152	51	53	59	43	13	8	75	42	54	138	94	782
1982	99	72	192	55	25	23	7	6	39	116	87	108	829
1983	106	72	80	60	36	2	40	43	47	90	165	159	900
1984	139	162	90	47	22	24	29	38	25	113	79	88	856
1985	127	216	57	68	10	19	10	3	58	185	115	148	1016
1986	82	93	137	43	0	36	2	57	37	182	109	111	889
1987	116	95	128	74	14	22	19	133	485	76	229	81	1472
1988	89	200	210	43	59	39	32	21	40	79	139	223	1174
1989	47	222	46	57	16	4	12	0	39	87	399	70	999
1990	82	109	190	65	25	9	1	92	32	121	84	137	947
1991	163	93	117	10	47	12	0	5	70	118	80	53	768
1992	86	30	61	42	0	1	7	0	23	53	73	82	458
1993	88	134	50	48	4	10	0	19	47	148	100	197	845
1994	148	48	287	41	15	13	43	56	12	113	42	202	1020
1995	138	87	234	98	24	66	5	7	24	187	175	414	1459
1996	215	178	99	41	67	4	192	18	27	171	120	194	1326
1997	181	84	68	80	37	80	38	41	65	96	196	107	1073
Statistics: 1976 - 1997 (22 years' data)													
Mean	128	104	125	55	25	18	22	36	68	114	125	130	949
CV%	41	55	60	43	73	119	189	93	144	38	62	62	27
Minimum value	47	30	40	10	0	0	0	0	12	53	42	53	458
25th percentile	87	72	63	41	14	4	3	9	33	81	80	81	805
Median	122	87	108	52	23	12	8	29	45	113	110	101	895
75th percentile	151	128	177	67	37	23	27	55	59	141	139	156	1060
Maximum value	243	222	287	98	67	80	192	133	485	187	399	414	1472

APPENDIX 3

FIELD AREAS FOR MONDI'S MIDLANDS SUGARCANE ENTERPRISE

TABLE OF CONTENTS

LIST OF TABLES

A3.1 Kranskop

Table A3.1: Field areas – Salem estate.

Field number	SASA areas (ha)	CAD areas (ha)	GIS areas (ha)	New GIS areas (ha)
AS01	2.1	2.6	2.2	2.3
AS02	1.8	1.9	1.5	1.7
AS03	3.5	3.6	2.6	3.2
AS04	5.6	5.8	4.3	5.1
AS05	3.9	4.3	3.6	3.9
AS06	4.6	5.0	4.2	4.6
AS07	3.4	3.8	3.2	3.4
AS08	2.5	4.1	2.6	3.0
AS09	4.6	4.8	3.7	4.2
BS01	4.4	4.9	4.6	4.5
BS02	4.9	5.3	3.7	4.5
BS03	3.5	3.7	2.9	3.3
BS04	1.4	1.6	1.0	1.3
BS05	5.0	5.0	4.3	4.6
BS06	4.9	4.8	4.1	4.4
BS07	4.5	4.7	3.9	4.2
BS08	2.2	2.3	1.8	2.0
BS09	4.2	4.5	3.5	4.0
CS01	5.4	5.5	5.1	5.1
CS02	3.0	3.4	2.8	3.0
CS03	1.0	0.9	0.6	0.8
CS04	4.3	4.8	4.3	4.4
CS05	5.8	4.8	4.5	4.4
CS06	5.5	5.7	4.8	5.3
CS07	7.2	7.4	6.6	7.0
CS08	4.6	4.8	3.9	4.3
CS09	3.7	3.8	3.1	3.4
CS10	4.7	4.9	4.4	4.5
CS11	6.1	6.4	5.6	5.9
CS12	7.0	7.2	6.0	6.5
CS13	4.5	4.7	3.5	4.1
Total: 31	129.8	137.0	112.9	122.9

Table A3.2: Field areas – Sutherlands estate.

Field number	SASA areas (ha)	CAD areas (ha)	GIS areas (ha)	New GIS areas (ha)
AS01	2.7	3.4	2.9	3.1
AS02	4.9	5.8	5.0	5.3
AS03	6.3	6.2	5.3	5.7
AS04	4.6	4.8	4.0	4.3
AS05	2.3	2.3	1.7	1.9
AS06	3.7	3.3	2.5	2.8
AS07	1.8	1.6	1.2	1.4
AS08	2.9	3.0	2.3	2.5
AS09	8.4	8.6	7.2	7.6
AS10	4.6	4.9	3.8	4.1
AS11	3.7	3.7	2.9	3.1
AS12	1.7	2.3	1.8	2.0
AS13	5.9	6.3	5.4	5.7
AS14	4.5	4.6	3.8	4.1
AS15	3.5	4.0	3.0	3.3
AS16	5.0	6.4	4.7	5.3
AS17	5.0	5.2	4.1	4.5
AS18	9.0	9.2	6.8	7.6
AS19	2.8	3.3	2.5	2.8
AS20	7.0	8.2	6.7	7.3
AS21	4.2	5.2	4.2	4.5
AS22	3.6	5.1	4.0	4.4
AS23	5.4	7.4	5.3	6.0
AS24	6.9	6.5	5.9	6.1
Total: 24	110.4	121.3	97.0	105.4

A3.2 Umvoti

Table A3.3: Field areas – Canema estate.

Field number	SASA areas (ha)	CAD areas (ha)	GIS areas (ha)	New GIS areas (ha)
DS01	24.6	25.0	20.9	23.4
DS02	17.8	18.6	15.9	17.1
DS03	7.9	9.6	7.9	8.7
DS04	9.9	10.2	8.9	12.5
DS05	6.1	6.2	5.4	5.7
DS06	9.0	9.2	8.6	8.8
DS07	9.6	11.8	8.4	10.3
DS08	7.3	6.8	4.4	5.7
DS09	6.0	5.7	4.5	5.1
DS10	5.4	5.4	4.5	5.0
DS11	7.8	8.1	6.7	7.4
DS12	6.7	6.6	5.3	6.0
DS13	9.7	9.2	7.9	8.6
DS14	19.2	20.7	16.2	18.8
DS15	8.9	9.2	7.4	8.2
DS16	7.5	8.0	6.3	7.2
DS17	18.5	19.4	15.6	17.7
DS18	8.3	8.4	6.6	7.5
DS19	7.9	8.2	6.4	7.2
DS20	9.6	9.9	8.5	9.1
DS21	20.5	20.9	17.6	19.5
DS22	9.7	9.8	7.5	8.7
DS23	10.4	8.0	6.9	7.4
DS24	19.6	20.2	18.7	19.4
DS25	18.0	21.1	18.2	19.8
DS26	13.5	13.8	11.4	12.7
DS27	33.9	35.2	27.5	31.5
DS28	20.9	21.6	17.0	19.6
DS29	14.5	13.5	12.5	12.8
DS30	11.1	11.5	9.9	10.6
DS31	7.9	8.4	6.9	7.5
DS32	14.1	14.6	12.3	13.4
DS33	6.5	6.7	5.3	6.1
ES01	4.7	4.5	3.7	4.0
Total: 34	413.0	426.0	351.7	393.0

A3.3 Richmond

Table A3.4: Field areas – Uplands estate.

Field number	SASA areas (ha)	CAD areas (ha)	GIS areas (ha)	New GIS areas (ha)
AS01	3.2	3.3	2.9	3.1
AS02	8.5	9.1	8.3	8.5
AS03	5.3	5.9	5.4	5.5
AS04	4.9	5.3	5.1	5.0
AS05	3.7	4.4	4.0	4.0
AS06	10.5	10.6	10.1	10.3
AS07	5.7	6.1	5.5	5.7
AS08	5.3	5.4	4.7	4.9
AS09	6.6	7.6	7.2	7.1
AS10	5.4	6.1	5.5	5.7
AS11	4.7	5.6	5.3	5.2
AS12	1.9	2.3	1.9	2.0
AS13	3.7	4.3	4.1	4.0
AS14	7.8	8.4	7.9	8.0
AS15	0.8	0.9	0.6	0.8
AS16	1.7	2.1	1.8	1.9
AS17	2.1	2.5	2.0	2.2
AS18	5.6	6.4	5.6	5.9
AS19	7.3	8.4	7.8	7.9
AS20	7.4	7.7	7.1	7.3
BS01	7.3	8.6	7.6	8.1
BS02	8.5	9.6	8.5	9.0
BS03	4.3	4.3	4.0	4.1
BS04	6.4	6.4	5.6	5.9
BS05	6.7	7.2	6.4	6.7
BS06	1.4	1.4	1.1	1.2
BS07	2.2	2.2	1.7	1.9
BS08	4.1	4.6	4.0	4.1
BS09	2.5	2.7	2.6	2.5
BS10	3.8	4.1	4.0	3.9
BS11	4.6	4.8	2.5	4.4
BS12	7.1	7.1	6.1	6.5
BS13	6.0	6.4	5.6	5.9
BS14	7.0	7.0	6.3	6.5
BS15	6.0	6.4	5.4	5.8
BS16	9.9	11.1	9.6	10.1
BS17	8.4	9.3	8.3	8.7
BS18	5.1	6.1	5.1	5.6
BS19	5.3	6.3	5.2	5.7
BS20	3.9	4.5	3.6	4.0
CS01	9.2	10.1	8.9	9.5
CS02	13.5	11.5	10.2	10.7
CS03	9.2	9.9	8.7	9.4
CS04	15.6	17.4	15.2	16.2
CS05	11.3	12.4	9.9	11.4
CS06	9.1	10.7	8.9	9.6
Total: 46	280.5	304.5	267.8	282.4

Table A3.5: Field areas – Greenhill estate.

Field number	SASA areas (ha)	CAD areas (ha)	GIS areas (ha)	New GIS areas (ha)
ES01	8.5	9.2	8.4	8.8
ES02	7.5	8.1	7.2	7.7
ES03	6.0	7.6	6.7	7.2
ES04	4.1	4.8	4.1	4.5
ES05	7.3	8.5	7.3	8.0
ES06	6.2	7.6	6.5	7.1
ES07	5.1	7.8	6.6	7.2
ES08	6.7	7.3	6.4	6.9
ES09	3.4	4.3	3.7	4.0
ES10	5.1	5.1	4.1	4.6
ES11	4.7	4.9	4.4	4.6
Total: 11	64.6	75.2	65.4	70.6

APPENDIX 4

COMPARISON OF MEAN RELATIVE SUCROSE AND WEIGHTED MEAN RELATIVE SUCROSE YIELDS

LIST OF TABLES

Table A4.1: Mean relative sucrose content and weighted mean relative sucrose content for sugarcane harvested from fields on Canema at Umvoti during the 1997 harvest season.

Field number	Cane yield (t cane field^{-1})	Sucrose yield (t sucrose field^{-1})		Relative sucrose content (%)	
		weighted mean	numeric mean	weighted mean	numeric mean
DS03	1026	134.29	134.41	13.1	13.1
DS05	294	39.06	39.10	13.3	13.3
DS06	772	102.76	101.90	13.3	13.2
DS07	779	105.06	105.17	13.5	13.5
DS08	789	104.96	97.84	13.3	12.4
DS09	367	48.65	48.44	13.3	13.2
DS10	427	56.32	56.36	13.2	13.2
DS11	595	77.87	77.95	13.1	13.1
DS12	341	41.10	40.92	12.1	12.0
DS14	275	34.15	35.48	12.4	12.9
DS15	214	27.33	27.29	12.8	12.8
DS16	689	93.72	93.70	13.6	13.6
DS17	1785	231.31	232.05	13.0	13.0
DS30	734	94.44	93.95	12.9	12.8
ES01	374	51.98	50.86	13.9	13.6
Total: 15	**9461**	**1243.00**	**1235.42**	**13.1**	**13.1**
DS01 †	1982	224.26	204.15	11.3	10.3

† The large difference between the weighted mean and numeric mean relative sucrose content does not correspond to the trend observed for all the other fields harvested during 1997 on Canema. Managers reported problems calculating the weighted mean sucrose yield for this field and therefore the observation is presented separately.

APPENDIX 5

CROP CYCLES FOR MONDI'S MIDLANDS SUGARCANE ENTERPRISE

LIST OF TABLES

Page

Page

Table A5.1: Crop cycles for nine to 40 month old summer-ratoon-crop sugarcane classified according to ratoon month (cycle start) and harvest age.

Harvest Age	Ratoon (start) month						
	January	February	March	September	October	November	December
9	RS.S*1S	RS.S*1S	RS.S*1S	RS.W*1S	RS.W*1S	RS.W*1S	RS.W*1S
10	RS.S*1S	RS.S*1S	RS.S*1S	RS.W*1S	RS.W*1S	RS.W*1S	RS.S*1S
11	RS.S*1S	RS.S*1S	RS.S*1S	RS.W*1S	RS.W*1S	RS.S*1S	RS.S*1S
12	RS.S*1S	RS.S*1S	RS.S*1S	RS.W*1S	RS.S*1S	RS.S*1S	RS.S*1S
13	RS.S*1S	RS.S*1S	RS.S*1S	RS.S*1S	RS.S*1S	RS.S*1S	RS.S*1S
14	RS.S*1S	RS.S*1S	RS.W*1S	RS.S*1S	RS.S*1S	RS.S*1S	RS.S*1S
15	RS.S*1S	RS.W*1S	RS.W*1S	RS.S*1S	RS.S*1S	RS.S*1S	RS.S*1S
16	RS.W*1S	RS.W*1S	RS.W*1S	RS.S*2S	RS.S*2S	RS.S*2S	RS.S*2S
17	RS.W*1S	RS.W*1S	RS.W*1S	RS.S*2S	RS.S*2S	RS.S*2S	RS.W*2S
18	RS.W*1S	RS.W*1S	RS.W*1S	RS.S*2S	RS.S*2S	RS.W*2S	RS.W*2S
19	RS.W*1S	RS.W*1S	RS.S*1S	RS.S*2S	RS.W*2S	RS.W*2S	RS.W*2S
20	RS.W*1S	RS.S*1S	RS.S*1S	RS.W*2S	RS.W*2S	RS.W*2S	RS.W*2S
21	RS.S*2S	RS.S*2S	RS.S*2S	RS.W*2S	RS.W*2S	RS.W*2S	RS.W*2S
22	RS.S*2S	RS.S*2S	RS.S*2S	RS.W*2S	RS.W*2S	RS.W*2S	RS.S*2S
23	RS.S*2S	RS.S*2S	RS.S*2S	RS.W*2S	RS.W*2S	RS.S*2S	RS.S*2S
24	RS.S*2S	RS.S*2S	RS.S*2S	RS.W*2S	RS.S*2S	RS.S*2S	RS.S*2S
25	RS.S*2S	RS.S*2S	RS.S*2S	RS.S*2S	RS.S*2S	RS.S*2S	RS.S*2S
26	RS.S*2S	RS.S*2S	RS.W*2S	RS.S*2S	RS.S*2S	RS.S*2S	RS.S*2S
27	RS.S*2S	RS.W*2S	RS.W*2S	RS.S*2S	RS.S*2S	RS.S*2S	RS.S*2S
28	RS.W*2S	RS.W*2S	RS.W*2S	RS.S*3S	RS.S*3S	RS.S*3S	RS.S*3S
29	RS.W*2S	RS.W*2S	RS.W*2S	RS.S*3S	RS.S*3S	RS.S*3S	RS.W*3S
30	RS.W*2S	RS.W*2S	RS.W*2S	RS.S*3S	RS.S*3S	RS.W*3S	RS.W*3S
31	RS.W*2S	RS.W*2S	RS.S*2S	RS.S*3S	RS.W*3S	RS.W*3S	RS.W*3S
32	RS.W*2S	RS.S*2S	RS.S*2S	RS.W*3S	RS.W*3S	RS.W*3S	RS.W*3S
33	RS.S*3S	RS.S*3S	RS.S*3S	RS.W*3S	RS.W*3S	RS.W*3S	RS.W*3S
34	RS.S*3S	RS.S*3S	RS.S*3S	RS.W*3S	RS.W*3S	RS.W*3S	RS.S*3S
35	RS.S*3S	RS.S*3S	RS.S*3S	RS.W*3S	RS.W*3S	RS.S*3S	RS.S*3S
36	RS.S*3S	RS.S*3S	RS.S*3S	RS.W*3S	RS.S*3S	RS.S*3S	RS.S*3S
37	RS.S*3S	RS.S*3S	RS.S*3S	RS.S*3S	RS.S*3S	RS.S*3S	RS.S*3S
38	RS.S*3S	RS.S*3S	RS.W*3S	RS.S*3S	RS.S*3S	RS.S*3S	RS.S*3S
39	RS.S*3S	RS.W*3S	RS.W*3S	RS.S*3S	RS.S*3S	RS.S*3S	RS.S*3S
40	RS.W*3S	RS.W*3S	RS.W*3S	RS.S*4S	RS.S*4S	RS.S*4S	RS.S*4S

Table A5.2: Crop cycles for nine to 40 month old winter-ratoon-crop sugarcane classified according to ratoon month (cycle start) and harvest age.

Harvest Age	Ratoon (start) month				
	April	May	June	July	August
9	RW.S*1S	RW.S*1S	RW.S*1S	RW.S*1S	RW.W*1S
10	RW.S*1S	RW.S*1S	RW.S*1S	RW.W*1S	RW.W*1S
11	RW.S*1S	RW.S*1S	RW.W*1S	RW.W*1S	RW.W*1S
12	RW.S*1S	RW.W*1S	RW.W*1S	RW.W*1S	RW.W*1S
13	RW.W*1S	RW.W*1S	RW.W*1S	RW.W*1S	RW.W*1S
14	RW.W*1S	RW.W*1S	RW.W*1S	RW.W*1S	RW.S*1S
15	RW.W*1S	RW.W*1S	RW.W*1S	RW.S*1S	RW.S*1S
16	RW.W*1S	RW.W*1S	RW.S*1S	RW.S*1S	RW.S*1S
17	RW.W*1S	RW.S*1S	RW.S*1S	RW.S*1S	RW.S*2S
18	RW.S*1S	RW.S*1S	RW.S*1S	RW.S*2S	RW.S*2S
19	RW.S*1S	RW.S*1S	RW.S*2S	RW.S*2S	RW.S*2S
20	RW.S*1S	RW.S*2S	RW.S*2S	RW.S*2S	RW.S*2S
21	RW.S*2S	RW.S*2S	RW.S*2S	RW.S*2S	RW.W*2S
22	RW.S*2S	RW.S*2S	RW.S*2S	RW.W*2S	RW.W*2S
23	RW.S*2S	RW.S*2S	RW.W*2S	RW.W*2S	RW.W*2S
24	RW.S*2S	RW.W*2S	RW.W*2S	RW.W*2S	RW.W*2S
25	RW.W*2S	RW.W*2S	RW.W*2S	RW.W*2S	RW.W*2S
26	RW.W*2S	RW.W*2S	RW.W*2S	RW.W*2S	RW.S*2S
27	RW.W*2S	RW.W*2S	RW.W*2S	RW.S*2S	RW.S*2S
28	RW.W*2S	RW.W*2S	RW.S*2S	RW.S*2S	RW.S*2S
29	RW.W*2S	RW.S*2S	RW.S*2S	RW.S*2S	RW.S*3S
30	RW.S*2S	RW.S*2S	RW.S*2S	RW.S*3S	RW.S*3S
31	RW.S*2S	RW.S*2S	RW.S*3S	RW.S*3S	RW.S*3S
32	RW.S*2S	RW.S*3S	RW.S*3S	RW.S*3S	RW.S*3S
33	RW.S*3S	RW.S*3S	RW.S*3S	RW.S*3S	RW.W*3S
34	RW.S*3S	RW.S*3S	RW.S*3S	RW.W*3S	RW.W*3S
35	RW.S*3S	RW.S*3S	RW.W*3S	RW.W*3S	RW.W*3S
36	RW.S*3S	RW.W*3S	RW.W*3S	RW.W*3S	RW.W*3S
37	RW.W*3S	RW.W*3S	RW.W*3S	RW.W*3S	RW.W*3S
38	RW.W*3S	RW.W*3S	RW.W*3S	RW.W*3S	RW.S*3S
39	RW.W*3S	RW.W*3S	RW.W*3S	RW.S*3S	RW.S*3S
40	RW.W*3S	RW.W*3S	RW.S*3S	RW.S*3S	RW.S*3S

Table A5.3: Crop cycles for nine to 40 month old plant-crop sugarcane classified according to planting month (cycle start) and harvest age.

Harvest Age	Plant (start) months							
	January	February	March	April	September	October	November	December
9	Check	Check	Check	PA.S*1S	PS.W*1S	PS.W*1S	PS.W*1S	PS.W*1S
10	Check	Check	PA.S*1S	PA.S*1S	PS.W*1S	PS.W*1S	PS.W*1S	PS.S*1S
11	Check	PA.S*1S	PA.S*1S	PA.S*1S	PS.W*1S	PS.W*1S	PS.S*1S	PS.S*1S
12	PA.S*1S	PA.S*1S	PA.S*1S	PA.S*1S	PS.W*1S	PS.S*1S	PS.S*1S	PS.S*1S
13	PA.S*1S	PA.S*1S	PA.S*1S	PA.W*1S	PS.S*1S	PS.S*1S	PS.S*1S	PS.S*1S
14	PA.S*1S	PA.S*1S	PA.W*1S	PA.W*1S	PS.S*1S	PS.S*1S	PS.S*1S	PS.S*1S
15	PA.S*1S	PA.W*1S	PA.W*1S	PA.W*1S	PS.S*1S	PS.S*1S	PS.S*1S	PS.S*1S
16	PA.W*1S	PA.W*1S	PA.W*1S	PA.W*1S	PS.S*2S	PS.S*2S	PS.S*2S	PS.S*2S
17	PA.W*1S	PA.W*1S	PA.W*1S	PA.W*1S	PS.S*2S	PS.S*2S	PS.S*2S	PS.W*2S
18	PA.W*1S	PA.W*1S	PA.W*1S	PA.S*1S	PS.S*2S	PS.S*2S	PS.W*2S	PS.W*2S
19	PA.W*1S	PA.W*1S	PA.S*1S	PA.S*1S	PS.S*2S	PS.W*2S	PS.W*2S	PS.W*2S
20	PA.W*1S	PA.S*1S	PA.S*1S	PA.S*1S	PS.W*2S	PS.W*2S	PS.W*2S	PS.W*2S
21	PA.S*1S	PA.S*1S	PA.S*1S	PA.S*2S	PS.W*2S	PS.W*2S	PS.W*2S	PS.W*2S
22	PA.S*1S	PA.S*1S	PA.S*2S	PA.S*2S	PS.W*2S	PS.W*2S	PS.W*2S	PS.S*2S
23	PA.S*1S	PA.S*2S	PA.S*2S	PA.S*2S	PS.W*2S	PS.W*2S	PS.S*2S	PS.S*2S
24	PA.S*2S	PA.S*2S	PA.S*2S	PA.S*2S	PS.W*2S	PS.S*2S	PS.S*2S	PS.S*2S
25	PA.S*2S	PA.S*2S	PA.S*2S	PA.W*2S	PS.S*2S	PS.S*2S	PS.S*2S	PS.S*2S
26	PA.S*2S	PA.S*2S	PA.W*2S	PA.W*2S	PS.S*2S	PS.S*2S	PS.S*2S	PS.S*2S
27	PA.S*2S	PA.W*2S	PA.W*2S	PA.W*2S	PS.S*2S	PS.S*2S	PS.S*2S	PS.S*2S
28	PA.W*2S	PA.W*2S	PA.W*2S	PA.W*2S	PS.S*3S	PS.S*3S	PS.S*3S	PS.S*3S
29	PA.W*2S	PA.W*2S	PA.W*2S	PA.W*2S	PS.S*3S	PS.S*3S	PS.S*3S	PS.W*3S
30	PA.W*2S	PA.W*2S	PA.W*2S	PA.S*2S	PS.S*3S	PS.S*3S	PS.W*3S	PS.W*3S
31	PA.W*2S	PA.W*2S	PA.S*2S	PA.S*2S	PS.S*3S	PS.W*3S	PS.W*3S	PS.W*3S
32	PA.W*2S	PA.S*2S	PA.S*2S	PA.S*2S	PS.W*3S	PS.W*3S	PS.W*3S	PS.W*3S
33	PA.S*2S	PA.S*2S	PA.S*2S	PA.S*3S	PS.W*3S	PS.W*3S	PS.W*3S	PS.W*3S
34	PA.S*2S	PA.S*2S	PA.S*3S	PA.S*3S	PS.W*3S	PS.W*3S	PS.W*3S	PS.S*3S
35	PA.S*2S	PA.S*3S	PA.S*3S	PA.S*3S	PS.W*3S	PS.W*3S	PS.S*3S	PS.S*3S
36	PA.S*3S	PA.S*3S	PA.S*3S	PA.S*3S	PS.W*3S	PS.S*3S	PS.S*3S	PS.S*3S
37	PA.S*3S	PA.S*3S	PA.S*3S	PA.W*3S	PS.S*3S	PS.S*3S	PS.S*3S	PS.S*3S
38	PA.S*3S	PA.S*3S	PA.W*3S	PA.W*3S	PS.S*3S	PS.S*3S	PS.S*3S	PS.S*3S
39	PA.S*3S	PA.W*3S	PA.W*3S	PA.W*3S	PS.S*3S	PS.S*3S	PS.S*3S	PS.S*3S
40	PA.W*3S	PA.W*3S	PA.W*3S	PA.W*3S	PS.S*4S	PS.S*4S	PS.S*4S	PS.S*4S

Table A5.4: Summer months for nine to 40 month old sugarcane crop cycles classified according to ratoon / plant month (cycle start) and harvest age.

Harvest Age	Ratoon Month												Plant Month							
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	Sep	Oct	Nov	Dec
9	4	4	4	4	5	6	7	7	7	6	5	4	Check	Check	Check	4	7	6	5	4
10	5	5	5	5	6	7	7	7	7	6	5	5	Check	Check	4	5	7	6	5	5
11	6	6	6	6	7	7	7	7	7	6	6	6	Check	4	5	6	7	6	6	6
12	7	7	7	7	7	7	7	7	7	7	7	7	4	5	6	7	7	7	7	7
13	8	8	8	7	7	7	7	7	8	8	8	8	5	6	7	7	8	8	8	8
14	9	9	8	7	7	7	7	8	9	9	9	9	6	7	7	7	9	9	9	9
15	10	9	8	7	7	7	8	9	10	10	10	10	7	7	7	7	10	10	10	10
16	10	9	8	7	7	8	9	10	11	11	11	11	7	7	7	7	11	11	11	11
17	10	9	8	7	8	9	10	11	12	12	12	11	7	7	7	7	12	12	12	11
18	10	9	8	8	9	10	11	12	13	13	12	11	7	7	7	8	13	13	12	11
19	10	9	9	9	10	11	12	13	14	13	12	11	7	7	8	9	14	13	12	11
20	10	10	10	10	11	12	13	14	14	13	12	11	7	8	9	10	14	13	12	11
21	11	11	11	11	12	13	14	14	14	13	12	11	8	9	10	11	14	13	12	11
22	12	12	12	12	13	14	14	14	14	13	12	12	9	10	11	12	14	13	12	12
23	13	13	13	13	14	14	14	14	14	13	13	13	10	11	12	13	14	13	13	13
24	14	14	14	14	14	14	14	14	14	14	14	14	11	12	13	14	14	14	14	14
25	15	15	15	14	14	14	14	14	15	15	15	15	12	13	14	14	15	15	15	15
26	16	16	15	14	14	14	14	15	16	16	16	16	13	14	14	14	16	16	16	16
27	17	16	15	14	14	14	15	16	17	17	17	17	14	14	14	14	17	17	17	17
28	17	16	15	14	14	15	16	17	18	18	18	18	14	14	14	14	18	18	18	18
29	17	16	15	14	15	16	17	18	19	19	19	18	14	14	14	14	19	19	19	18
30	17	16	15	15	16	17	18	19	20	20	19	18	14	14	14	15	20	20	19	18
31	17	16	16	16	17	18	19	20	21	20	19	18	14	14	15	16	21	20	19	18
32	17	17	17	17	18	19	20	21	21	20	19	18	14	15	16	17	21	20	19	18
33	18	18	18	18	19	20	21	21	21	20	19	18	15	16	17	18	21	20	19	18
34	19	19	19	19	20	21	21	21	21	20	19	19	16	17	18	19	21	20	19	19
35	20	20	20	20	21	21	21	21	21	20	20	20	17	18	19	20	21	20	20	20
36	21	21	21	21	21	21	21	21	21	21	21	21	18	19	20	21	21	21	21	21
37	22	22	22	21	21	21	21	21	22	22	22	22	19	20	21	21	22	22	22	22
38	23	23	22	21	21	21	21	22	23	23	23	23	20	21	21	21	23	23	23	23
39	24	23	22	21	21	21	22	23	24	24	24	24	21	21	21	21	24	24	24	24
40	24	23	22	21	21	22	23	24	25	25	25	25	21	21	21	21	25	25	25	25

Check: These crop cycles have experienced less than four summer months' growth.

Table A5.5: Winter months for nine to 40 month old sugarcane crop cycles classified according to ratoon / plant month (cycle start) and harvest age.

Harvest Age	Ratoon Month												Plant Month							
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	Sep	Oct	Nov	Dec
9	5	5	5	5	4	3	2	2	2	3	4	5	Check	Check	Check	5	2	3	4	5
10	5	5	5	5	4	3	3	3	3	4	5	5	Check	Check	6	5	3	4	5	5
11	5	5	5	5	4	4	4	4	4	5	5	5	Check	7	6	5	4	5	5	5
12	5	5	5	5	5	5	5	5	5	5	5	5	8	7	6	5	5	5	5	5
13	5	5	5	6	6	6	6	6	5	5	5	5	8	7	6	6	5	5	5	5
14	5	5	6	7	7	7	7	6	5	5	5	5	8	7	7	7	5	5	5	5
15	5	6	7	8	8	8	7	6	5	5	5	5	8	8	8	8	5	5	5	5
16	6	7	8	9	9	8	7	6	5	5	5	5	9	9	9	9	5	5	5	5
17	7	8	9	10	9	8	7	6	5	5	5	6	10	10	10	10	5	5	5	6
18	8	9	10	10	9	8	7	6	5	5	6	7	11	11	11	10	5	5	6	7
19	9	10	10	10	9	8	7	6	5	6	7	8	12	12	11	10	5	6	7	8
20	10	10	10	10	9	8	7	6	6	7	8	9	13	12	11	10	6	7	8	9
21	10	10	10	10	9	8	7	7	7	8	9	10	13	12	11	10	7	8	9	10
22	10	10	10	10	9	8	8	8	8	9	10	10	13	12	11	10	8	9	10	10
23	10	10	10	10	9	9	9	9	9	10	10	10	13	12	11	10	9	10	10	10
24	10	10	10	10	10	10	10	10	10	10	10	10	13	12	11	10	10	10	10	10
25	10	10	10	11	11	11	11	11	10	10	10	10	13	12	11	11	10	10	10	10
26	10	10	11	12	12	12	12	11	10	10	10	10	13	12	12	12	10	10	10	10
27	10	11	12	13	13	13	12	11	10	10	10	10	13	13	13	13	10	10	10	10
28	11	12	13	14	14	13	12	11	10	10	10	10	14	14	14	14	10	10	10	10
29	12	13	14	15	14	13	12	11	10	10	10	11	15	15	15	15	10	10	10	11
30	13	14	15	15	14	13	12	11	10	10	11	12	16	16	16	15	10	10	11	12
31	14	15	15	15	14	13	12	11	10	11	12	13	17	17	16	15	10	11	12	13
32	15	15	15	15	14	13	12	11	11	12	13	14	18	17	16	15	11	12	13	14
33	15	15	15	15	14	13	12	12	12	13	14	15	18	17	16	15	12	13	14	15
34	15	15	15	15	14	13	13	13	13	14	15	15	18	17	16	15	13	14	15	15
35	15	15	15	15	14	14	14	14	14	15	15	15	18	17	16	15	14	15	15	15
36	15	15	15	15	15	15	15	15	15	15	15	15	18	17	16	15	15	15	15	15
37	15	15	15	16	16	16	16	16	15	15	15	15	18	17	16	16	15	15	15	15
38	15	15	16	17	17	17	17	16	15	15	15	15	18	17	17	17	15	15	15	15
39	15	16	17	18	18	18	17	16	15	15	15	15	18	18	18	18	15	15	15	15
40	16	17	18	19	19	18	17	16	15	15	15	15	19	19	19	19	15	15	15	15

Check: These crop cycles have experienced less than four summer months' growth.

Table A5.6: Crop cycle schedule – RS.S*1S.

Ratoon Month	Ages	Summer Months		Harvest Details						
September	13-15	8-10	Harvest Month	Jan	Feb	Mar	Apr	Oct	Nov	Dec
			Age	4	5	6	7	13	14	15
			Summer Months	4	5	6	7	8	9	10
October	12-15	7-10	Harvest Month	Feb	Mar	Apr	Oct	Nov	Dec	Jan
			Age	4	5	6	12	13	14	15
			Summer Months	4	5	6	7	8	9	10
November	11-15	6-10	Harvest Month	Mar	Apr	Oct	Nov	Dec	Jan	Feb
			Age	4	5	11	12	13	14	15
			Summer Months	4	5	6	7	8	9	10
December	10-15	5-10	Harvest Month	Apr	Oct	Nov	Dec	Jan	Feb	Mar
			Age	4	10	11	12	13	14	15
			Summer Months	4	5	6	7	8	9	10
January	9-15	4-10	Harvest Month	Oct	Nov	Dec	Jan	Feb	Mar	Apr
			Age	9	10	11	12	13	14	15
			Summer Months	4	5	6	7	8	9	10
February	9-14; 20	4-10	Harvest Month	Nov	Dec	Jan	Feb	Mar	Apr	Oct
			Age	9	10	11	12	13	14	20
			Summer Months	4	5	6	7	8	9	10
March	9-13; 19-20	4-10	Harvest Month	Dec	Jan	Fab	Mar	Apr	Oct	Nov
			Age	9	10	11	12	13	19	20
			Summer Months	4	5	6	7	8	9	10

Table A5.7: Crop cycle schedule – RS.S*2S.

Ratoon Month	Ages	Summer Months		Harvest Details						
September	16-19; 25-27	11-17	Harvest Month	Jan	Feb	Mar	Apr	Oct	Nov	Dec
			Age	16	17	18	19	25	26	27
			Summer Months	11	12	13	14	15	16	17
October	16-18; 24-27	11-17	Harvest Month	Feb	Mar	Apr	Oct	Nov	Dec	Jan
			Age	16	17	18	24	25	26	27
			Summer Months	11	12	13	14	15	16	17
November	16-17; 23-27	11-17	Harvest Month	Mar	Apr	Oct	Nov	Dec	Jan	Feb
			Age	16	17	23	24	25	26	27
			Summer Months	11	12	13	14	15	16	17
December	16; 22-27	11-17	Harvest Month	Apr	Oct	Nov	Dec	Jan	Feb	Mar
			Age	16	22	23	24	25	26	27
			Summer Months	11	12	13	14	15	16	17
January	21-27	11-17	Harvest Month	Oct	Nov	Dec	Jan	Feb	Mar	Apr
			Age	21	22	23	24	25	26	27
			Summer Months	11	12	13	14	15	16	17
February	21-26; 32	11-17	Harvest Month	Nov	Dec	Jan	Feb	Mar	Apr	Oct
			Age	21	22	23	24	25	26	32
			Summer Months	11	12	13	14	15	16	17
March	21-25; 31-32	11-17	Harvest Month	Dec	Jan	Feb	Mar	Apr	Oct	Nov
			Age	21	22	23	24	25	31	32
			Summer Months	11	12	13	14	15	16	17

Table A5.8: Crop cycle schedule – RS.S*3S.

Ratoon Month	Ages	Summer Months		Harvest Details						
September	28-31; 37-39	18-24	Harvest Month	Jan	Feb	Mar	Apr	Oct	Nov	Dec
			Age	28	29	30	31	37	38	39
			Summer Months	18	19	20	21	22	23	24
October	28-30; 36-39	18-24	Harvest Month	Feb	Mar	Apr	Oct	Nov	Dec	Jan
			Age	28	29	30	36	37	38	39
			Summer Months	18	19	20	21	22	23	24
November	28-29; 35-39	18-24	Harvest Month	Mar	Apr	Oct	Nov	Dec	Jan	Feb
			Age	28	29	35	36	37	38	39
			Summer Months	18	19	20	21	22	23	24
December	28; 34-39	18-24	Harvest Month	Apr	Oct	Nov	Dec	Jan	Feb	Mar
			Age	28	34	35	36	37	38	39
			Summer Months	18	19	20	21	22	23	24
January	33-39	18-24	Harvest Month	Oct	Nov	Dec	Jan	Feb	Mar	Apr
			Age	33	34	35	36	37	38	39
			Summer Months	18	19	20	21	22	23	24
February	33-38; 44	18-24	Harvest Month	Nov	Dec	Jan	Feb	Mar	Apr	Oct
			Age	33	34	35	36	37	38	44
			Summer Months	18	19	20	21	22	23	24
March	33-37; 43-44	18-24	Harvest Month	Dec	Jan	Feb	Mar	Apr	Oct	Nov
			Age	33	34	35	36	37	43	44
			Summer Months	18	19	20	21	22	23	24

Table A5.9: Crop cycle schedule – RS.W*1S.

Ratoon Month	Ages	Summer Months		Harvest Details				
September	9-12	7	Harvest Month	May	Jun	Jul	Aug	Sep
			Age	8	9	10	11	12
			Summer Months	7	7	7	7	7
October	9-11	6	Harvest Month	May	Jun	Jul	Aug	Sep
			Age	7	8	9	10	11
			Summer Months	6	6	6	6	6
November	9-10	5	Harvest Month	May	Jun	Jul	Aug	Sep
			Age	6	7	8	9	10
			Summer Months	5	5	5	5	5
December	9	4	Harvest Month	May	Jun	Jul	Aug	Sep
			Age	5	6	7	8	9
			Summer Months	4	4	4	4	4
January	16-20	10	Harvest Month	May	Jun	Jul	Aug	Sep
			Age	16	17	18	19	20
			Summer Months	10	10	10	10	10
February	15-19	9	Harvest Month	May	Jun	Jul	Aug	Sep
			Age	15	16	17	18	19
			Summer Months	9	9	9	9	9
March	14-18	8	Harvest Month	May	Jun	Jul	Aug	Sep
			Age	14	15	16	17	18
			Summer Months	8	8	8	8	8

Table A5.10: Crop cycle schedule – RS.W*2S.

Ratoon Month	Ages	Summer Months		Harvest Details				
September	20-24	14	Harvest Month	May	Jun	Jul	Aug	Sep
			Age	20	21	22	23	24
			Summer Months	14	14	14	14	14
October	19-23	13	Harvest Month	May	Jun	Jul	Aug	Sep
			Age	19	20	21	22	23
			Summer Months	13	13	13	13	13
November	18-22	12	Harvest Month	May	Jun	Jul	Aug	Sep
			Age	18	19	20	21	22
			Summer Months	12	12	12	12	12
December	17-21	11	Harvest Month	May	Jun	Jul	Aug	Sep
			Age	17	18	19	20	21
			Summer Months	11	11	11	11	11
January	28-32	17	Harvest Month	May	Jun	Jul	Aug	Sep
			Age	28	29	30	31	32
			Summer Months	17	17	17	17	17
February	27-31	16	Harvest Month	May	Jun	Jul	Aug	Sep
			Age	27	28	29	30	31
			Summer Months	16	16	16	16	16
March	26-30	15	Harvest Month	May	Jun	Jul	Aug	Sep
			Age	26	27	28	29	30
			Summer Months	15	15	15	15	15

Table A5.11: Crop cycle schedule – RS.W*3S.

Ratoon Month	Ages	Summer Months		Harvest Details				
September	32-36	21	Harvest Month	May	Jun	Jul	Aug	Sep
			Age	32	33	34	35	36
			Summer Months	21	21	21	21	21
October	31-35	20	Harvest Month	May	Jun	Jul	Aug	Sep
			Age	31	32	33	34	35
			Summer Months	20	20	20	20	20
November	30-34	19	Harvest Month	May	Jun	Jul	Aug	Sep
			Age	30	31	32	33	34
			Summer Months	19	19	19	19	19
December	29-33	18	Harvest Month	May	Jun	Jul	Aug	Sep
			Age	29	30	31	32	33
			Summer Months	18	18	18	18	18
January	40-44	24	Harvest Month	May	Jun	Jul	Aug	Sep
			Age	40	41	42	43	44
			Summer Months	24	24	24	24	24
February	39-43	23	Harvest Month	May	Jun	Jul	Aug	Sep
			Age	39	40	41	42	43
			Summer Months	23	23	23	23	23
March	38-42	22	Harvest Month	May	Jun	Jul	Aug	Sep
			Age	38	39	40	41	42
			Summer Months	22	22	22	22	22

Table A5.12: Crop cycle schedule – RW.S*1S.

Ratoon Month	Ages	Summer Months		Harvest Details						
April	9-12; 18-20	4-10	Harvest Month	Jan	Feb	Mar	Apr	Oct	Nov	Dec
			Age	9	10	11	12	18	19	20
			Summer Months	4	5	6	7	8	9	10
May	9-11; 17-19	5-10	Harvest Month	Jan	Feb	Mar	Apr	Oct	Nov	Dec
			Age	8	9	10	11	17	18	19
			Summer Months	4	5	6	7	8	9	10
June	9-10; 16-18	6-10	Harvest Month	Jan	Feb	Mar	Apr	Oct	Nov	Dec
			Age	7	8	9	10	16	17	18
			Summer Months	4	5	6	7	8	9	10
July	9; 15-17	7-10	Harvest Month	Jan	Feb	Mar	Apr	Oct	Nov	Dec
			Age	6	7	8	9	15	16	17
			Summer Months	4	5	6	7	8	9	10
August	14-16	8-10	Harvest Month	Jan	Feb	Mar	Apr	Oct	Nov	Dec
			Age	5	6	7	8	14	15	16
			Summer Months	4	5	6	7	8	9	10

Table A5.13: Crop cycle schedule – RW.S*2S.

Ratoon Month	Ages	Summer Months		Harvest Details						
April	21-24; 30-32	11-17	Harvest Month	Jan	Feb	Mar	Apr	Oct	Nov	Dec
			Age	21	22	23	24	30	31	32
			Summer Months	11	12	13	14	15	16	17
May	20-23; 29-31	11-17	Harvest Month	Jan	Feb	Mar	Apr	Oct	Nov	Dec
			Age	20	21	22	23	29	30	31
			Summer Months	11	12	13	14	15	16	17
June	19-22; 28-30	11-17	Harvest Month	Jan	Feb	Mar	Apr	Oct	Nov	Dec
			Age	19	20	21	22	28	29	30
			Summer Months	11	12	13	14	15	16	17
July	18-21; 27-29	11-17	Harvest Month	Jan	Feb	Mar	Apr	Oct	Nov	Dec
			Age	18	19	20	21	27	28	29
			Summer Months	11	12	13	14	15	16	17
August	17-20; 26-28	11-17	Harvest Month	Jan	Feb	Mar	Apr	Oct	Nov	Dec
			Age	17	18	19	20	26	27	28
			Summer Months	11	12	13	14	15	16	17

Table A5.14: Crop cycle schedule – RW.S*3S.

Ratoon Month	Ages	Summer Months		Harvest Details						
April	33-36; 42-44	18-24	Harvest Month	Jan	Feb	Mar	Apr	Oct	Nov	Dec
			Age	33	34	35	36	42	43	44
			Summer Months	18	19	20	21	22	23	24
May	32-35; 41-43	18-24	Harvest Month	Jan	Feb	Mar	Apr	Oct	Nov	Dec
			Age	32	33	34	35	41	42	43
			Summer Months	18	19	20	21	22	23	24
June	31-34; 40-42	18-24	Harvest Month	Jan	Feb	Mar	Apr	Oct	Nov	Dec
			Age	31	32	33	34	40	41	42
			Summer Months	18	19	20	21	22	23	24
July	30-33; 39-41	18-24	Harvest Month	Jan	Feb	Mar	Apr	Oct	Nov	Dec
			Age	30	31	32	33	39	40	41
			Summer Months	18	19	20	21	22	23	24
August	29-32; 38-40	18-24	Harvest Month	Jan	Feb	Mar	Apr	Oct	Nov	Dec
			Age	29	30	31	32	38	39	40
			Summer Months	18	19	20	21	22	23	24

Table A5.15: Crop cycle schedule – RW.W*1S.

Ratoon Month	Ages	Summer Months		Harvest Details				
April	13-17	7	Harvest Month	May	Jun	Jul	Aug	Sep
			Age	13	14	15	16	17
			Summer Months	7	7	7	7	7
May	12-16	7	Harvest Month	May	Jun	Jul	Aug	Sep
			Age	12	13	14	15	16
			Summer Months	7	7	7	7	7
June	11-15	7	Harvest Month	May	Jun	Jul	Aug	Sep
			Age	11	12	13	14	15
			Summer Months	7	7	7	7	7
July	10-14	7	Harvest Month	May	Jun	Jul	Aug	Sep
			Age	10	11	12	13	14
			Summer Months	7	7	7	7	7
August	9-13	7	Harvest Month	May	Jun	Jul	Aug	Sep
			Age	9	10	11	12	13
			Summer Months	7	7	7	7	7

Table A5.16: Crop cycle schedule – RW.W*2S.

Ratoon Month	Ages	Summer Months		Harvest Details				
April	25-29	14	Harvest Month	May	Jun	Jul	Aug	Sep
			Age	25	26	27	28	29
			Summer Months	14	14	14	14	14
May	24-28	14	Harvest Month	May	Jun	Jul	Aug	Sep
			Age	24	25	26	27	28
			Summer Months	14	14	14	14	14
June	23-27	14	Harvest Month	May	Jun	Jul	Aug	Sep
			Age	23	24	25	26	27
			Summer Months	14	14	14	14	14
July	22-26	14	Harvest Month	May	Jun	Jul	Aug	Sep
			Age	22	23	24	25	26
			Summer Months	14	14	14	14	14
August	21-25	14	Harvest Month	May	Jun	Jul	Aug	Sep
			Age	21	22	23	24	25
			Summer Months	14	14	14	14	14

Table A5.17: Crop cycle schedule – RW.W*3S.

Ratoon Month	Ages	Summer Months		Harvest Details				
April	37-41	21	Harvest Month	May	Jun	Jul	Aug	Sep
			Age	37	38	39	40	41
			Summer Months	21	21	21	21	21
May	36-40	21	Harvest Month	May	Jun	Jul	Aug	Sep
			Age	36	37	38	39	40
			Summer Months	21	21	21	21	21
June	35-39	21	Harvest Month	May	Jun	Jul	Aug	Sep
			Age	35	36	37	38	39
			Summer Months	21	21	21	21	21
July	34-38	21	Harvest Month	May	Jun	Jul	Aug	Sep
			Age	34	35	36	37	38
			Summer Months	21	21	21	21	21
August	33-37	21	Harvest Month	May	Jun	Jul	Aug	Sep
			Age	33	34	35	36	37
			Summer Months	21	21	21	21	21

Table A5.18: Crop cycle schedule – PA.S*1S.

Ratoon Month	Ages	Summer Months		Harvest Details						
January	12-15; 21-23	4-10	Harvest Month	Jan	Feb	Mar	Apr	Oct	Nov	Dec
			Age	12	13	14	15	21	22	23
			Summer Months	4	5	6	7	8	9	10
February	11-14; 20-22	4-10	Harvest Month	Jan	Feb	Mar	Apr	Oct	Nov	Dec
			Age	11	12	13	14	20	21	22
			Summer Months	4	5	6	7	8	9	10
March	10-13; 19-21	4-10	Harvest Month	Jan	Feb	Mar	Apr	Oct	Nov	Dec
			Age	10	11	12	13	19	20	21
			Summer Months	4	5	6	7	8	9	10
April	9-12; 18-20	4-10	Harvest Month	Jan	Feb	Mar	Apr	Oct	Nov	Dec
			Age	9	10	11	12	18	19	20
			Summer Months	4	5	6	7	8	9	10

Table A5.19: Crop cycle schedule – PA.S*2S.

Ratoon Month	Ages	Summer Months		Harvest Details						
January	24-27; 33-35	11-17	Harvest Month	Jan	Feb	Mar	Apr	Oct	Nov	Dec
			Age	24	25	26	27	33	34	35
			Summer Months	11	12	13	14	15	16	17
February	23-26; 32-34	11-17	Harvest Month	Jan	Feb	Mar	Apr	Oct	Nov	Dec
			Age	23	24	25	26	32	33	34
			Summer Months	11	12	13	14	15	16	17
March	22-25; 31-33	11-17	Harvest Month	Jan	Feb	Mar	Apr	Oct	Nov	Dec
			Age	22	23	24	25	31	32	33
			Summer Months	11	12	13	14	15	16	17
April	21-24; 30-32	11-17	Harvest Month	Jan	Feb	Mar	Apr	Oct	Nov	Dec
			Age	21	22	23	24	30	31	32
			Summer Months	11	12	13	14	15	16	17

Table A5.20: Crop cycle schedule – PA.S*3S.

Ratoon Month	Ages	Summer Months		Harvest Details						
January	36-39; 45-47	18-24	Harvest Month	Jan	Feb	Mar	Apr	Oct	Nov	Dec
			Age	36	37	38	39	45	46	47
			Summer Months	18	19	20	21	22	23	24
February	35-38; 44-46	18-24	Harvest Month	Jan	Feb	Mar	Apr	Oct	Nov	Dec
			Age	35	36	37	38	44	45	46
			Summer Months	18	19	20	21	22	23	24
March	34-37; 43-45	18-24	Harvest Month	Jan	Feb	Mar	Apr	Oct	Nov	Dec
			Age	34	35	36	37	43	44	45
			Summer Months	18	19	20	21	22	23	24
April	33-36; 42-44	18-24	Harvest Month	Jan	Feb	Mar	Apr	Oct	Nov	Dec
			Age	33	34	35	36	42	43	44
			Summer Months	18	19	20	21	22	23	24

Table A5.21: Crop cycle schedule – PA.W*1S.

Ratoon Month	Ages	Summer Months		Harvest Details				
January	16-20	7	Harvest Month	May	Jun	Jul	Aug	Sep
			Age	16	17	18	19	20
			Summer Months	7	7	7	7	7
February	15-19	7	Harvest Month	May	Jun	Jul	Aug	Sep
			Age	15	16	17	18	19
			Summer Months	7	7	7	7	7
March	14-18	7	Harvest Month	May	Jun	Jul	Aug	Sep
			Age	14	15	16	17	18
			Summer Months	7	7	7	7	7
April	13-17	7	Harvest Month	May	Jun	Jul	Aug	Sep
			Age	13	14	15	16	17
			Summer Months	7	7	7	7	7

Table A5.22: Crop cycle schedule – PA.W*2S.

Ratoon Month	Ages	Summer Months		Harvest Details				
January	28-32	14	Harvest Month	May	Jun	Jul	Aug	Sep
			Age	28	29	30	31	32
			Summer Months	14	14	14	14	14
February	27-31	14	Harvest Month	May	Jun	Jul	Aug	Sep
			Age	27	28	29	30	31
			Summer Months	14	14	14	14	14
March	26-30	14	Harvest Month	May	Jun	Jul	Aug	Sep
			Age	26	27	28	29	30
			Summer Months	14	14	14	14	14
April	25-29	14	Harvest Month	May	Jun	Jul	Aug	Sep
			Age	25	26	27	28	29
			Summer Months	14	14	14	14	14

Table A5.23: Crop cycle schedule – PA.W*3S.

Ratoon Month	Ages	Summer Months		Harvest Details				
January	40-44	21	Harvest Month	May	Jun	Jul	Aug	Sep
			Age	40	41	42	43	44
			Summer Months	21	21	21	21	21
February	39-43	21	Harvest Month	May	Jun	Jul	Aug	Sep
			Age	39	40	41	42	43
			Summer Months	21	21	21	21	21
March	38-42	21	Harvest Month	May	Jun	Jul	Aug	Sep
			Age	38	39	40	41	42
			Summer Months	21	21	21	21	21
April	37-41	21	Harvest Month	May	Jun	Jul	Aug	Sep
			Age	37	38	39	40	41
			Summer Months	21	21	21	21	21

Table A5.24: Crop cycle schedule – PS.S*1S.

Ratoon Month	Ages	Summer Months		Harvest Details						
September	13-15	8-10	Harvest Month	Jan	Feb	Mar	Apr	Oct	Nov	Dec
			Age	4	5	6	7	13	14	15
			Summer Months	4	5	6	7	8	9	10
October	12-15	7-10	Harvest Month	Feb	Mar	Apr	Oct	Nov	Dec	Jan
			Age	4	5	6	12	13	14	15
			Summer Months	4	5	6	7	8	9	10
November	11-15	6-10	Harvest Month	Mar	Apr	Oct	Nov	Dec	Jan	Feb
			Age	4	5	11	12	13	14	15
			Summer Months	4	5	6	7	8	9	10
December	10-15	5-10	Harvest Month	Apr	Oct	Nov	Dec	Jan	Feb	Mar
			Age	4	10	11	12	13	14	15
			Summer Months	4	5	6	7	8	9	10

Table A5.25: Crop cycle schedule – PS.S*2S.

Ratoon Month	Ages	Summer Months		Harvest Details						
September	16-19; 25-27	11-17	Harvest Month	Jan	Feb	Mar	Apr	Oct	Nov	Dec
			Age	16	17	18	19	25	26	27
			Summer Months	11	12	13	14	15	16	17
October	16-18; 24-27	11-17	Harvest Month	Feb	Mar	Apr	Oct	Nov	Dec	Jan
			Age	16	17	18	24	25	26	27
			Summer Months	11	12	13	14	15	16	17
November	16-17; 23-27	11-17	Harvest Month	Mar	Apr	Oct	Nov	Dec	Jan	Feb
			Age	16	17	23	24	25	26	27
			Summer Months	11	12	13	14	15	16	17
December	16; 22-27	11-17	Harvest Month	Apr	Oct	Nov	Dec	Jan	Feb	Mar
			Age	16	22	23	24	25	26	27
			Summer Months	11	12	13	14	15	16	17

Table A5.26: Crop cycle schedule – PS.S*3S.

Ratoon Month	Ages	Summer Months		Harvest Details						
September	28-31; 37-39	18-24	Harvest Month	Jan	Feb	Mar	Apr	Oct	Nov	Dec
			Age	28	29	30	31	37	38	39
			Summer Months	18	19	20	21	22	23	24
October	28-30; 36-39	18-24	Harvest Month	Feb	Mar	Apr	Oct	Nov	Dec	Jan
			Age	28	29	30	36	37	38	39
			Summer Months	18	19	20	21	22	23	24
November	28-29; 35-39	18-24	Harvest Month	Mar	Apr	Oct	Nov	Dec	Jan	Feb
			Age	28	29	35	36	37	38	39
			Summer Months	18	19	20	21	22	23	24
December	28; 34-39	18-24	Harvest Month	Apr	Oct	Nov	Dec	Jan	Feb	Mar
			Age	28	34	35	36	37	38	39
			Summer Months	18	19	20	21	22	23	24

Table A5.27: Crop cycle schedule – PS.W*1S.

Ratoon Month	Ages	Summer Months		Harvest Details				
September	9-12	7	Harvest Month	May	Jun	Jul	Aug	Sep
			Age	8	9	10	11	12
			Summer Months	7	7	7	7	7
October	9-11	6	Harvest Month	May	Jun	Jul	Aug	Sep
			Age	7	8	9	10	11
			Summer Months	6	6	6	6	6
November	9-10	5	Harvest Month	May	Jun	Jul	Aug	Sep
			Age	6	7	8	9	10
			Summer Months	5	5	5	5	5
December	9	4	Harvest Month	May	Jun	Jul	Aug	Sep
			Age	5	6	7	8	9
			Summer Months	4	4	4	4	4

Table A5.28: Crop cycle schedule – PS.W*2S.

Ratoon Month	Ages	Summer Months		Harvest Details				
September	20-24	14	Harvest Month	May	Jun	Jul	Aug	Sep
			Age	20	21	22	23	24
			Summer Months	14	14	14	14	14
October	19-23	13	Harvest Month	May	Jun	Jul	Aug	Sep
			Age	19	20	21	22	23
			Summer Months	13	13	13	13	13
November	18-22	12	Harvest Month	May	Jun	Jul	Aug	Sep
			Age	18	19	20	21	22
			Summer Months	12	12	12	12	12
December	17-21	11	Harvest Month	May	Jun	Jul	Aug	Sep
			Age	17	18	19	20	21
			Summer Months	11	11	11	11	11

Table A5.29: Crop cycle schedule – PS.W*3S.

Ratoon Month	Ages	Summer Months		Harvest Details				
September	32-36	21	Harvest Month	May	Jun	Jul	Aug	Sep
			Age	32	33	34	35	36
			Summer Months	21	21	21	21	21
October	31-35	20	Harvest Month	May	Jun	Jul	Aug	Sep
			Age	31	32	33	34	35
			Summer Months	20	20	20	20	20
November	30-34	19	Harvest Month	May	Jun	Jul	Aug	Sep
			Age	30	31	32	33	34
			Summer Months	19	19	19	19	19
December	29-33	18	Harvest Month	May	Jun	Jul	Aug	Sep
			Age	29	30	31	32	33
			Summer Months	18	18	18	18	18

APPENDIX 6

SUGARCANE YIELD (TCH) RESPONSE TO SELECTED VARIABLES

LIST OF FIGURES

Page

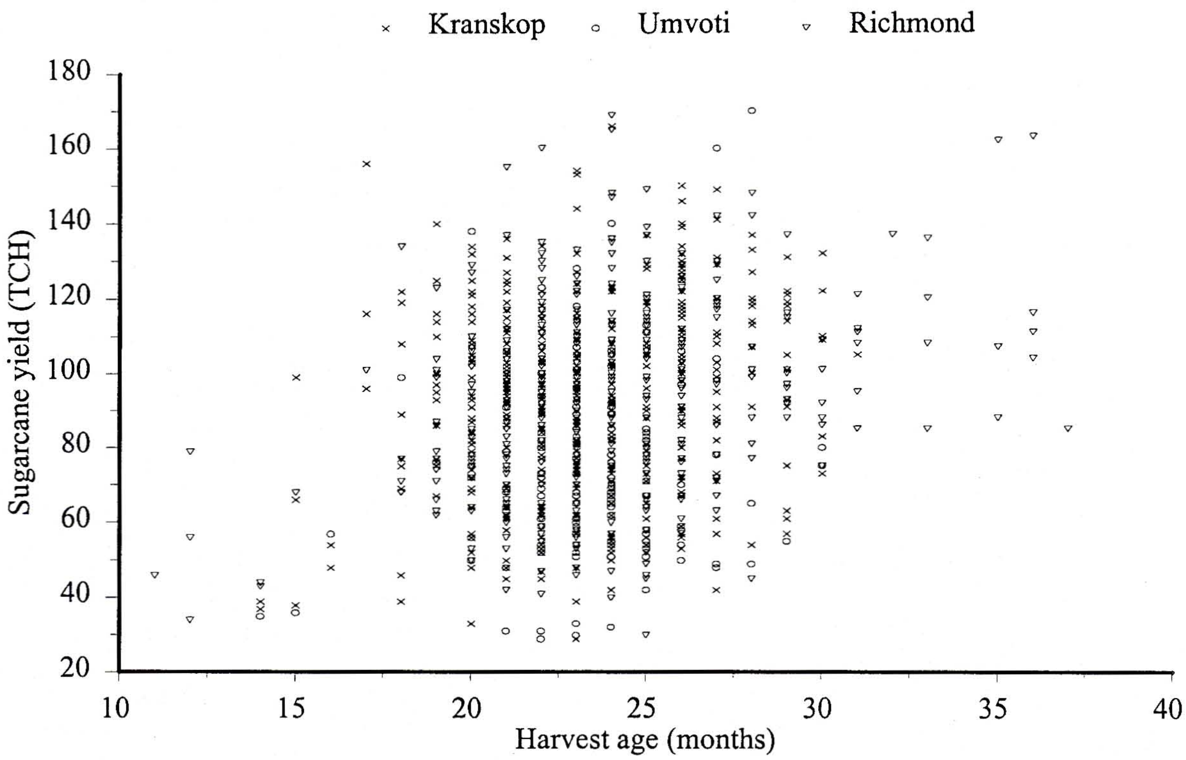


Figure A6.1a: Graph of sugarcane yield versus harvest age at the district level (for all 984 observations – Table 10, Chapter 4).

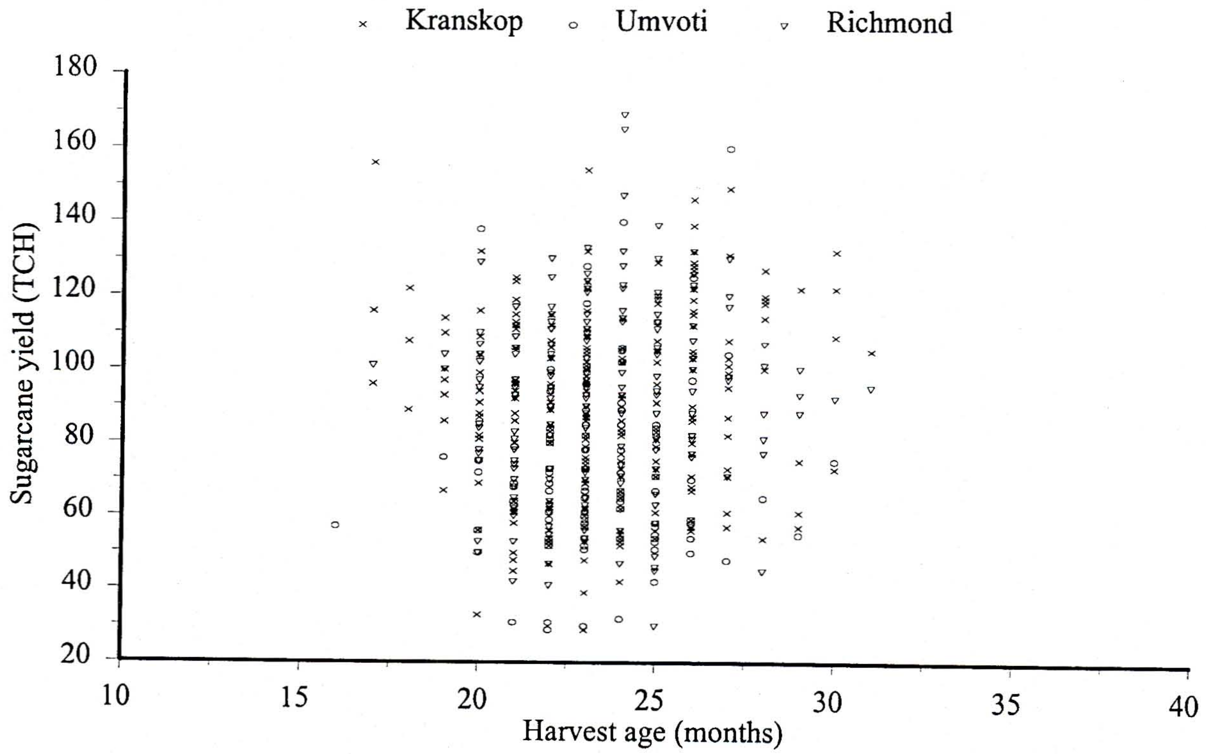


Figure A6.1b: Graph of sugarcane yield versus harvest age at the district level (for the restricted data set of 543 observations – Table 14, Section 4.1.1).

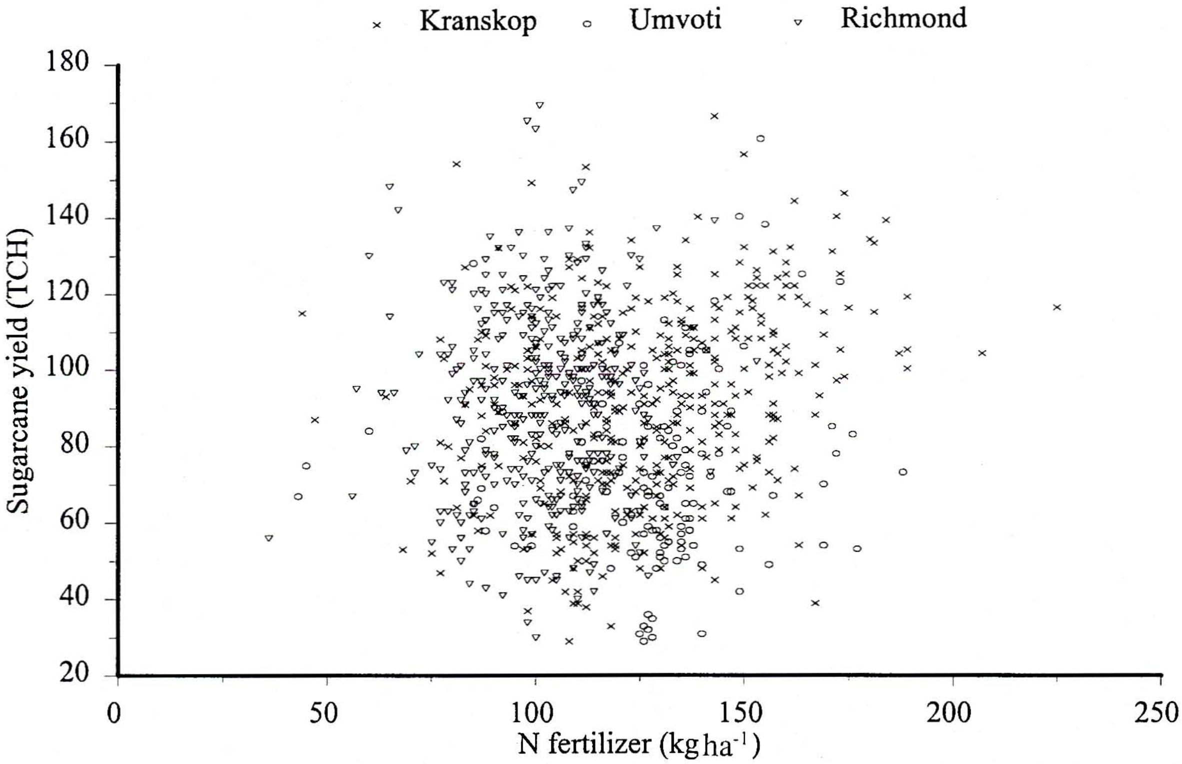


Figure A6.2a: Graph of sugarcane yield versus N fertilizer at the district level (for all 887 observations which have fertilizer records – Table 10, Chapter 4).

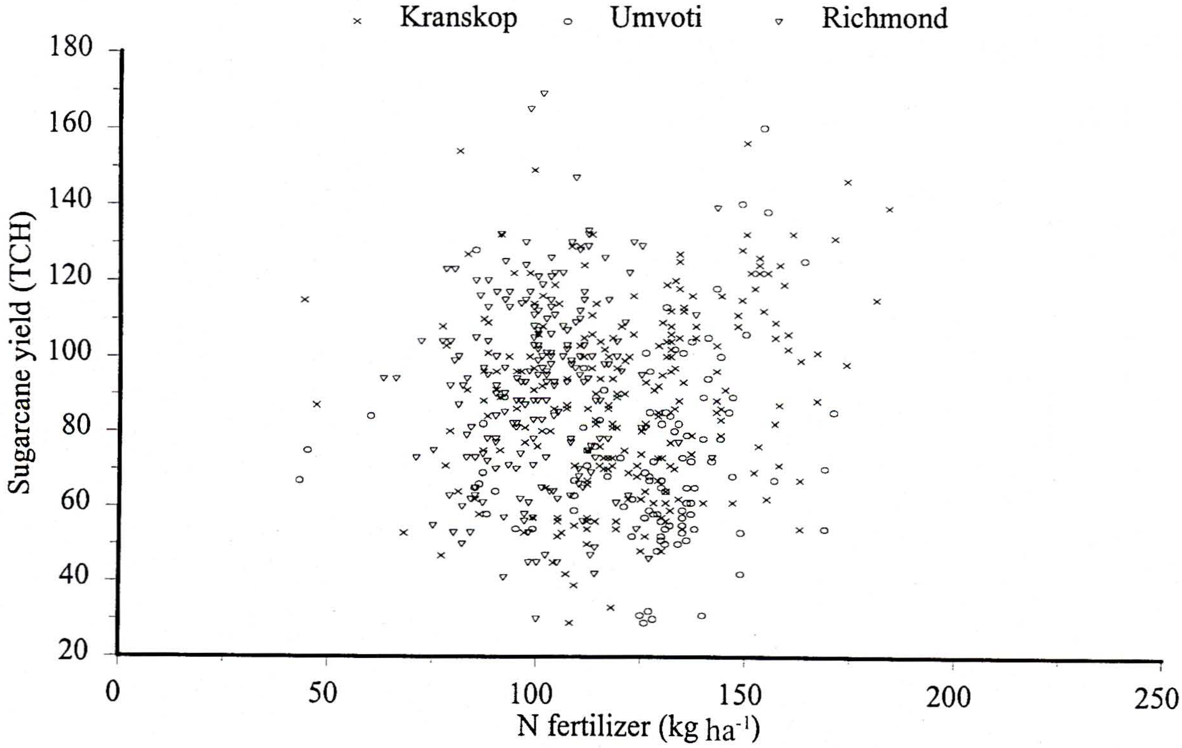


Figure A6.2b: Graph of sugarcane yield versus N fertilizer at the district level (for the restricted data set of 543 observations – Table 14, Section 4.1.1).

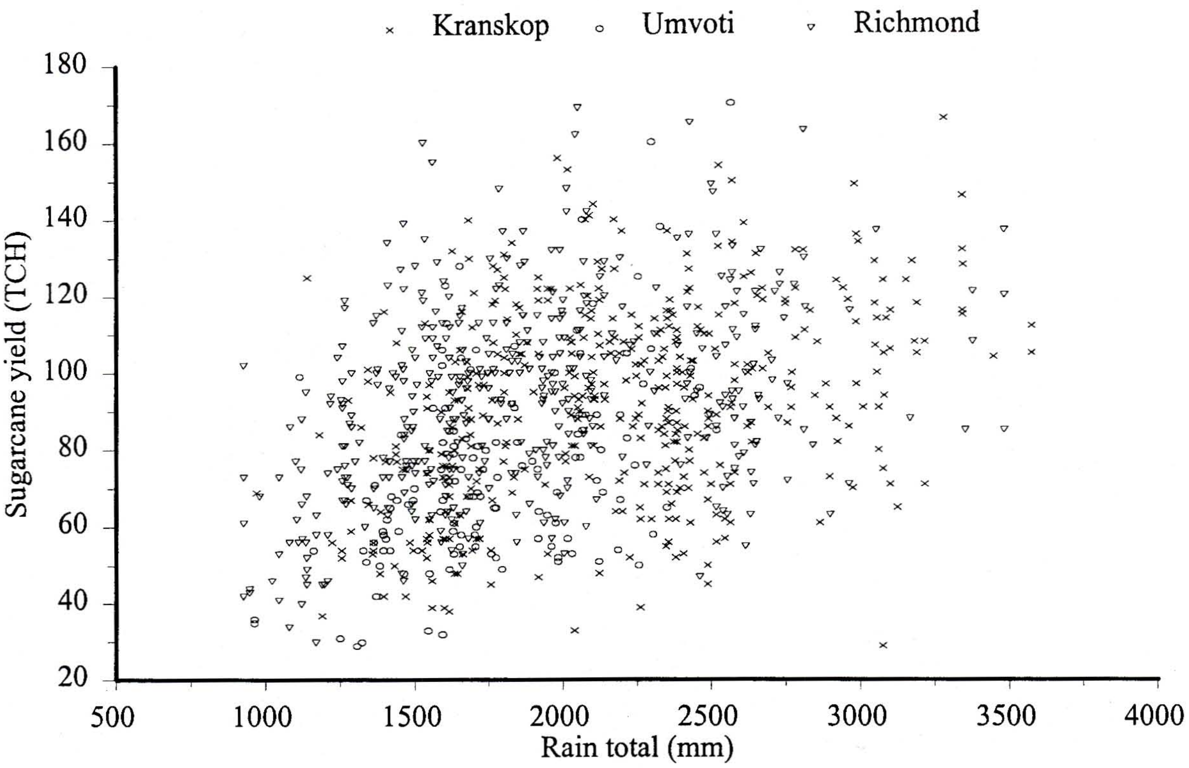


Figure A6.3a: Graph of sugarcane yield versus rain total accumulated for each crop cycle at the district level (for all 984 observations – Table 10, Chapter 4).

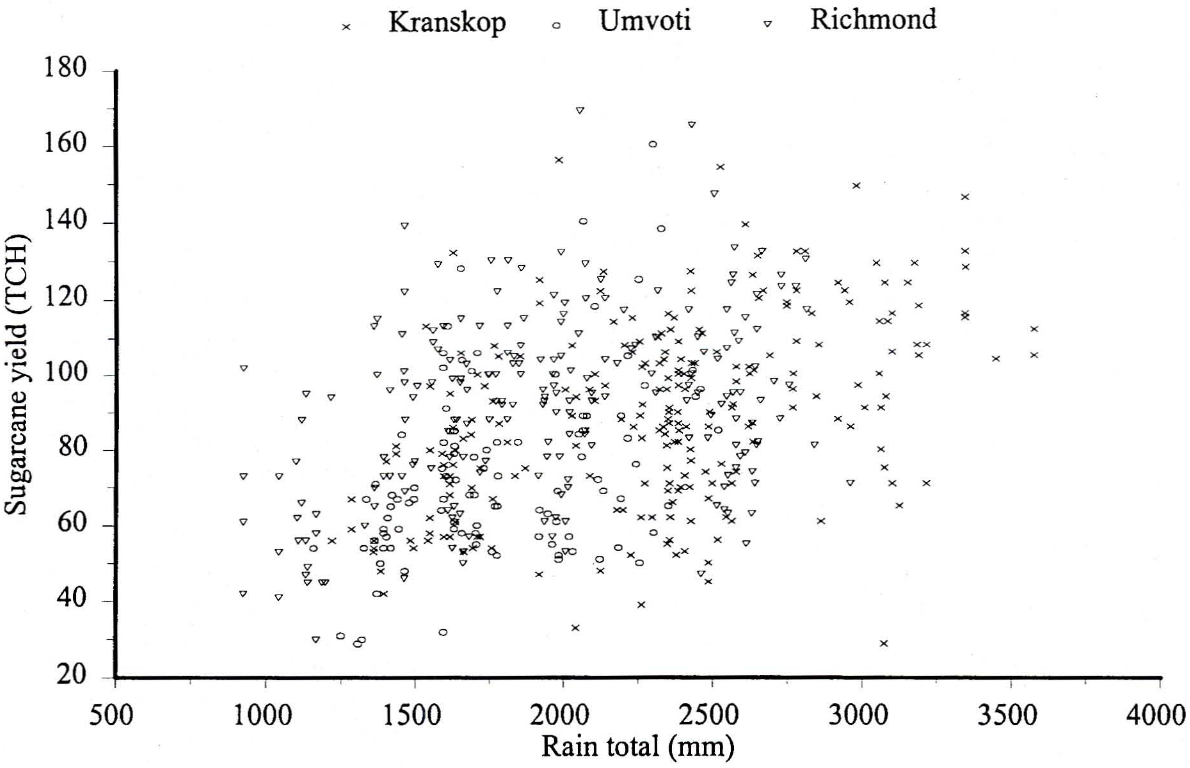


Figure A6.3b: Graph of sugarcane yield versus rain total accumulated for each crop cycle at the district level (for the restricted data set of 543 observations – Table 14, Section 4.1.1).

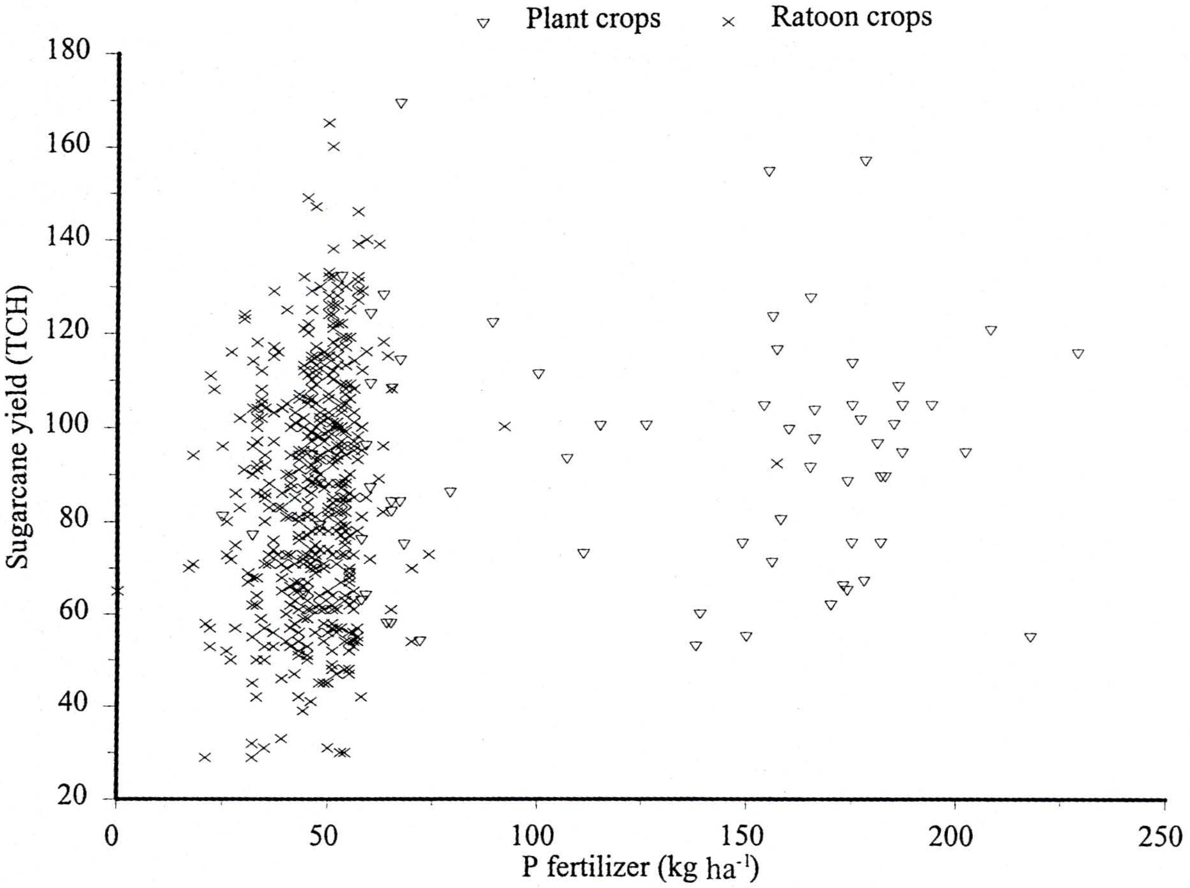


Figure A6.4: Graph of sugarcane yield versus P fertilizer for plant and ratoon sugarcane crops using the restricted data set of 543 observations (Section 4.2).

APPENDIX 7

SUMMARY STATISTICS OF THE VARIABLES AVAILABLE FOR STATISTICAL ANALYSIS

TABLE OF CONTENTS

LIST OF TABLES

A7.1 Dependent variables

Table A7.1: Summary statistics of the potential dependent variables.

Variable	No. of observations	Missing values	Mean	Median	Min	Lower quartile	Upper quartile	Max	Range	SD	CV%	Skewness	Kurtosis
TCH	543	0	87.3	88	29	67	105	169	140	25.2	28.9	0.179	-0.298
TSH	533	10	11.30	11.1	3.6	8.6	13.8	23.5	19.9	3.4	30.5	0.382	-0.007
TCHM	543	0	3.75	3.7	1.2	2.8	4.5	9.2	8.0	1.1	30.4	0.402	0.420
TSHM	533	10	0.49	0.5	0.2	0.4	0.6	1.2	1.0	0.2	31.1	0.435	0.310

A7.2 Independent variables

Table A7.2: Summary statistics of potential categorical variables.

Variable	Group name	Levels	Number of observations
Crop cycle (Section 3.11)	Cycle	PS.S*2S PS.W*2S RS.S*2S RS.W*2S RW.S*2S RW.W*2S	45 25 59 157 61 196
	Harvest season [1]	Summer Winter	165 378
Plant / ratoon crop (Section 3.6.6)	Status	Plant crops Ratoon crops	70 473
	Crop number	Plant Ratoon 1 Ratoon 2 Ratoon 3 Ratoon 4 Ratoon 5 Ratoon 6 Ratoon 7 Ratoon 8 (only at Kranskop) Ratoon 9 (only at Kranskop	70 79 87 78 83 72 48 19 4 3

---

[1] Levels not significantly different from each other were grouped during the initial regression analyses. These groupings suggested a crop cycle classification according to harvest season.

Table A7.2: Summary statistics of potential categorical variables (continued).

Variable	Group name	Levels	Number of observations
Variety (Section 3.6.2)	Variety	N12	169
		N16	36
		NCo293	220
		NCo376	118
Harvest year (Section 2.3.8)	Season	1985	25
		1986	18
		1987	41
		1988	52
		1989	41
		1990	39
		1991	41
		1992	41
		1993	46
		1994	59
		1995	46
		1996	46
		1997	48
Location (Section 3.1)	District	Kranskop	229
		Umvoti	102
		Richmond	212
	Estate	Salem	129
		Sutherlands	100
		Canema	102
		Uplands	174
		Greenhill	38
Terrain (Section 3.1)	Landscape position	Crest - Midslope	193
		Footslope - Midslope	105
		Midslope - Footslope	41
		Midslope - Bottomland	204
Aspect (Section 3.5.1)	Four levels	North	128
		East	105
		South	196
		West	114
	Two levels	North and West	242
		South and East	301
Tree effects (Section 3.5.5)	Trees along any field boundary	Present	411
		Absent	132
	Trees along lower field boundary	Present	336
		Absent	207

Table A7.2: Summary statistics of potential categorical variables (continued).

Variable	Group name	Levels	Number of observations
Frost damage (Section 3.5.5)	Damage class [2]	Winter green Winter brown Winter dead	332 153 58
Soil analysis laboratory (Section 3.7.1)	Laboratory	SASEX GCI	499 44
Soil N (Section 3.7.1)	Soil mineralization	Category II Category III Category IV Missing data	9 204 234 96
Soil carbon	Estimated carbon content	Class I (1.4 - 1.8%) Class II (1.8 - 2.0%)	56 487
Soil type (Section 3.3)	Dominant soil type	Orthic red Humic red Humic yellow Humic red / yellow	49 246 185 63
	Dominant soil colour	Red Yellow Red / yellow	295 185 63
	Humus content	Orthic A-horizon Humic A-horizon	49 494
Lithology (Section 3.3)	Dominant soil parent material	Sandstone Dolerite	331 212
Texture (Section 3.3)	Dominant topsoil texture	Fine sandy clay-loam Clay-loam Clay	331 170 42
Data description (Section 4.1; Table 17)	Data quality	Checked data Remaining data	345 198

---

[2] The frost damage class was not recorded for each observation in the field records and this classification is derived from the field observations made during the 1995 winter.

Table A7.3: Summary statistics of potential continuous variables.

Variable (units)	No. of observations	Missing values	Mean	Median	Minimum	Lower quartile	Upper quartile	Maximum	Range	Standard deviation	CV%	Skewness	Kurtosis
FIELD ATTRIBUTES (area weighted means)													
Altitude (m a.s.l.)	543	0	1025	1032	916	970	1071	1121	205	56.1	5.5	-0.092	-1.309
Slope (%)	543	0	11.2	12	3	8	15	19	16	3.9	34.7	-0.188	-0.748
Effective rooting depth ERD (cm)	543	0	125.1	129	52	114	137	151	99	19.0	15.2	-1.023	0.721
Total available water TAW (mm)	543	0	97.4	103	36	85	111	123	87	16.2	16.6	-0.920	0.281
Topsoil clay content (%)	543	0	31.2	32	25	28	33	43	18	4.1	13.3	0.444	0.557
Subsoil clay content (%)	543	0	41.7	43	26	36	49	53	27	6.4	15.4	-0.433	-0.537
Profile gravel (% 2 – 20 mm fragments)	543	0	6.8	5	0	2	10	30	30	5.9	86.9	1.006	0.626
Profile rocks (% 20 – 500 mm fragments)	543	0	0.621	0	0	0	2	17	17	1.9	299.4	5.389	37.443
Latitude (°S)	543	0	29.4	29.2	29.0	29.0	29.9	29.9	0.9	0.4	1.3	0.374	-1.770
Longitude (°E)	543	0	30.7	30.7	30.3	30.4	30.9	30.9	0.6	0.3	0.9	-0.155	-1.721
CROP ATTRIBUTES													
Harvest age (months)	543	0	23.5	23	16	21	25	31	15	2.5	10.5	0.261	0.241
Summer months growth (months)	543	0	13.8	14	11	13	15	17	6	1.2	8.5	-0.015	1.147
Fallow period (months)	543	0	0.6	0	0	0	4	7	7	1.5	274.8	2.637	5.623
AGRONOMIC INPUTS													
N fertilizer (kg ha^{-1})	543	0	114.2	112	43	98	131	184	141	23.1	20.2	0.191	0.039
P fertilizer (kg ha^{-1})	543	0	56.8	50	0	43	55	229	229	34.8	61.2	2.888	7.854
K fertilizer (kg ha^{-1})	543	0	120.7	128	0	95	167	337	337	66.7	55.3	-0.409	-0.278
Zn fertilizer (kg ha^{-1})	543	0	1.5	1	0	0	2	7	7	1.6	104.6	0.955	0.380
Calmag (kg ha^{-1})	543	0	5.9	0	0	0	93	136	136	23.0	389.0	3.929	14.414
Lime (t ha^{-1})	543	0	0.6	0	0	0	2	14	14	1.5	235.3	3.849	20.419
Gypsum (t ha^{-1})	543	0	0.9	0	0	0	5	7	7	2.0	214.1	1.794	1.440
Diuron herbicide (kg ha^{-1})	389	154	1.6	1.6	0	0.8	2.0	3.6	3.6	0.5	30.9	-0.501	5.481
Paraquat herbicide (kg ha^{-1})	389	154	0.1	0	0	0	0.1	0.4	0.4	0.1	160.7	1.885	3.012
Hexazinone herbicide (kg ha^{-1})	389	154	0.4	0.6	0	0	0.7	1.4	1.4	0.4	79.1	-0.130	-1.041
RAIN													
Two months before start of cycle (mm)	543	0	109.8	75	0	34	151	656	656	114.7	104.5	2.160	6.046
After 12 months' growth (mm)	543	0	1037	961	436	821	1255	2001	1565	333.5	32.2	0.537	-0.217
One month before harvest (mm)	543	0	51.5	26	0	12	70	560	560	62.8	12.8	3.018	15.606
Total accumulated for crop cycle (mm)	543	0	2070	2039	925	1631	2449	3573	2648	533.8	25.8	0.235	-0.521
LEAF ANALYSES													
N (% of SASEX threshold)	450	93	113.5	112.3	64	100.6	124	186	122	18.2	16.0	0.396	0.271
P (% of SASEX threshold)	449	94	100.6	100.0	53	89	113	169	116	17.8	17.7	0.403	0.727
K (% of SASEX threshold)	450	93	131.0	130.5	13	111	150	209	196	25.8	19.7	-0.155	0.206
Ca (% of SASEX threshold)	450	93	202.4	173.3	87	147	243	940	853	86.1	42.5	2.220	11.687
Mg (% of SASEX threshold)	450	93	265.8	250.0	125	200	300	988	863	91.2	34.3	2.944	15.532
S (% of SASEX threshold)	418	125	168.7	167.0	100	150	192	233	133	24.3	14.4	0.281	-0.171
Zn (% of SASEX threshold)	394	149	229.5	138	8	115	169	3823	3815	467.6	203.8	5.502	31.243

Table A7.3: Summary statistics of potential continuous variables (continued).

Variable (units)	No. of observations	Missing values	Mean	Median	Minimum	Lower quartile	Upper quartile	Maximum	Range	Standard deviation	CV%	Skewness	Kurtosis
TOPSOIL ANALYSES (FAS and GCI laboratories)													
Phosphorus desorption index (PDI)	485	58	0.26	0.23	0.04	0.15	0.35	0.71	0.67	0.13	51.7	0.805	0.077
P (kg ha^{-1})	543	0	53.2	40.0	4	25	70	290	286	41.6	78.3	1.732	3.391
K (kg ha^{-1})	543	0	303.8	258.0	48	180	377	1153	1105	187.7	61.8	2.115	5.806
Ca (kg ha^{-1})	543	0	879.5	743.0	45	479	1080	4068	4023	591.6	67.3	1.818	4.855
Mg (kg ha^{-1})	543	0	197.6	165.0	20	102	253	791	771	132.6	67.1	1.584	3.311
Zn (ppm)	332	211	2.6	2.5	0.3	1.4	4.0	9.8	9.5	1.3	50.5	0.824	2.510
pH (water)	542	1	5.03	5.00	4.1	4.8	5.2	6.4	2.3	0.35	6.9	0.443	1.033
Cations (c mol$_c$ kg^{-1})	543	0	3.0	2.6	0.3	1.8	3.8	12.1	11.8	1.8	58.1	1.569	3.582
Bulk density (g mL^{-1})	44	499	0.947	0.915	0.790	0.885	0.965	1.220	0.430	0.104	11.0	1.053	0.303
Exchangeable acidity (c mol$_c$ L^{-1})	44	499	1.55	1.54	0.04	0.98	1.95	3.16	3.12	0.8	51.5	0.094	-0.499
Total cations (c mol$_c$ L^{-1})	44	499	5.2	5.0	3.5	4.2	5.9	8.9	5.4	1.1	22.0	0.980	1.254
Acid saturation (%)	44	499	32.5	30.0	1	16	46	79	78	18.5	56.7	0.268	-0.522
Exchangeable aluminium (ppm)	419	124	53.6	41.0	1	20	77	239	238	43.7	81.5	1.286	1.732
Aluminium saturation index (ASI) (%)	419	124	12.4	8.0	0	3	18	74	74	13	102.8	1.632	2.715
SUBSOIL ANALYSES (GCI laboratory only)													
Bulk density (g mL^{-1})	127	416	0.967	0.970	0.750	0.910	1.020	1.210	0.460	0.081	8.4	0.068	0.291
P (kg ha^{-1})	127	416	5.2	4.0	0	0	8	40	40	6.6	127.5	2.902	10.405
K (kg ha^{-1})	127	416	57.6	45.0	14	34	68	311	297	43.4	75.3	3.168	13.003
Ca (kg ha^{-1})	127	416	306.0	238.0	43	135	440	946	903	222.7	72.8	1.074	0.389
Mg (kg ha^{-1})	127	416	100.4	51.0	0	26	137	627	627	113.7	113.2	1.999	4.295
Zn (ppm)	127	416	0.44	0.40	0.0	0.2	0.6	4.6	4.6	0.49	111.1	5.686	43.110
Exchangeable acidity (c mol$_c$ L^{-1})	127	416	1.15	1.01	0.02	0.37	1.58	5.08	5.06	0.93	81.5	1.453	2.983
Total cations (c mol$_c$ L^{-1})	127	416	2.6	2.4	0.4	1.8	3.1	6.6	6.2	1.1	42.5	0.901	1.185
Acid saturation (%)	127	416	46.2	52.0	1	18	73	91	90	28.4	61.5	-0.205	-1.361
pH (KCl) (pH water equivalents)	127 (127)	416 (416)	4.5 (5.4)	4.3 (5.3)	3.1 (4.1)	4.2 (5.2)	4.6 (5.6)	5.9 (6.9)	2.8 (2.8)	0.4 (0.4)	9.3 (7.3)	0.942 (0.765)	2.799 (2.606)
NUTRIENT RATIOS													
Leaf N:P (using % nutrient content)	449	94	10.4	10.3	6.8	9.4	11.3	15.0	8.2	1.4	13.2	0.307	-0.151
Leaf K:Mg (using % nutrient content)	450	93	7.2	6.9	0.9	5.3	8.9	17.5	16.5	2.7	37.8	0.453	0.317
Leaf Ca:Mg (using % nutrient content)	450	93	1.5	1.4	0.6	1.1	1.7	9.4	8.8	0.6	41.2	5.527	64.764
Topsoil K:Mg (using kg ha^{-1} nutrient content) (Topsoil K:Mg using charge equivalents)	543 (543)	0 (0)	2.0 (0.6)	1.7 (0.5)	0.3 (0.1)	1.0 (0.3)	2.5 (0.8)	13.0 (4.1)	12.7 (4.0)	1.5 (0.5)	74.1 (74.1)	2.363 (2.367)	8.931 (8.961)
Topsoil Ca:Mg (using kg ha^{-1} nutrient content) (Topsoil Ca:Mg using charge equivalents)	543 (543)	0 (0)	5.3 (3.2)	4.0 (2.4)	1.4 (0.9)	3.1 (1.9)	6.0 (3.6)	34.9 (21.2)	33.5 (20.3)	4.0 (2.4)	75.2 (75.2)	3.169 (3.168)	13.302 (13.300)
Total P (kg ha^{-1}) (fertilizer application + topsoil analysis)	543	0	110	92	46	75	131	341	295	50.6	46.0	1.595	2.769
Total K (kg ha^{-1}) (fertilizer application + topsoil analysis)	543	0	425	402	104	304	479	1316	1212	195.7	46.1	1.969	5.590
Total Zn (kg ha^{-1}) (fertilizer application + topsoil analysis)	332	211	7.3	7.0	1.4	4.8	9.0	22.0	20.6	3.2	44.3	0.680	0.967

APPENDIX 8

OBSERVATIONS EXCLUDED AS OUTLIERS FROM THE EXPLANATORY AND PREDICTION MODELS

LIST OF TABLES

ble A8.1: Details of observations excluded as outliers from the explanatory and prediction models.

Observation excluded by:	Reason for exclusion	District	Field number	TCH	Harvest season	Variety	Crop number	Crop cycle	Dominant aspect	Dominant soil type	ERD (cm)	Altitude (m a.s.l.)	Rain (mm) 12 months	Rain (mm) for cycle	N fertilizer (kg ha^{-1})	Total K (kg ha^{-1})	Soil Ca:Mg (from kg ha^{-1})
oth models	1	Kranskop	CS03	29	1988	NCo293	Ratoon 3	RW.S*2S	South	Humic yellows	136	1062	1340	3073	108	619	3.59
oth models	2	Kranskop	AS04	154	1987	NCo376	Plant	PS.W*2S	South	Humic red	118	1055	1124	2524	81	308	3.01
oth models	3	Umvoti	DS23	160	1988	N12	Ratoon 1	RW.S*2S	South	Humic red	112	1105	990	2298	154	503	4.82
oth models	4	Richmond	ES11	169	1991	N12	Plant	PS.S*2S	North	Humic yellows	118	944	1195	2051	101	506	4.50
xplanatory model	5	Kranskop	BS04	156	1990	N12	Plant	PS.S*2S	South	Humic yellows	136	1052	1100	1980	150	514	2.88
xplanatory model	6	Umvoti	DS09	140	1989	NCo293	Ratoon 5	RW.W*2S	South	Humic red	107	1067	1237	2064	149	290	3.59
ediction model	7	Richmond	AS17	165	1987	NCo293	Ratoon 2	RS.S*2S	South	Humic yellows	81	938	926	2428	98	629	7.33
ediction model	8	Richmond	AS05	84	1995	NCo293	Ratoon 6	RS.W*2S	South	Humic yellows	97	956	1143	2069	104	477	34.94

ason for exclusion

Records could not be confirmed, lowest TCH recorded on Salem estate; a small field (0.8 ha) with an exceptionally low yield in spite of high rainfall.

Records could not be confirmed, highest TCH recorded on Sutherlands estate; probably includes some sugarcane from another field.

Records could not be confirmed, highest TCH recorded on Canema estate; probably includes some sugarcane from another field.

Records confirmed by manager, highest TCH ever recorded on Greenhill estate; probably includes some sugarcane from another field.

Records confirmed by manager, highest TCH ever recorded on Salem estate, 9.2 TCHM !; MUST include some sugarcane from another field.

Records could not be confirmed, unusually high TCH for Canema; second highest yield for Umvoti in data base.

Records could not be confirmed, highest TCH ever recorded on Uplands estate; probably includes some sugarcane from another field.

Records confirmed by manager, this observation from the Uplands estate has an exceptionally large topsoil calcium:magnesium ratio compared to all other observations; see summary statistics for topsoil Ca:Mg (from kg ha^{-1}) (Appendix 7).

APPENDIX 9

RESIDUAL PLOTS FOR THE EXPLANATORY AND PREDICTION MODELS

LIST OF FIGURES

Appendix 9: Residual plots for the sugarcane yield explanatory and prediction models.

Standardized residuals

```
     -
     I
 3.0 I                                 *  **
     I                    *       *  * * **      **
     I                     2    *  2 **     3         *
     I              * *    *  ** *2**2 22  *252**2*
     I           * *  *  2 *****22 342433*24*342 222*4 **
     I          * **  33 2 4*2**23*4554345255*42242233***  *
 0.0 I      *        * 362* 224*354*253555335362254332 *2*
     I           *  *    2532**4*3*46443*24*5554*632*2* **  *
     I            * *    4 32 344 5*53233*3*332*222**  2   *
     I                  ** 2     243*22 523**2 *2 *    *    *
     I                       * * *   *2  * * 2 * *   *
     I                                   *   ***    *
-3.0 I                                    *
     -+---------+---------+---------+---------+---------+---------+---
   20.0      40.0      60.0      80.0     100.0     120.0     140.0
```

Predicted values

Figure A9.1: Plot of standardized residuals versus predicted values for the final explanatory model (Model 1.1).

Standardized residuals

```
     -
     I
 3.0 I                                                    ** *
     I                                               222**
     I                                             542
     I                                         8986
     I                                     99992
     I                                79999
 0.0 I                            99999
     I                        99999
     I                    99999
     I               368992
     I           *2442
     I       *****
-3.0 I     *
     -+---------+---------+---------+---------+---------+---------+---
   -3.6      -2.4      -1.2       0.0       1.2       2.4       3.6
```

Standard normal ordinate

Figure A9.2: Normal plot for the final explanatory model (Model 1.1).

Standardized residuals

```
     -
     I
 3.0 I
     I                               ****    **    *
     I                   * *  **2 *   *   2**3 2       **
     I              *  * **  *3**  *2  2*3*2  3*6 **  *
     I       *    **  2 *   3 *32  *2*34**4 363* 4***** 2  *
     I         **  *2 * ** 23*22* 423239254484235532 23222*      **
 0.0 I            *2 *45** 23 2 623232*3445*3332324*24* 2* * 2*2
     I              ***2 5*3*3**545*43434*8235 44262222 2    2*
     I                 * 2** 24*53242   226432  2 22** 2 ** *
     I                    * **  *2 43* *2*3422** * *  ***
     I                   * ** 2   22       **  * **    *     *
     I                          *         **    **   *
-3.0 I
     -+---------+---------+---------+---------+---------+---------+---
   32.0      48.0      64.0      80.0      96.0     112.0     128.0
```

Predicted values

Figure A9.3: Plot of standardized residuals versus predicted values for the prediction model (Model 2).

Standardized residuals

```
     -
     I
 3.0 I
     I                                                 2**** *
     I                                            45442
     I                                        69982
     I                                    29997
     I                                99999
 0.0 I                            99998
     I                        99999
     I                    99992
     I                68993
     I           *2445
     I     * *****
-3.0 I
     -+---------+---------+---------+---------+---------+---------+---
   -3.6      -2.4      -1.2       0.0       1.2       2.4       3.6
```

Standard normal ordinate

Figure A9.4: Normal plot for the prediction model (Model 2).

APPENDIX 10

REVISED REGRESSION MODELS INCLUDING THE TOPSOIL PHOSPHORUS DESORPTION INDEX (PDI)

LIST OF TABLES

Table A10.1: The revised explanatory model.

Categorical variables	Levels	Regression coefficient differences	Standard deviation	t-value (455 d.f.)	Probability
Constant	Reference categories †	60.9	36.9	1.65	0.100
Season	**1** (1993, 1996)	-	(Reference category)		
	2 (1990, 1994, 1995, 1997)	11.98	2.45	4.89	<0.001
	3 (1991, 1992)	20.15	2.91	6.92	<0.001
	4 (1985, 1988)	27.95	3.63	7.70	<0.001
	5 (1986, 1987, 1989)	31.49	3.16	9.97	<0.001
Soil type	**1** (Humic yellow and red / yellows)	-	(Reference category)		
	2 (Humic red)	6.35	2.24	2.84	0.005
	3 (Orthic red)	14.61	3.30	4.43	<0.001
Variety	**1** (NCo376 and NCo293)	-	(Reference category)		
	2 (N16 and N12)	7.44	2.64	2.82	0.005
Plant / Ratoon crop	**1** (Plant, Ratoons 1 and 2)	-	(Reference category)		
	2 (Ratoons 3, 4 and 5)	-6.51	2.08	-3.13	0.002
	3 (Ratoons 6, 7, 8 and 9)	-14.69	3.03	-4.84	<0.001
District	Umvoti	-	(Reference category)		
	Kranskop	37.3	16.4	2.27	0.024
	Richmond	40.9	16.4	2.49	0.013
Crop cycle	Winter harvest	-	(Reference category)		
	Summer harvest	4.18	1.77	2.72	0.007
Aspect	North and West	-	(Reference category)		
	South and East	14.69	6.70	2.19	0.029
Continuous variables		Regression coefficient	Standard deviation	t-value (455 d.f.)	Probability
N fertilizer (kg ha^{-1})		0.1922	0.0401	4.79	<0.001
Rain total for cycle (mm)		0.0124	0.0027	4.68	<0.001
Rain × *North and West*		-	(Reference category)		
Rain × *South and East*		-0.0063	0.0031	-2.03	0.042
Altitude (m a.s.l.)		-0.0800	0.0274	-2.92	0.004
Total K (kg ha^{-1})		0.0160	0.0041	3.89	<0.001
ERD (cm)		0.2117	0.0949	2.23	0.026
ERD × *Umvoti*		-	(Reference category)		
ERD × *Kranskop*		-0.1870	0.1270	-1.47	0.141
ERD × *Richmond*		-0.253	0.1120	-2.25	0.025
Phosphorus desorption index (PDI)		-15.610	9.2600	-1.69	0.093
$R^2 = 0.536$	SE = 16.4 TCH				

† The constant term refers to the base yield for the reference categories.

Table A10.2: The revised prediction model.

Categorical variables	Levels	Regression coefficient differences	Standard deviation	t-value (463 d.f.)	Probability
Constant	Reference categories †	118.1	33.5	3.52	<0.001
Soil type	**1** (Humic yellow and red / yellows)	-	(Reference category)		
	2 (Humic red)	6.41	2.37	2.70	0.007
	3 (Orthic red)	15.55	3.57	4.35	<0.001
Plant / Ratoon crop	**1** (Plant, Ratoons 1 and 2)	-	(Reference category)		
	2 (Ratoons 3, 4 and 5)	-9.08	1.86	-4.88	<0.001
	3 (Ratoons 6, 7, 8 and 9)	-20.20	2.67	-7.58	<0.001
District	Umvoti	-	(Reference category)		
	Kranskop	7.36	3.51	2.10	0.037
	Richmond	2.34	5.56	0.42	0.674
Crop cycle	Winter harvest	-	(Reference category)		
	Summer harvest	6.34	1.88	3.37	<0.001
Continuous variables		Regression coefficient	Standard deviation	t-value (463 d.f.)	Probability
N fertilizer (kg ha^{-1})		0.2177	0.0414	5.26	<0.001
Rain for 12 months' growth (mm)		0.0595	0.0152	3.93	<0.001
(*Rain for 12 months' growth*)^2 (mm^2)		-0.00001464	0.00000645	-2.27	0.024
Altitude (m a.s.l.)		-0.0979	0.0284	-3.44	<0.001
Total K (kg ha^{-1})		0.0191	0.0045	4.27	<0.001
Topsoil Ca:Mg		-0.6960	0.2290	-3.04	0.003
Phosphorus desorption index (PDI)		-28.2300	9.3100	-3.03	0.003
R^2 = 0.440	SE = 17.9 TCH				

† The constant term refers to the base yield for the reference categories.

APPENDIX 11

MODEL VALIDATION DETAILS

LIST OF TABLES

LIST OF FIGURES

Table A11.1: Sugarcane production data for the 1998 harvest season used to validate the explanatory and prediction models.

Field number	**Estate** 1=Salem 2=Sutherlands 3=Canema 4=Uplands 5=Greenhill	**Yield** (TCH)	**Season**	**Soil type** 1=Humic yellows 2=Humic red 3=Apedal red	**Variety** 1=NCo376 and NCo293 2=N16 and N12	**Crop group** 1=Plant, Ratoons 1 and 2 2=Ratoons 3, 4 and 5 3=Ratoons 6, 7, 8, and 9	**District** 1=Umvoti 2=Kranskop 3=Richmond	**Harvest season** 1=Winter 2=Summer	**Aspect** 1=North 2=South	**N fertilizer** (kg ha^{-1})	**Rain for cycle** (mm)	**Rain for 12 months** (mm)	**Altitude** (m a.s.l.)	**Total K** (kg ha^{-1})	**Topsoil Ca:Mg ratio** (from kg ha^{-1})	**ERD** (cm)
DS13	3	127	1998	2	2	3	1	2	2	130	1678	970	1100	447	7.6	107
DS18	3	107	1998	2	2	1	1	2	2	136	1847	1015	1119	378	6.3	140
DS19	3	74	1998	2	2	2	1	1	1	137	1888	1084	1103	339	9.3	113
DS20	3	81	1998	2	2	2	1	1	2	133	1574	969	1111	447	9.1	80
DS22	3	99	1998	2	2	1	1	1	2	136	1745	1015	1099	345	7.1	142
DS24	3	79	1998	2	2	2	1	1	2	126	2030	1100	1121	444	9.1	127
DS25	3	69	1998	2	2	1	1	2	2	88	1634	1026	1108	441	5.1	128
DS26	3	74	1998	2	2	2	1	2	1	129	1928	1084	1080	416	6.0	94
DS33	3	83	1998	2	2	2	1	2	2	133	2106	1084	1071	407	8.6	147
AS04	1	99	1998	2	2	1	2	2	1	94	1933	1061	1073	317	11.1	114
AS07	1	106	1998	1	2	1	2	1	1	134	1941	1023	1070	446	13.6	135
BS01	1	89	1998	1	2	2	2	2	2	132	2051	1083	1084	322	6.8	136
CS01	1	101	1998	1	2	1	2	1	1	129	1818	1061	1061	203	4.0	136
CS04	1	110	1998	1	2	2	2	1	1	133	2125	1083	1051	278	10.6	122
CS05	1	95	1998	1	2	1	2	1	2	133	1772	1061	1071	376	3.1	136
CS06	1	84	1998	1	2	2	2	1	2	132	1980	1088	1068	324	4.3	134
CS07	1	102	1998	1	2	1	2	1	1	127	1987	1023	1071	246	6.0	135
CS09	1	131	1998	1	2	2	2	1	1	131	1973	1088	1066	292	5.7	136
CS10	1	133	1998	2	2	1	2	1	1	132	1973	1088	1076	299	9.3	134
CS11	1	93	1998	1	2	3	2	1	2	133	2125	1083	1092	255	9.2	136
AS05	2	107	1998	1	2	1	2	1	2	154	1994	861	1023	404	8.0	136
AS06	2	108	1998	1	2	1	2	1	2	142	1994	861	1041	394	15.6	115
AS08	2	116	1998	1	2	1	2	1	2	152	1994	861	1023	293	4.6	118
AS10	2	108	1998	1	2	2	2	1	1	145	2017	863	1008	306	5.7	128
AS13	2	88	1998	1	2	2	2	2	2	131	2047	861	961	239	9.4	136
AS14	2	98	1998	1	2	2	2	2	2	132	2047	861	969	436	20.1	136
AS15	2	112	1998	1	2	1	2	2	2	144	1670	1063	984	142	8.8	136
AS16	2	97	1998	1	2	1	2	2	2	148	1774	954	1002	304	21.1	131
AS17	2	118	1998	1	2	1	2	1	2	95	1643	1063	963	337	22.2	133
AS18	2	111	1998	1	2	2	2	1	2	145	2020	861	985	360	26.1	136
AS11	4	100	1998	2	2	1	3	2	2	95	2258	1098	938	426	7.3	115
AS18	4	127	1998	1	2	1	3	1	2	79	2176	1121	918	400	9.4	52
AS19	4	103	1998	3	1	3	3	1	1	93	2149	1159	946	325	6.1	103
BS06	4	148	1998	3	2	1	3	1	2	96	2216	1121	957	584	6.5	107
BS09	4	110	1998	3	2	1	3	1	1	84	2189	1159	978	329	19.2	96
BS10	4	151	1998	3	2	1	3	1	1	80	2189	1159	993	314	8.1	105
BS16	4	124	1998	2	1	3	3	2	1	101	2223	1252	1013	393	5.1	127
BS18	4	92	1998	2	2	1	3	1	1	95	2390	1176	1022	307	13.1	144
BS19	4	81	1998	1	2	1	3	1	1	95	2393	1176	1005	372	6.6	134
BS20	4	86	1998	2	2	1	3	1	1	103	2393	1176	988	420	5.9	109
CS02	4	141	1998	2	2	2	3	2	1	129	3241	1546	1030	705	4.1	146
CS06	4	110	1998	3	2	2	3	1	1	98	2538	1167	975	467	3.2	128
ES05	5	69	1998	1	2	2	3	1	1	93	1858	1160	956	533	5.6	126
ES06	5	136	1998	1	2	2	3	1	1	87	1978	1084	982	655	5.8	144
ES09	5	93	1998	1	2	2	3	1	1	90	1937	1084	933	368	3.8	151
ES10	5	96	1998	1	2	2	3	1	1	114	1937	1084	925	431	8.3	143
ES11	5	102	1998	1	2	2	3	2	1	82	1906	1084	944	634	10.0	118

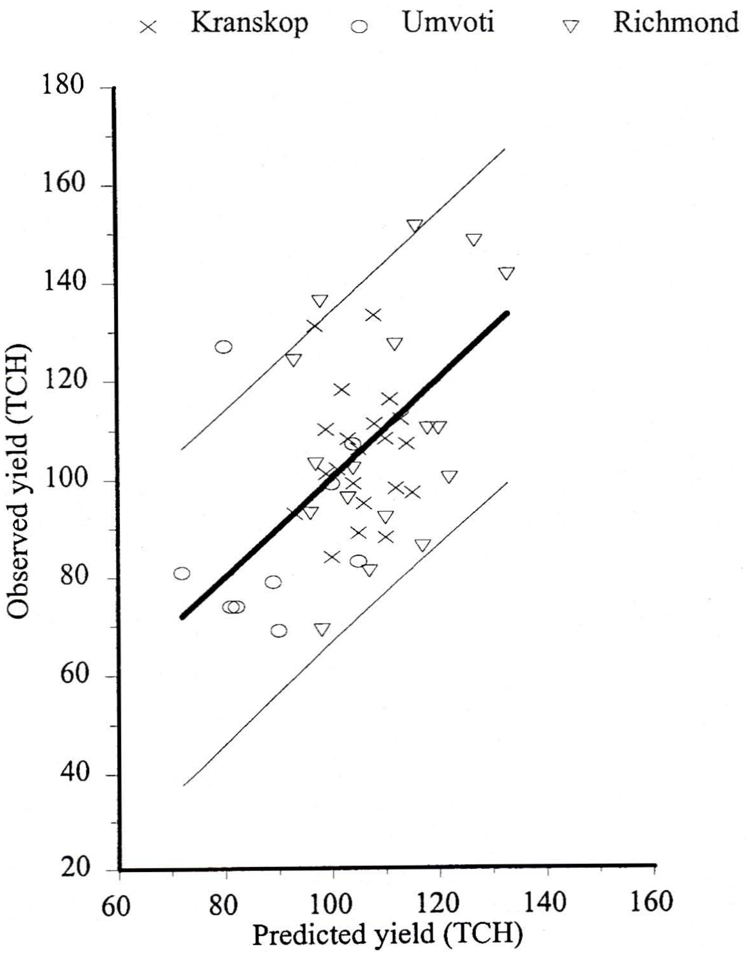


Figure A11.1: Graph of observed versus predicted sugarcane yield (TCH) for the 1998 harvest season (heavy solid line) using the explanatory model (Model 1.1). The 95% confidence limits for individual predictions (light solid lines) are also shown.

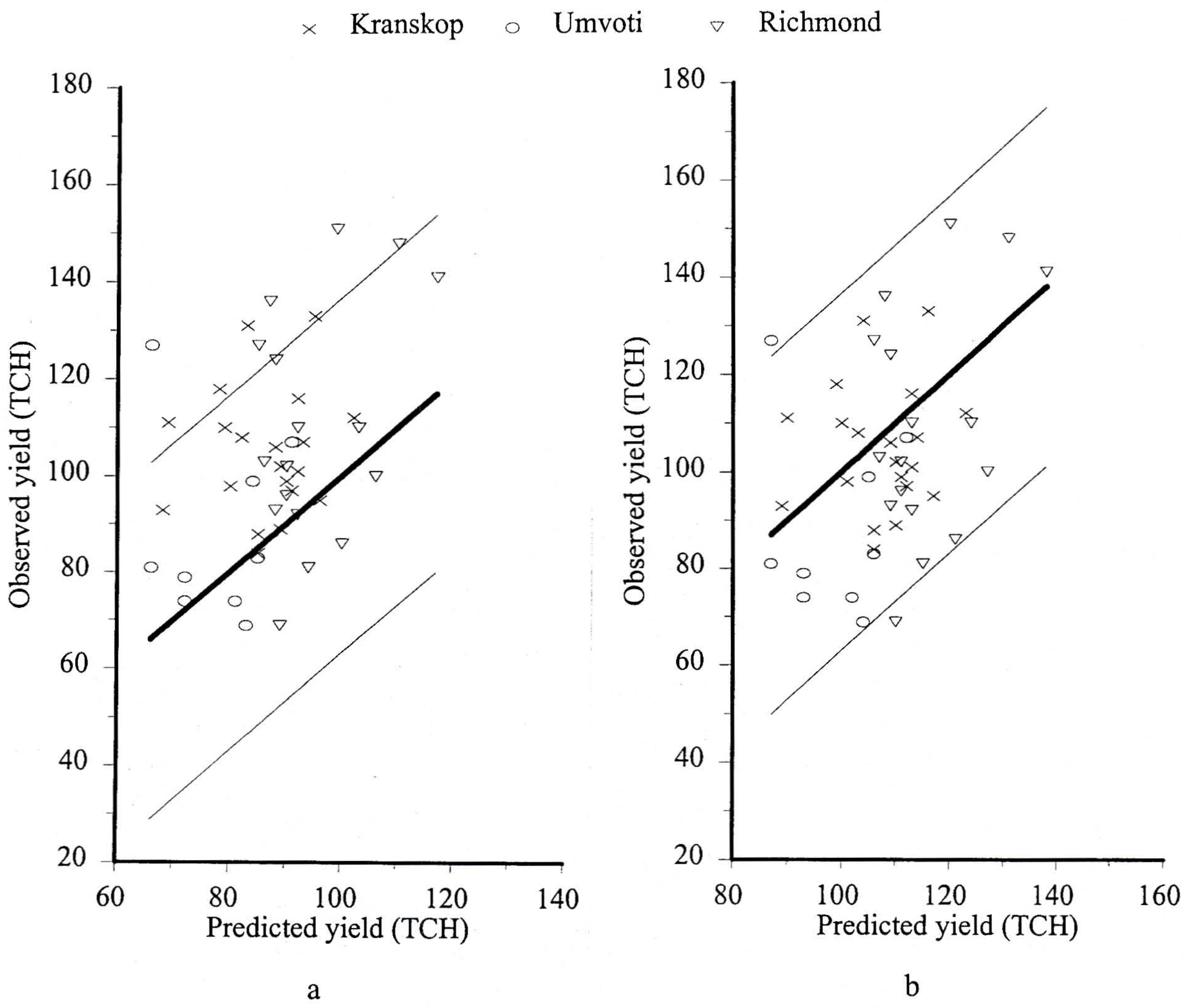


Figure A11.2: (a) Graph of observed versus predicted sugarcane yield (TCH) for the 1998 harvest season (heavy solid line) using the prediction model (Model 2). The 95% confidence limits for individual predictions (light solid lines) are also shown.

(b) Graph of observed versus predicted sugarcane yield (TCH) for the 1998 harvest season using the prediction model (Model 2) and additional information available after the first month's harvest, i.e. the intercept is shifted upwards by 21 TCH to account for the above-average yields. The 95% confidence limits for individual predictions (light solid lines) are also shown.

APPENDIX 12

THEMATIC MAPS OF PREDICTED SUGARCANE YIELDS (TCH) FOR MONDI'S MIDLANDS SUGARCANE ENTERPRISE

LIST OF FIGURES

NOTE: All yield predictions (Figures A12.1 – A12.10) are based on the assumptions specified in Section 4.4 of the thesis. Spatial variables were specified using results from interrogations of the Mondi GIS database (Figures A12.11 – A12.30).

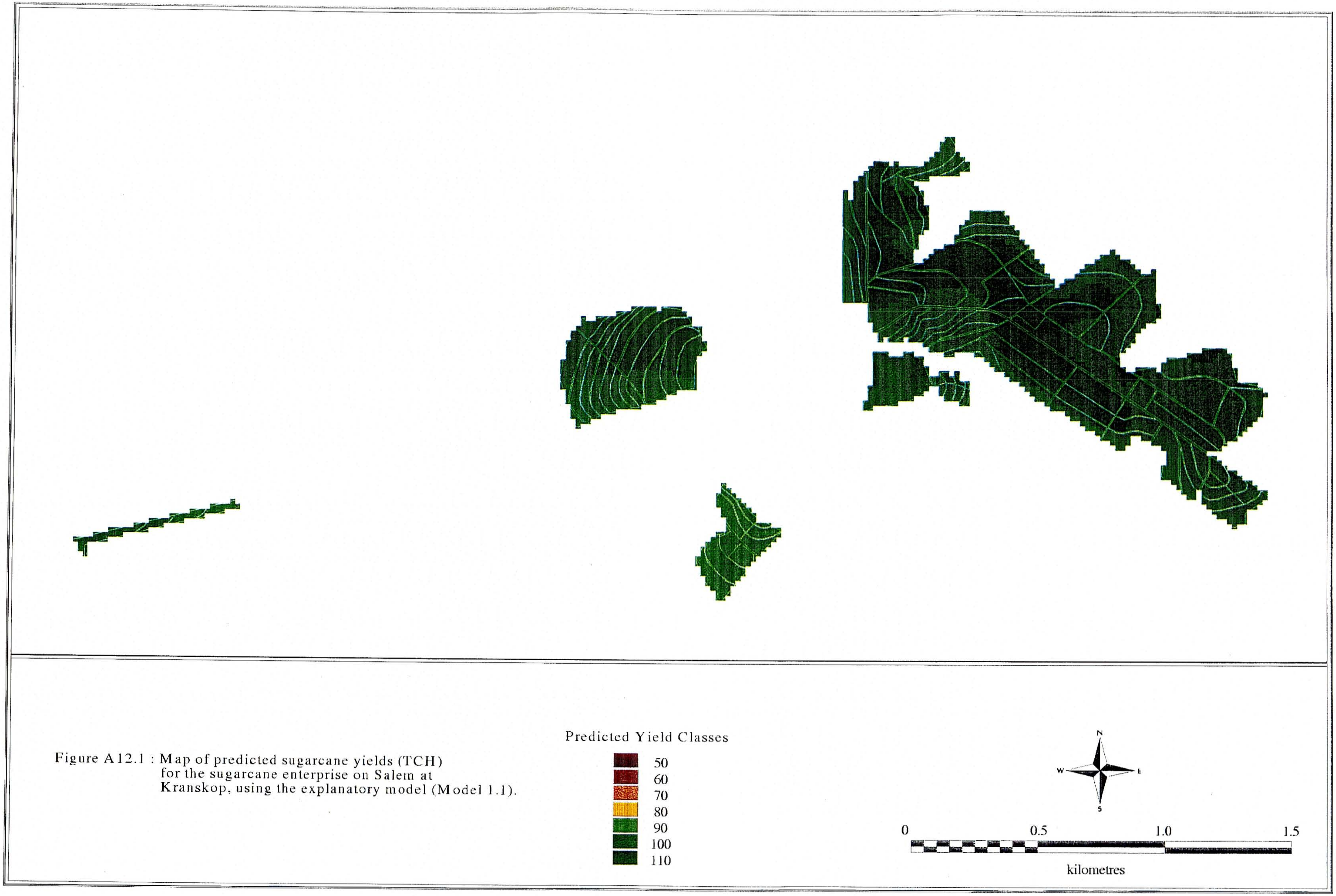


Figure A12.1 : Map of predicted sugarcane yields (TCH) for the sugarcane enterprise on Salem at Kranskop, using the explanatory model (Model 1.1).

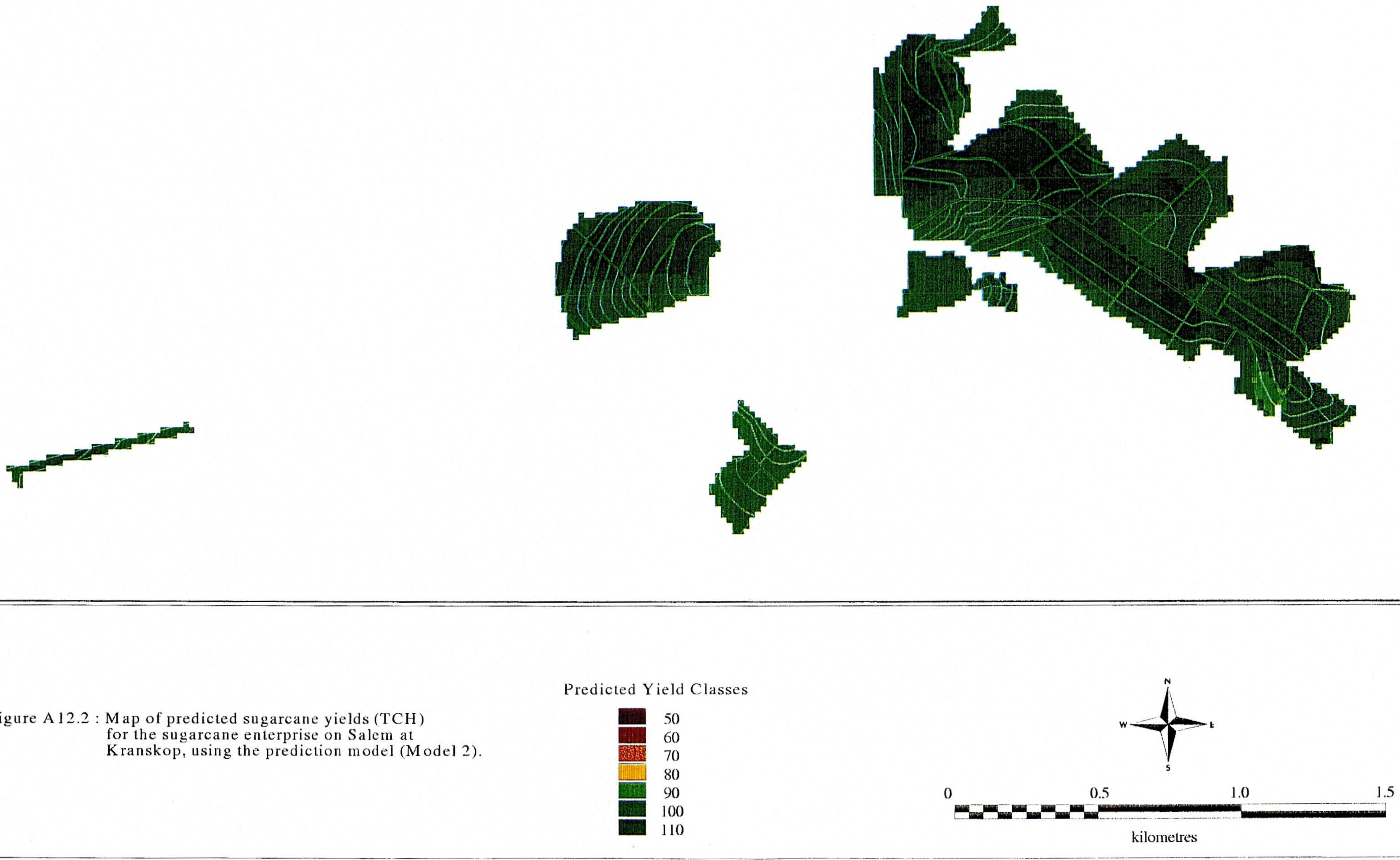


Figure A12.2 : Map of predicted sugarcane yields (TCH) for the sugarcane enterprise on Salem at Kranskop, using the prediction model (Model 2).

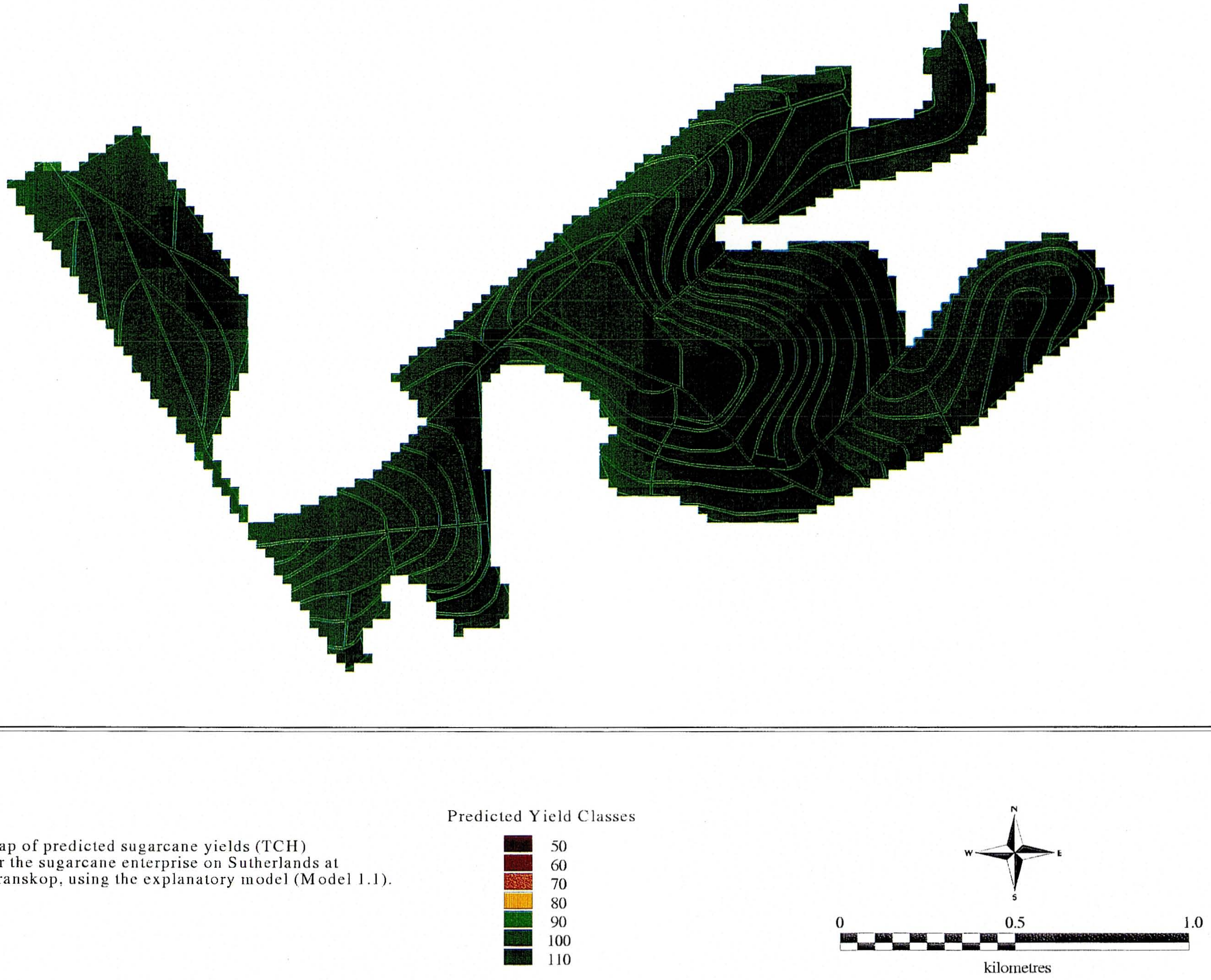


Figure A12.3 : Map of predicted sugarcane yields (TCH) for the sugarcane enterprise on Sutherlands at Kranskop, using the explanatory model (Model 1.1).

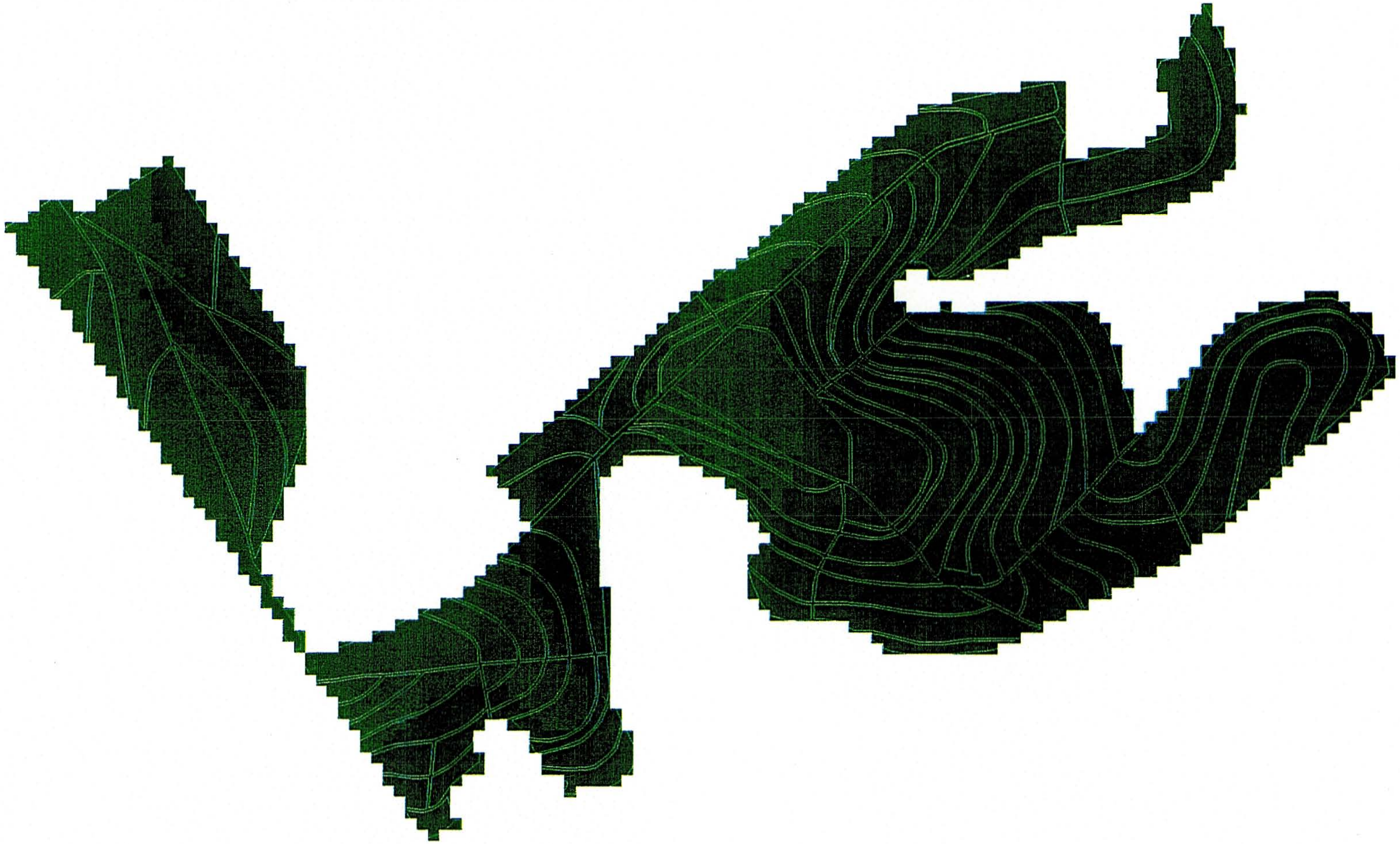


Figure A12.4 : Map of predicted sugarcane yields (TCH) for the sugarcane enterprise on Sutherlands at Kranskop, using the prediction model (Model 2).

Predicted Yield Classes

🟫 50
🟥 60
🟧 70
🟨 80
🟩 90
🟩 100
🟩 110

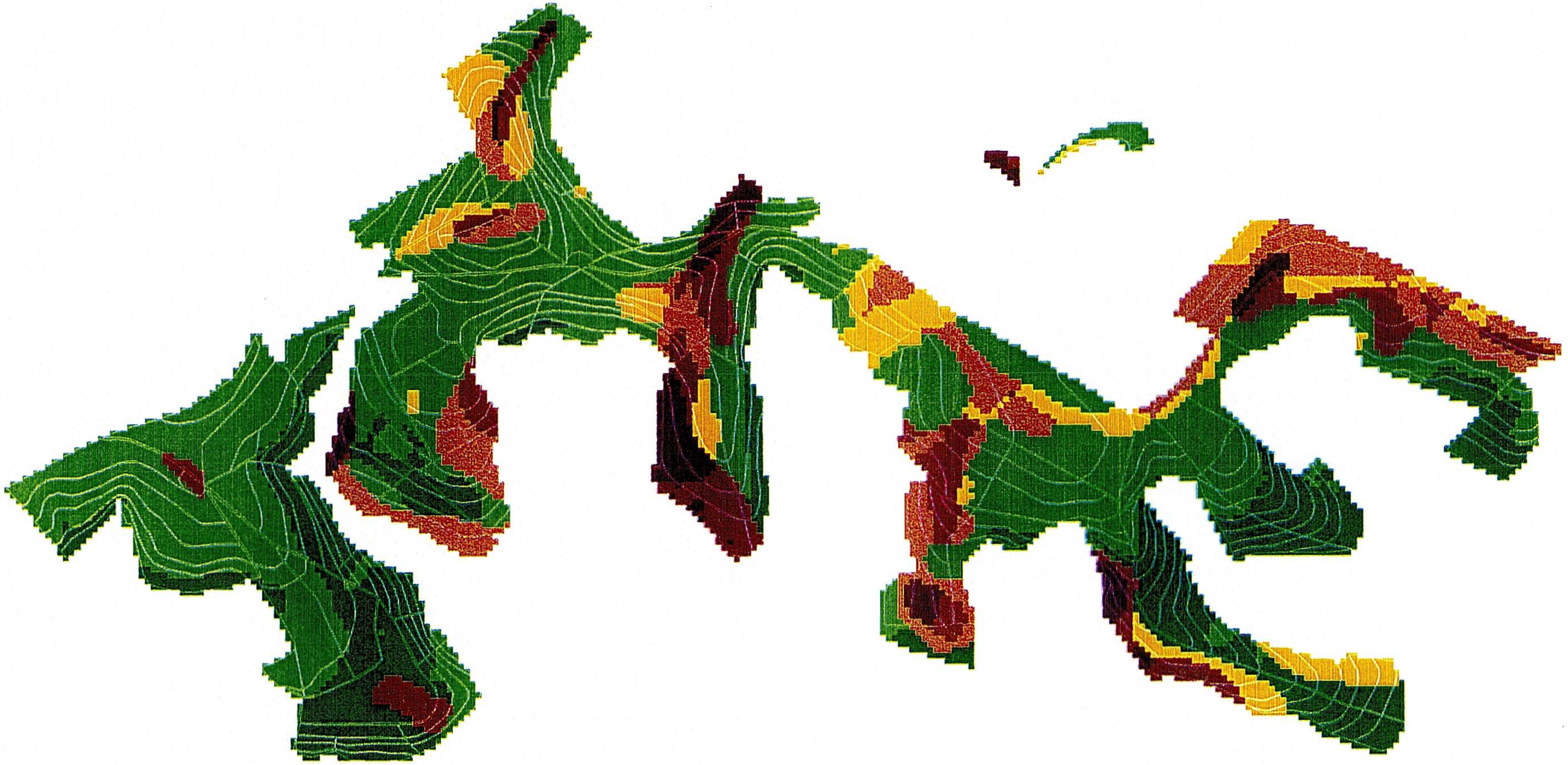


Figure A12.5 : Map of predicted sugarcane yields (TCH) for the sugarcane enterprise on Canema at Umvoti, using the explanatory model (Model 1.1).

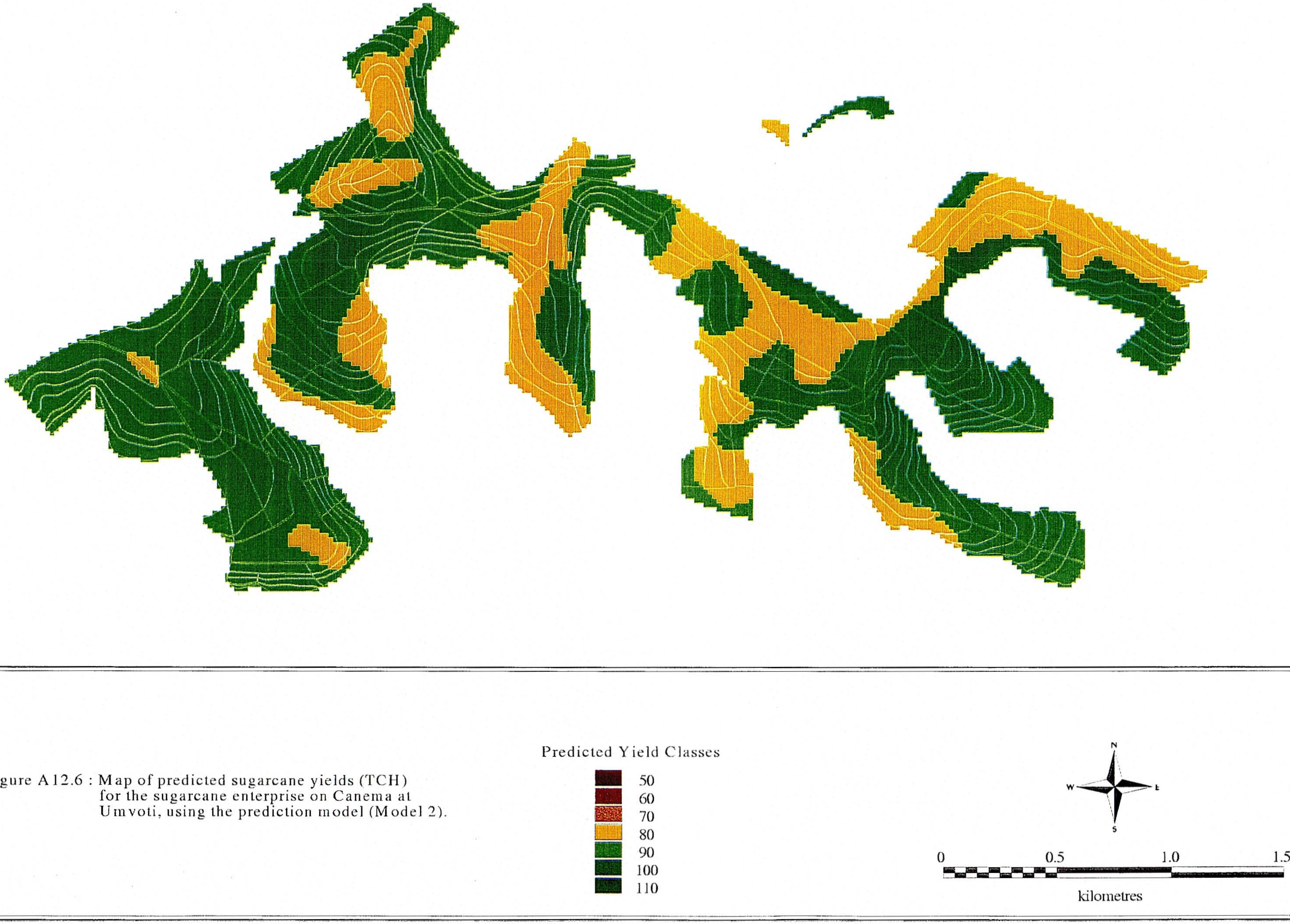


Figure A12.6 : Map of predicted sugarcane yields (TCH) for the sugarcane enterprise on Canema at Umvoti, using the prediction model (Model 2).

Figure A12.7 : Map of predicted sugarcane yields (TCH) for the sugarcane enterprise on Uplands at Richmond, using the explanatory model (Model 1.1).

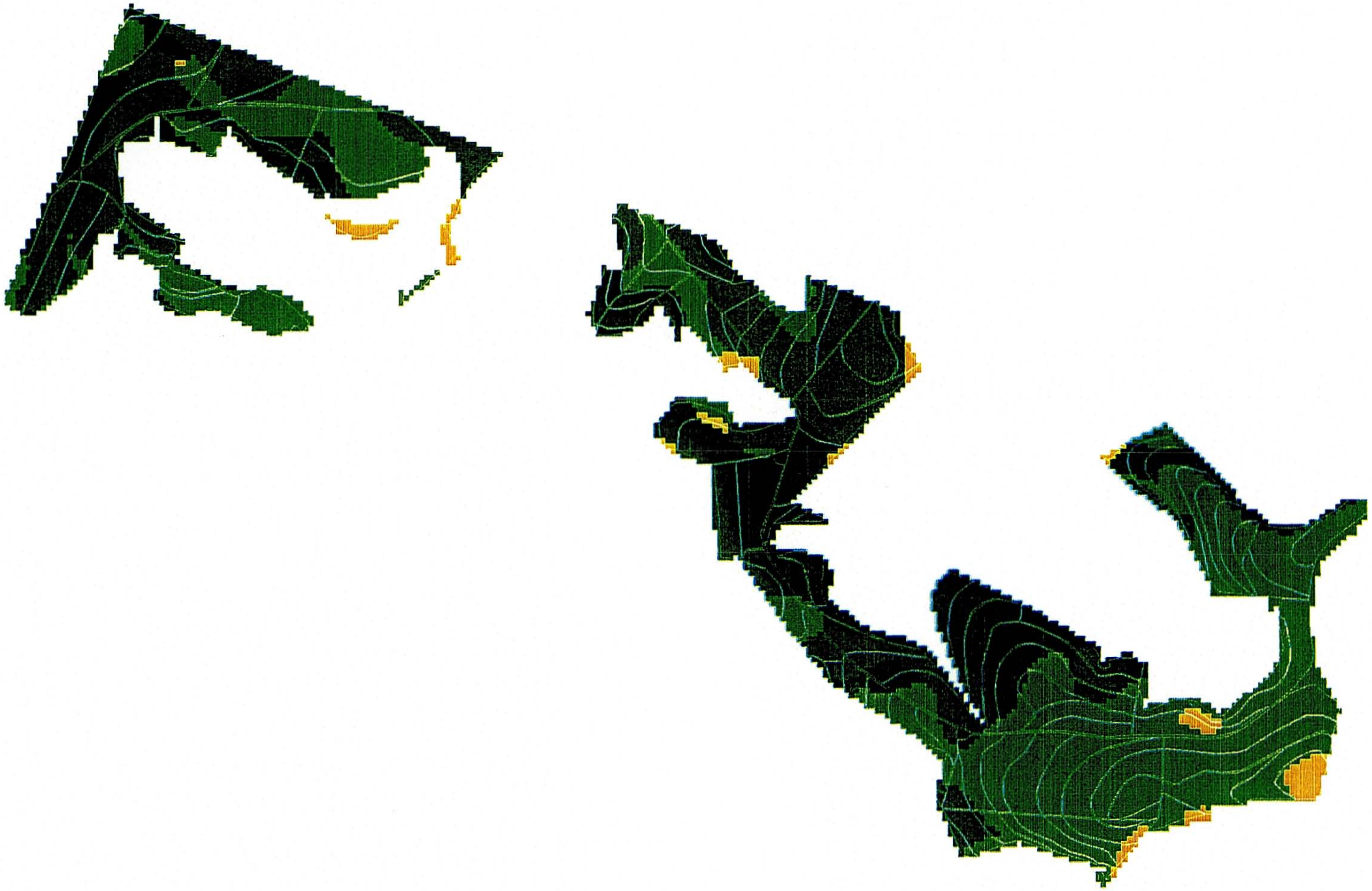


Figure A12.8 : Map of predicted sugarcane yields (TCH) for the sugarcane enterprise on Uplands at Richmond, using the prediction model (Model 2).

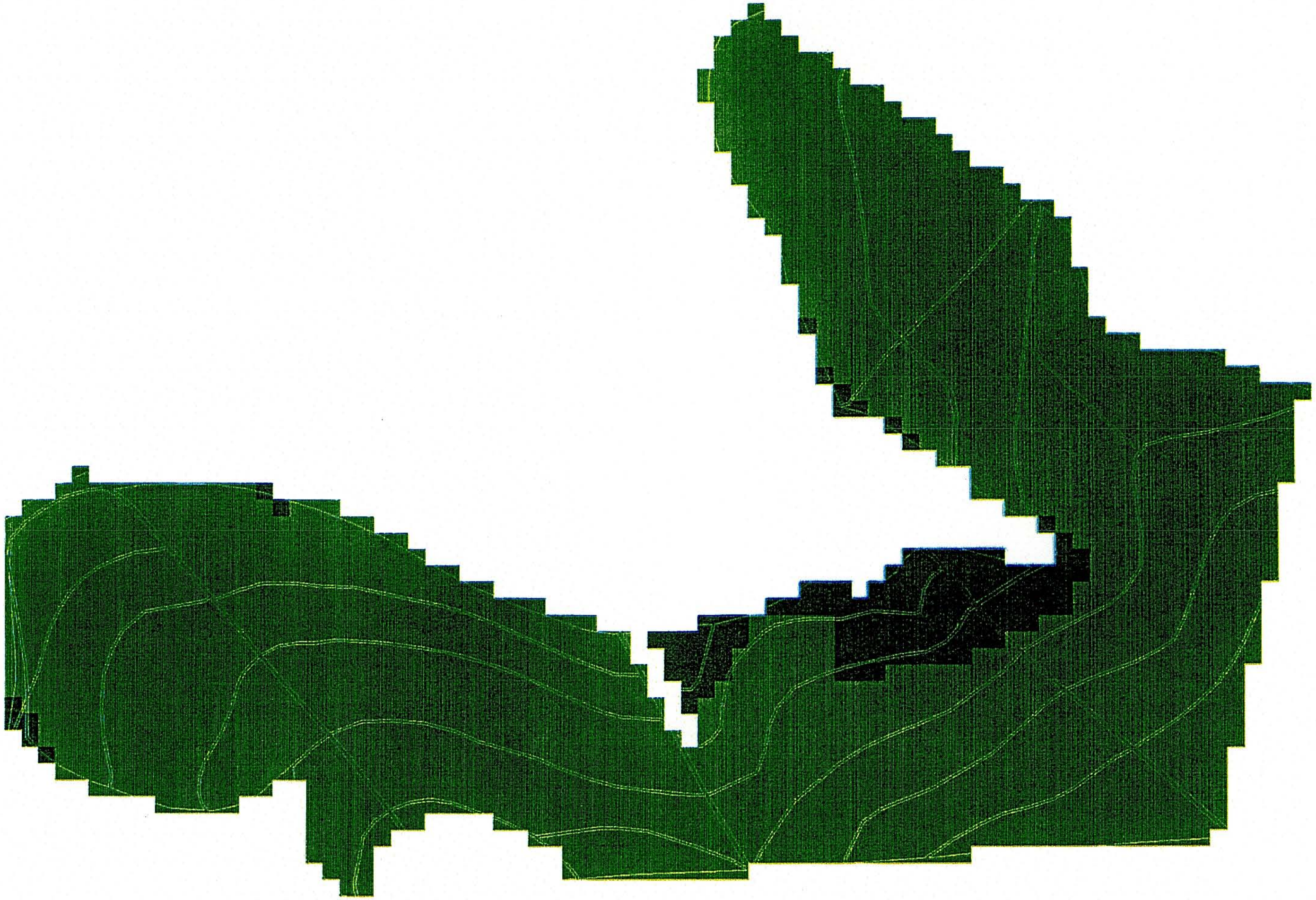
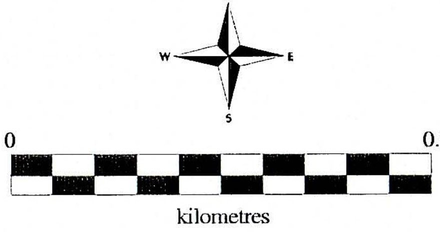


Figure A12.9 : Map of predicted sugarcane yields (TCH) for the sugarcane enterprise on Greenhill at Richmond, using the explanatory model (Model 1.1).

Predicted Yield Classes

🟥 70
🟧 80
🟨 90
🟩 100
■ 110
■ 120



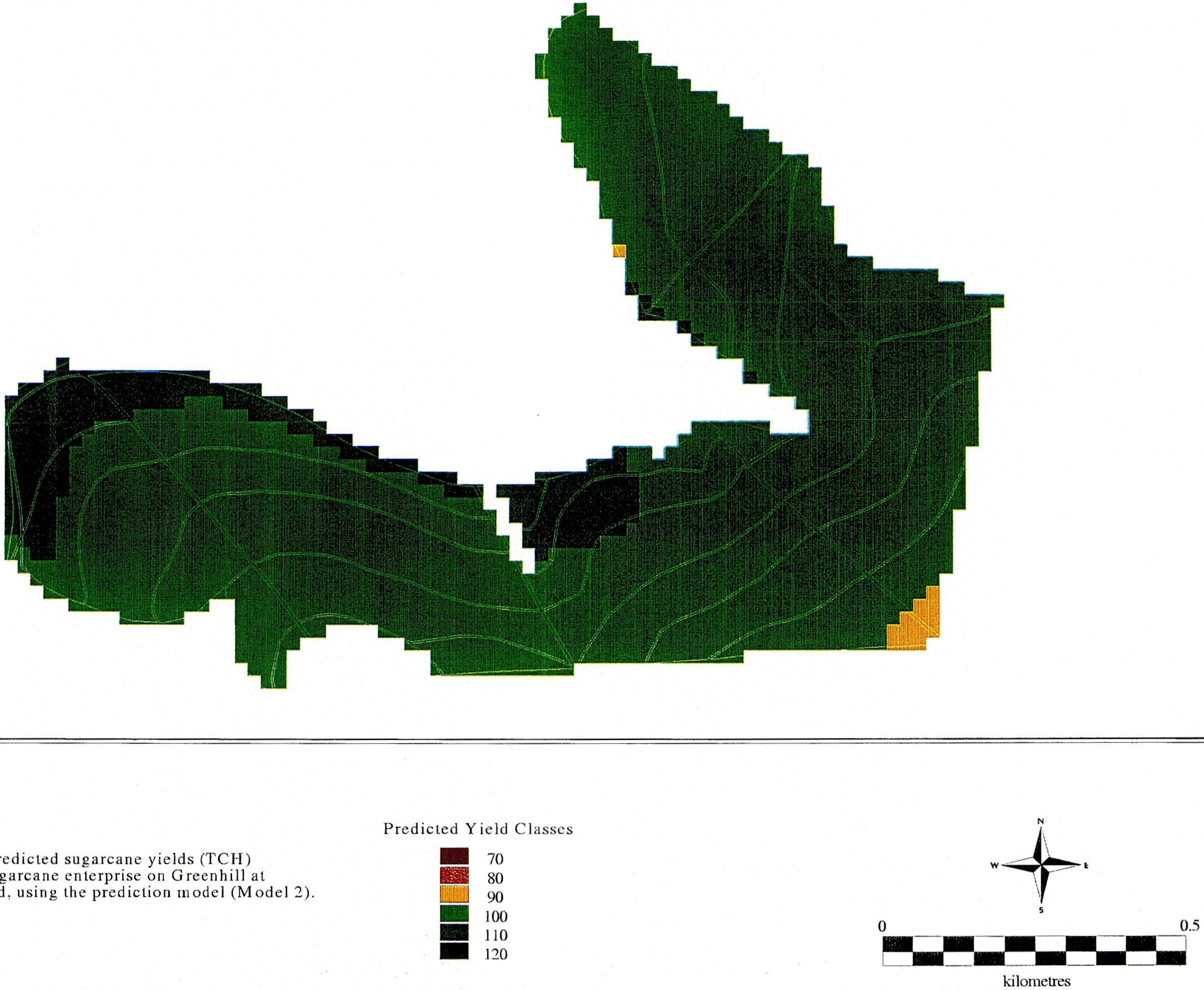


Figure A12.10 : Map of predicted sugarcane yields (TCH) for the sugarcane enterprise on Greenhill at Richmond, using the prediction model (Model 2).

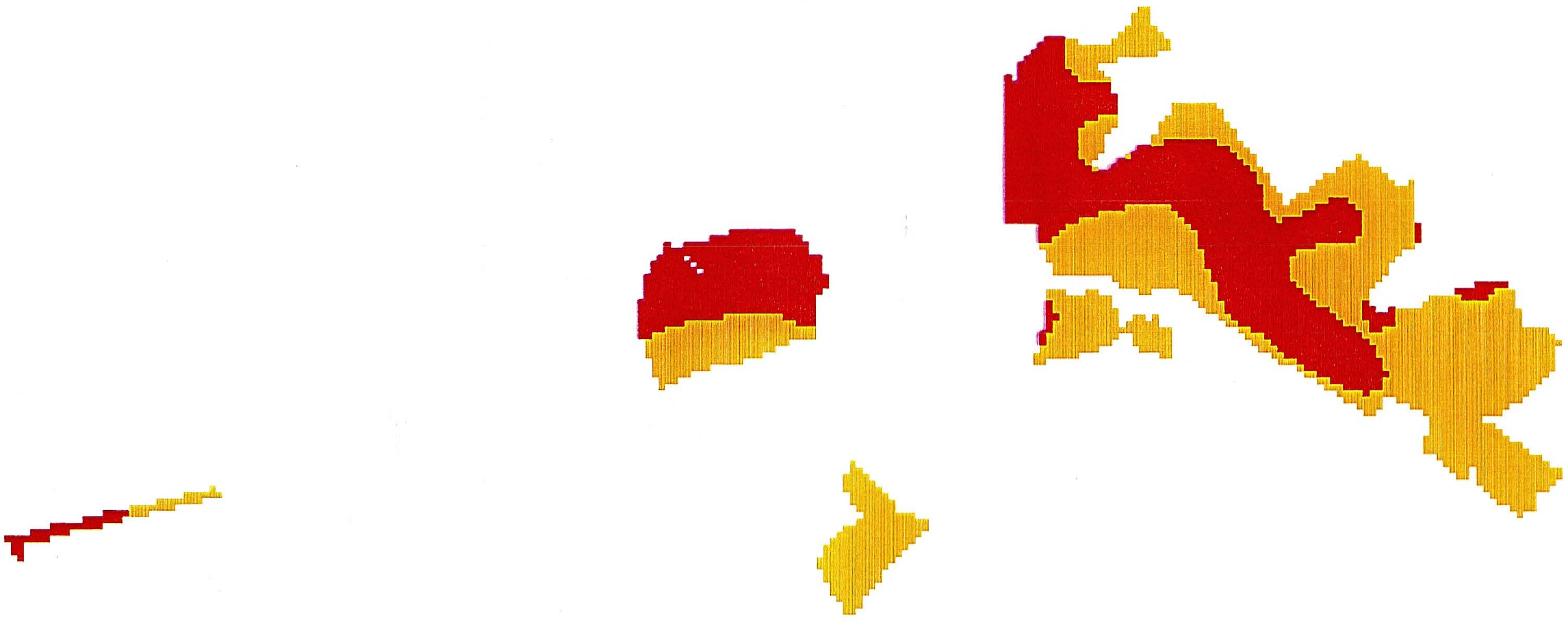


Figure A12.11
Map of soil type for the sugarcane enterprise on Salem at Kranskop.

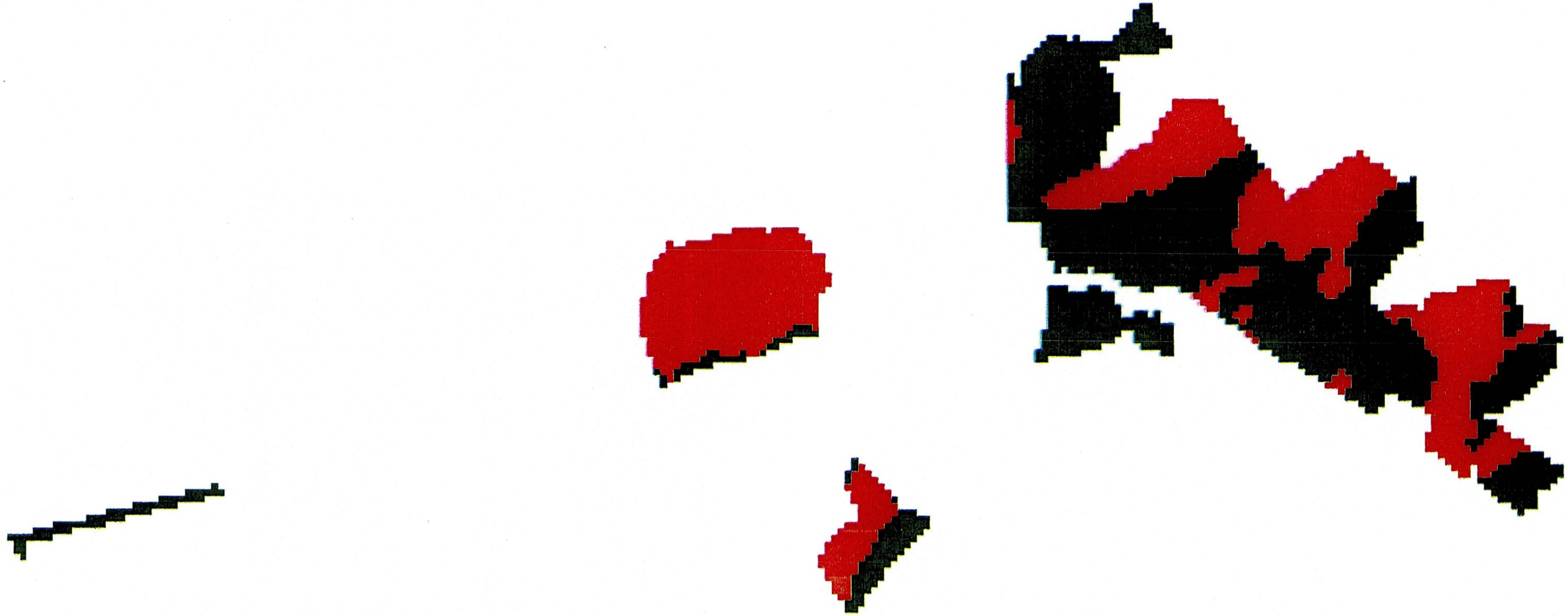


Figure A12.12
Map of aspect for the sugarcane enterprise on Salem at Kranskop.

Figure A12.13
Map of altitude sugarcane enterprise on Salem at Kranskop.

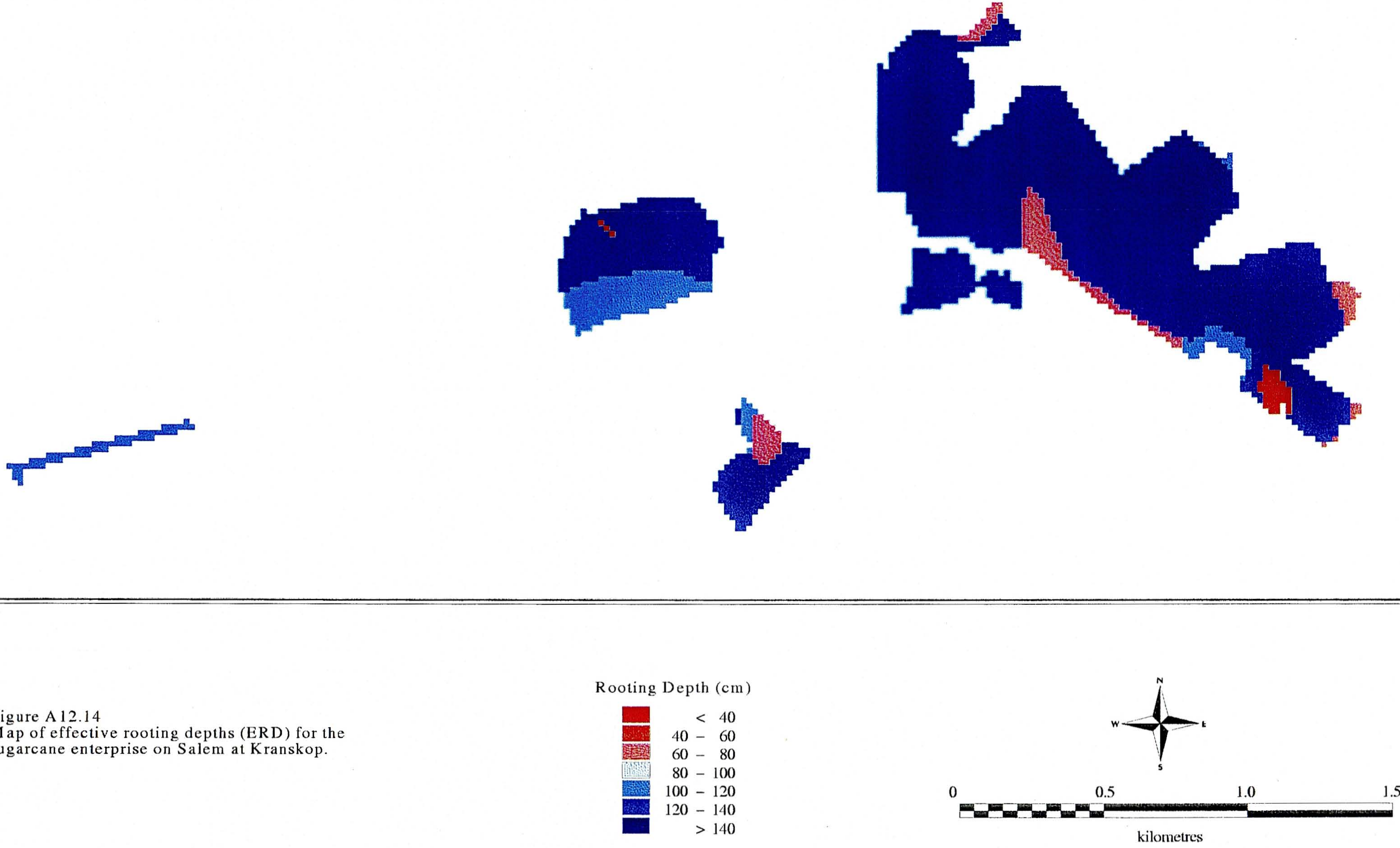


Figure A12.14
Map of effective rooting depths (ERD) for the sugarcane enterprise on Salem at Kranskop.

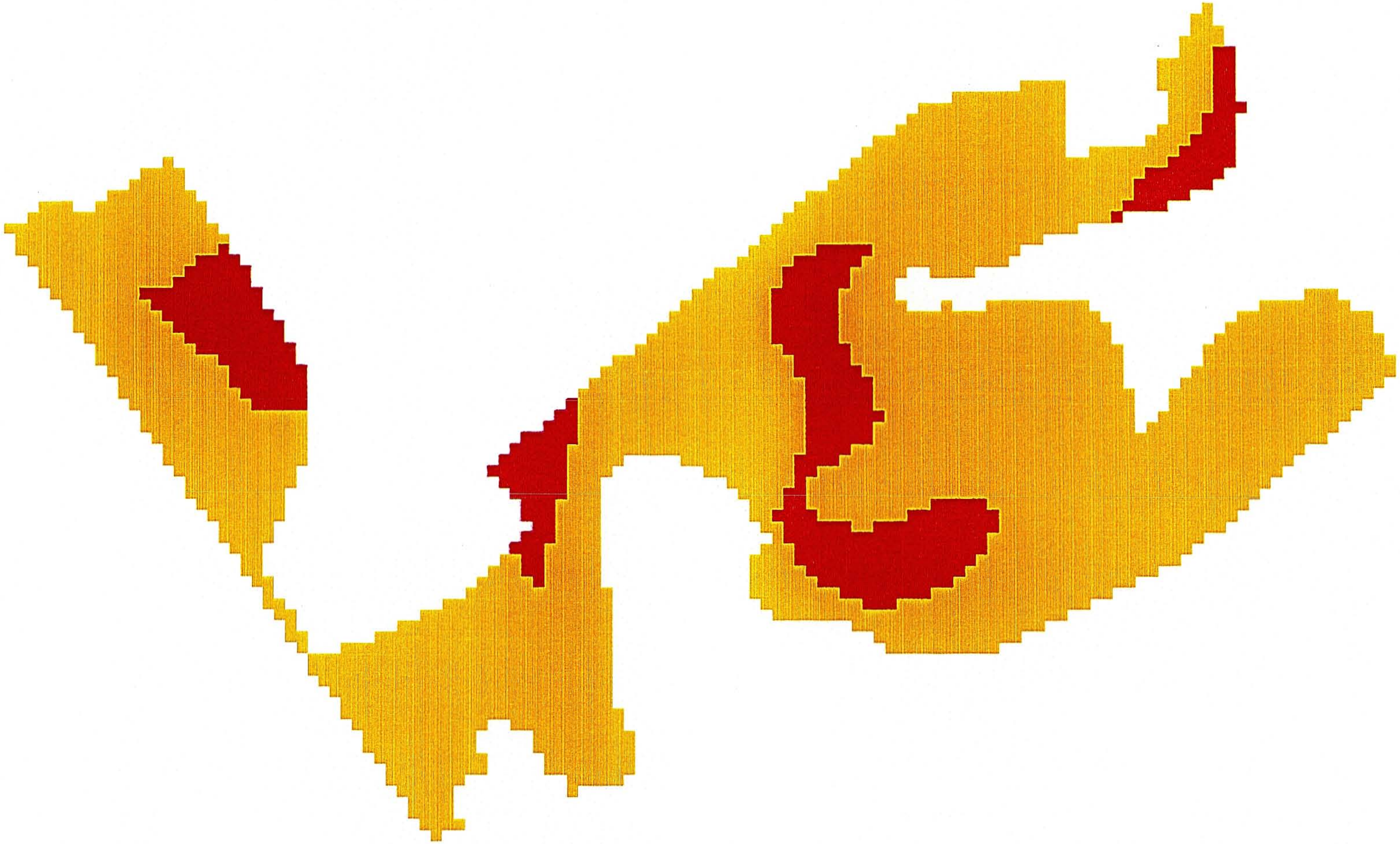
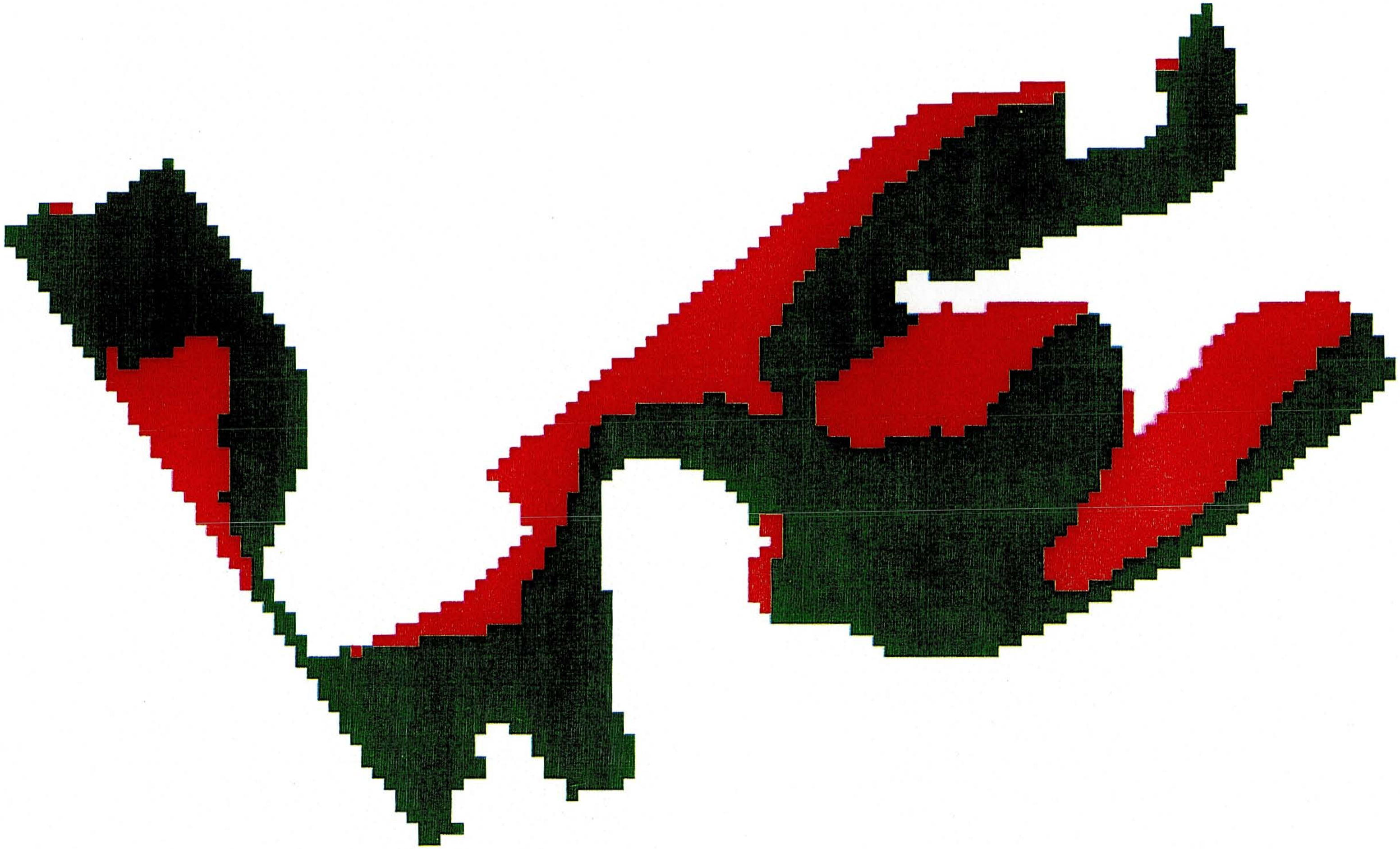


Figure A12.15
Map of soil type for the sugarcane enterprise on Sutherlands at Kranskop



Aspect

🟥 North West
🟩 South East

Figure A12.16
Map of aspect for the sugarcane enterprise on Sutherlands at Kranskop

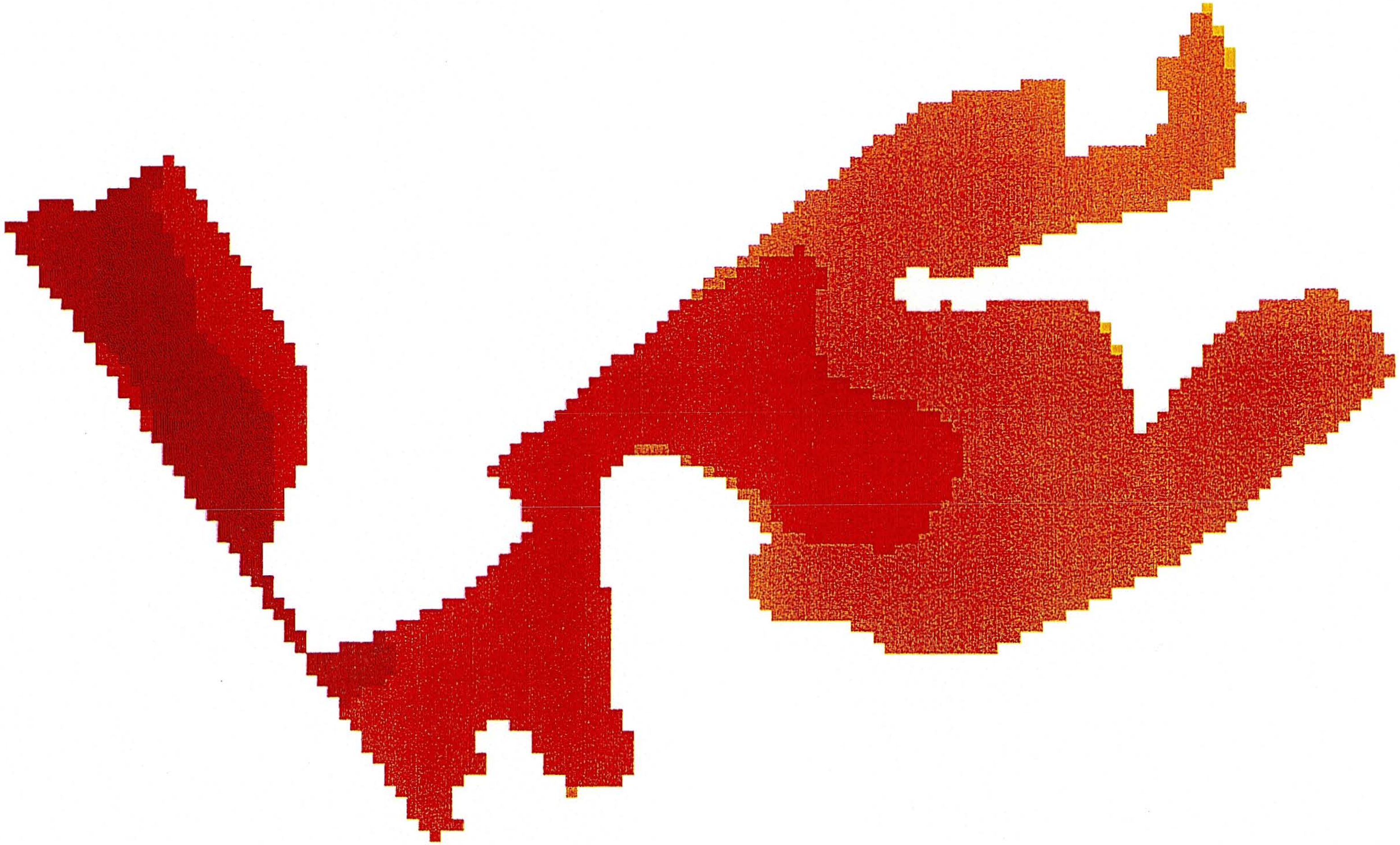


Figure A12.17
Map of altitude for the sugarcane enterprise on Sutherlands at Kranskop

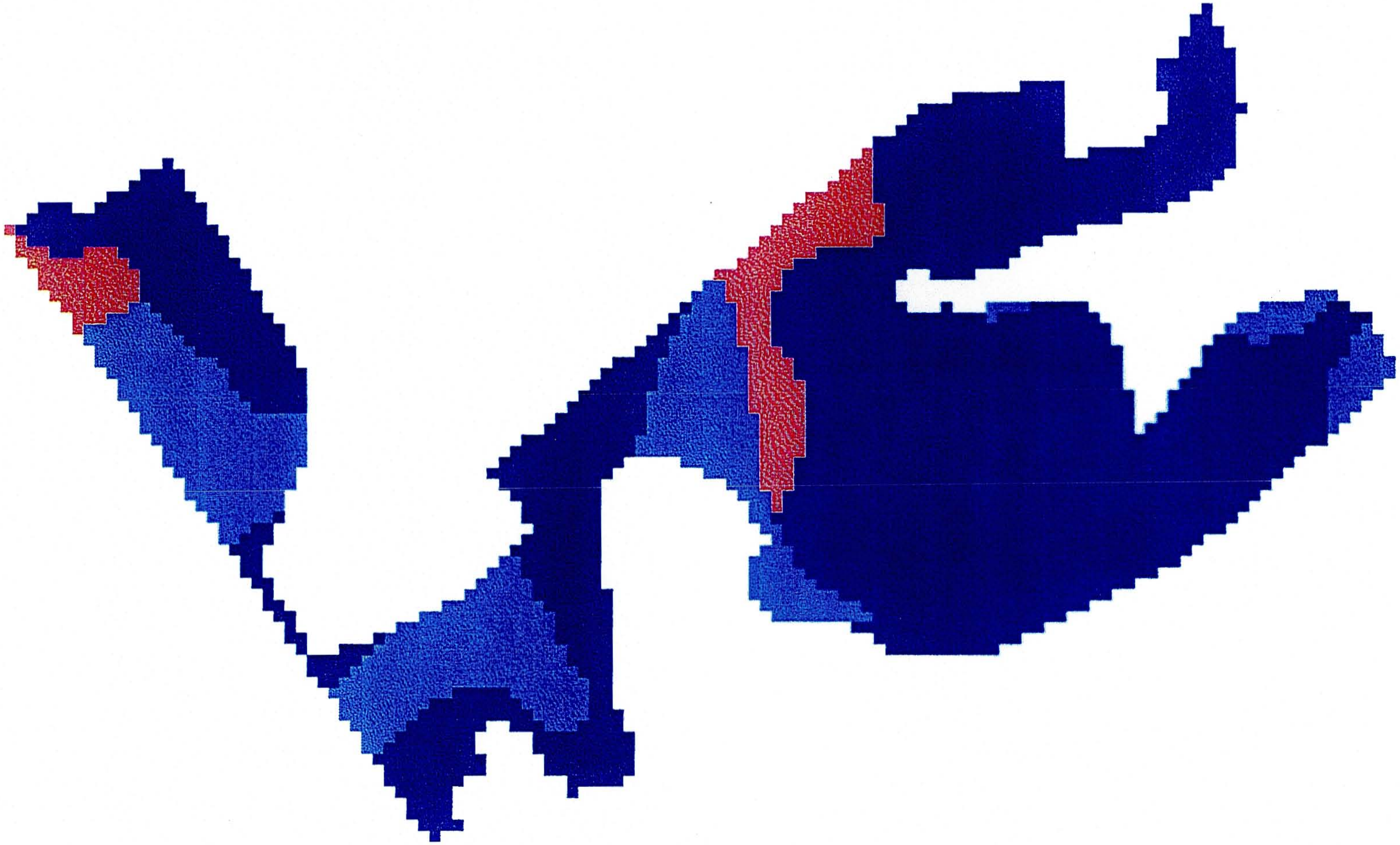


Figure A12.18
Map of effective rooting depths (ERD) for the sugarcane enterprise on Sutherlands at Kranskop

Figure A12.19
Map of soil types for the sugarcane enterprise on Canema at Umvoti.

Figure A12.20
Map of aspect for the sugarcane enterprise on Canema at Umvoti.

Aspect

🟥 North West
⬛ South East

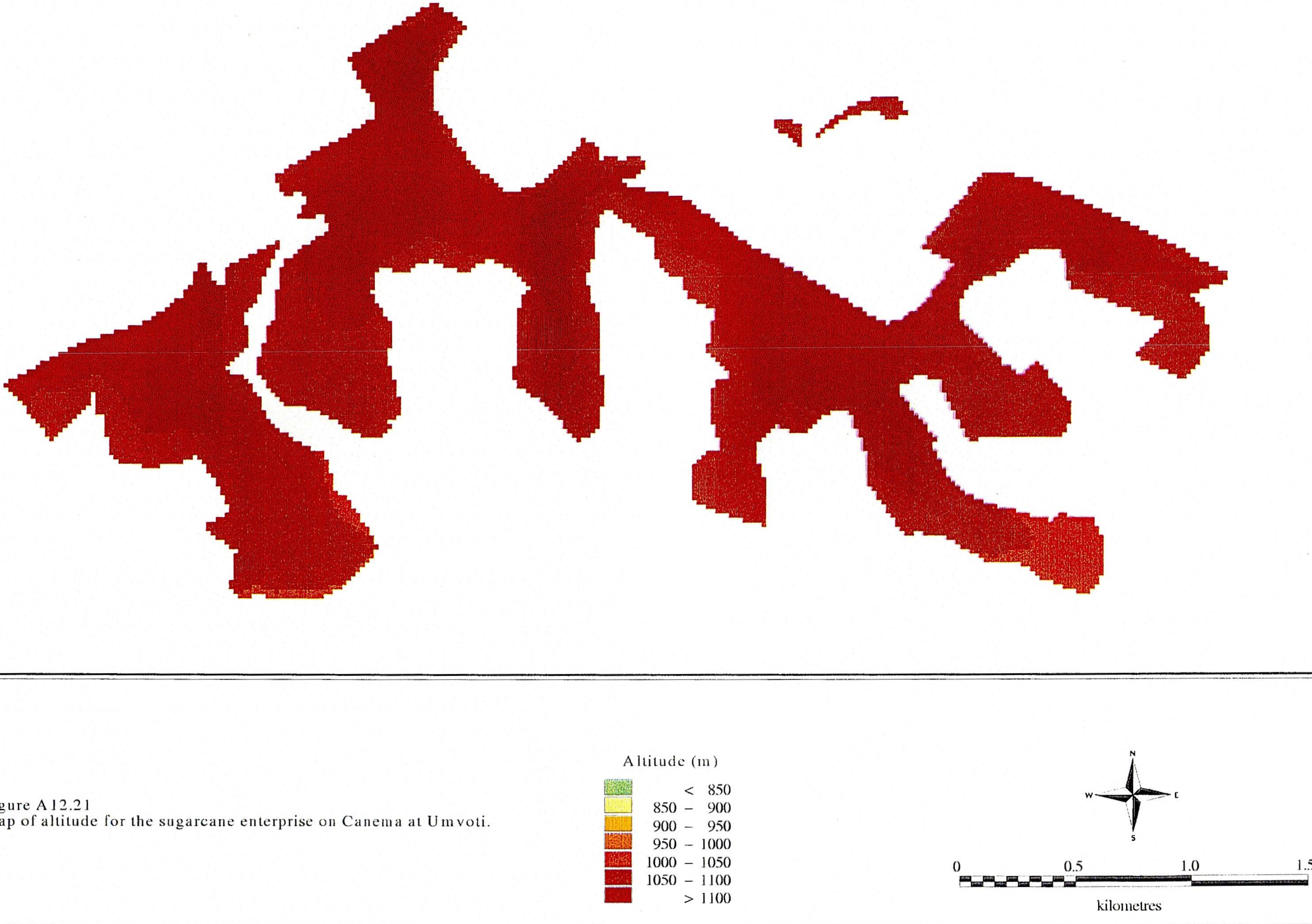


Figure A12.21
Map of altitude for the sugarcane enterprise on Canema at Umvoti.

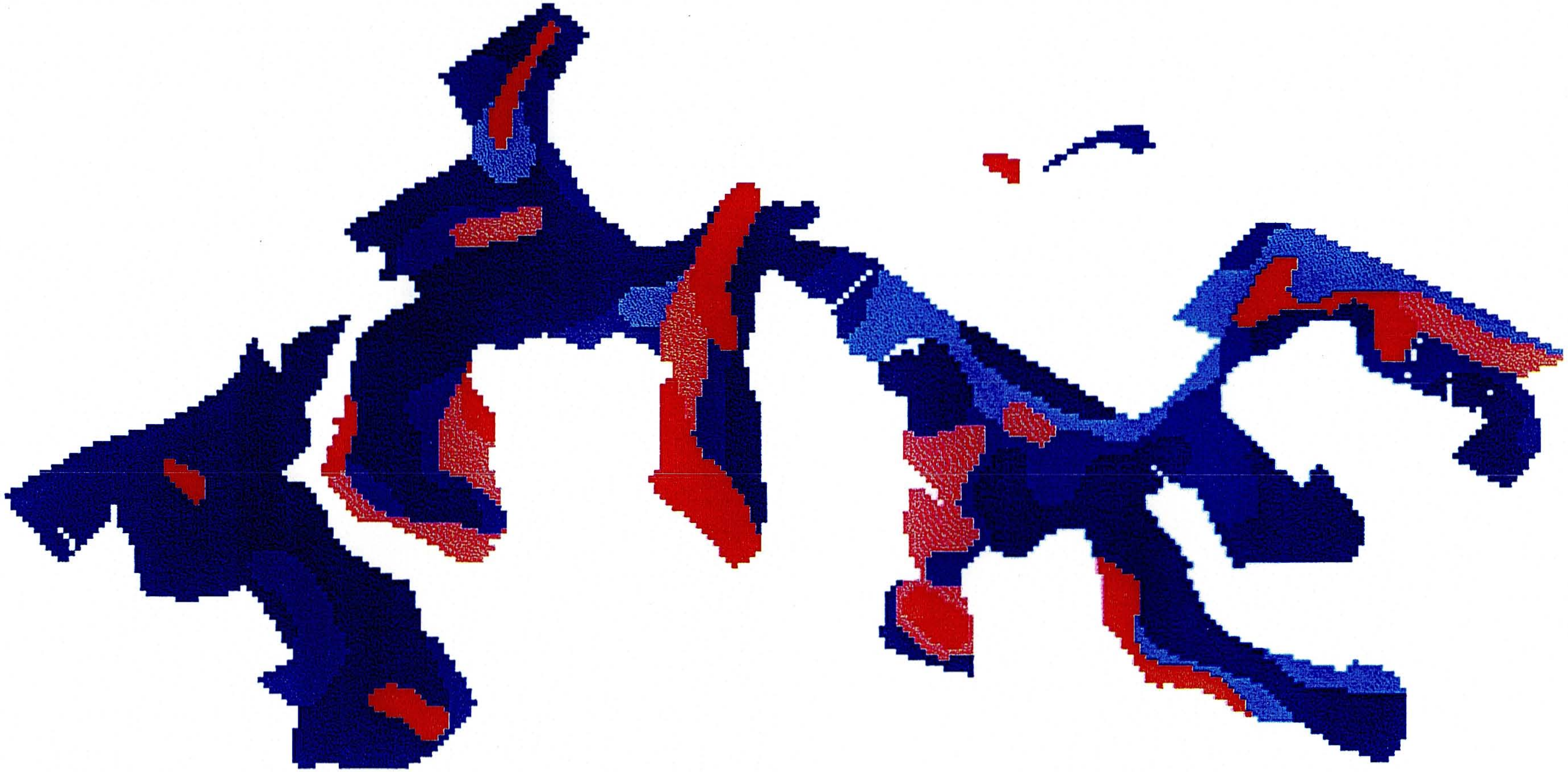


Figure A12.22
Map of effective rooting depth (ERD) for the sugarcane enterprise on Canema at Umvoti.

Figure A12.23
Map of soil types for the sugarcane enterprise on Uplands at Richmond.

Figure A12.24
Map of aspect for the sugarcane enterprise on Uplands at Richmond.

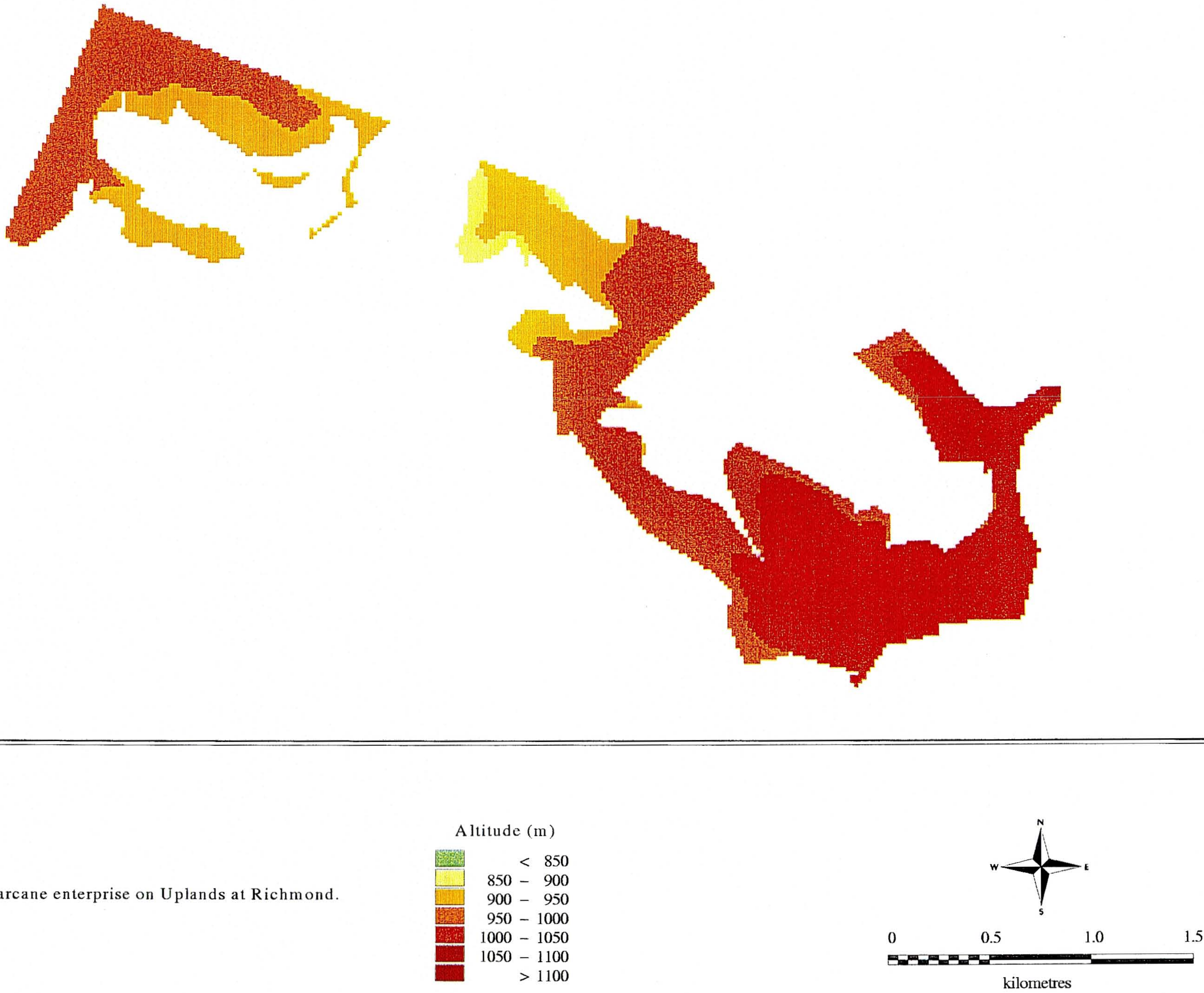


Figure A12.25
Map of altitude for the sugarcane enterprise on Uplands at Richmond.

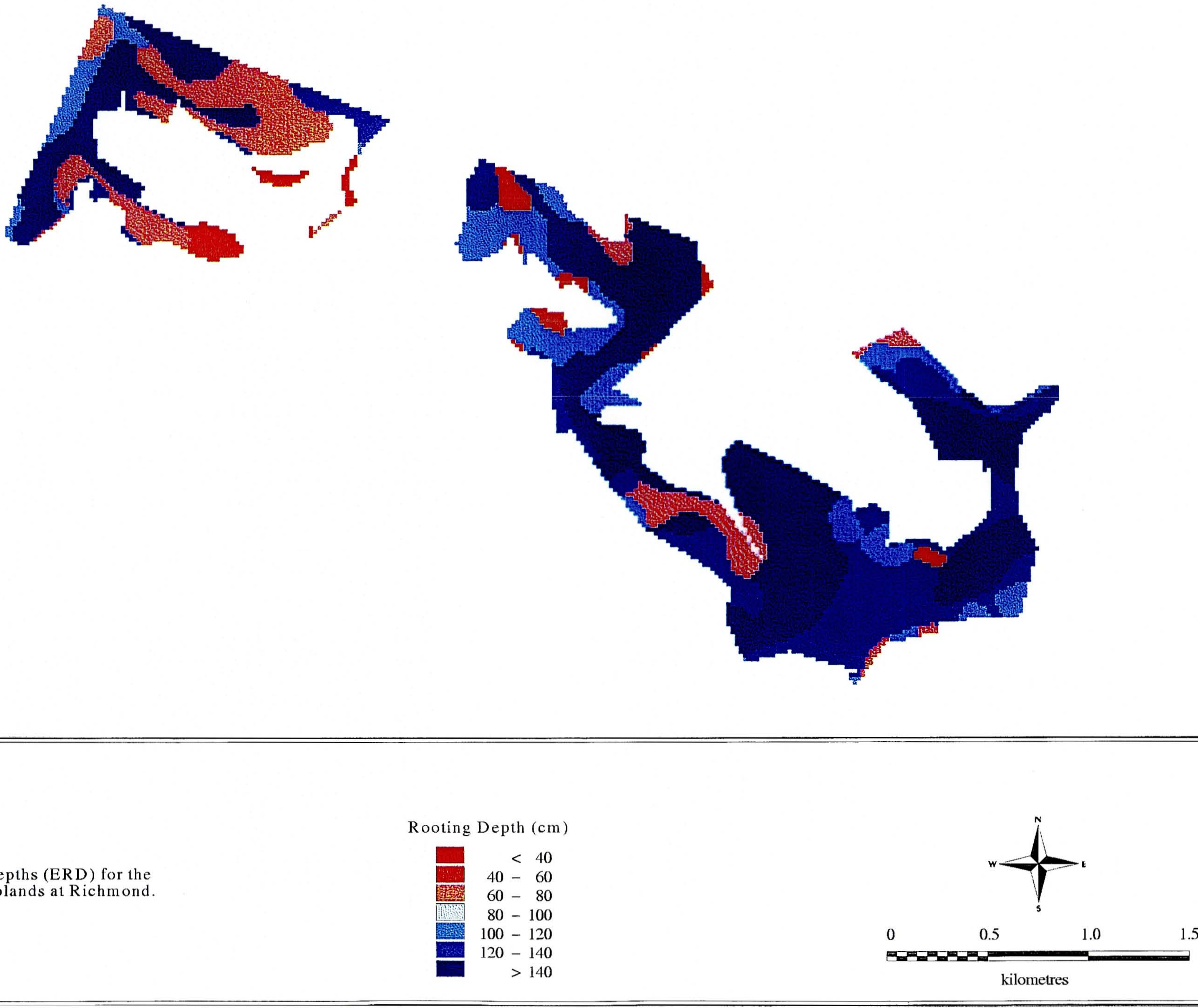


Figure A12.26
Map of effective rooting depths (ERD) for the sugarcane enterprise on Uplands at Richmond.

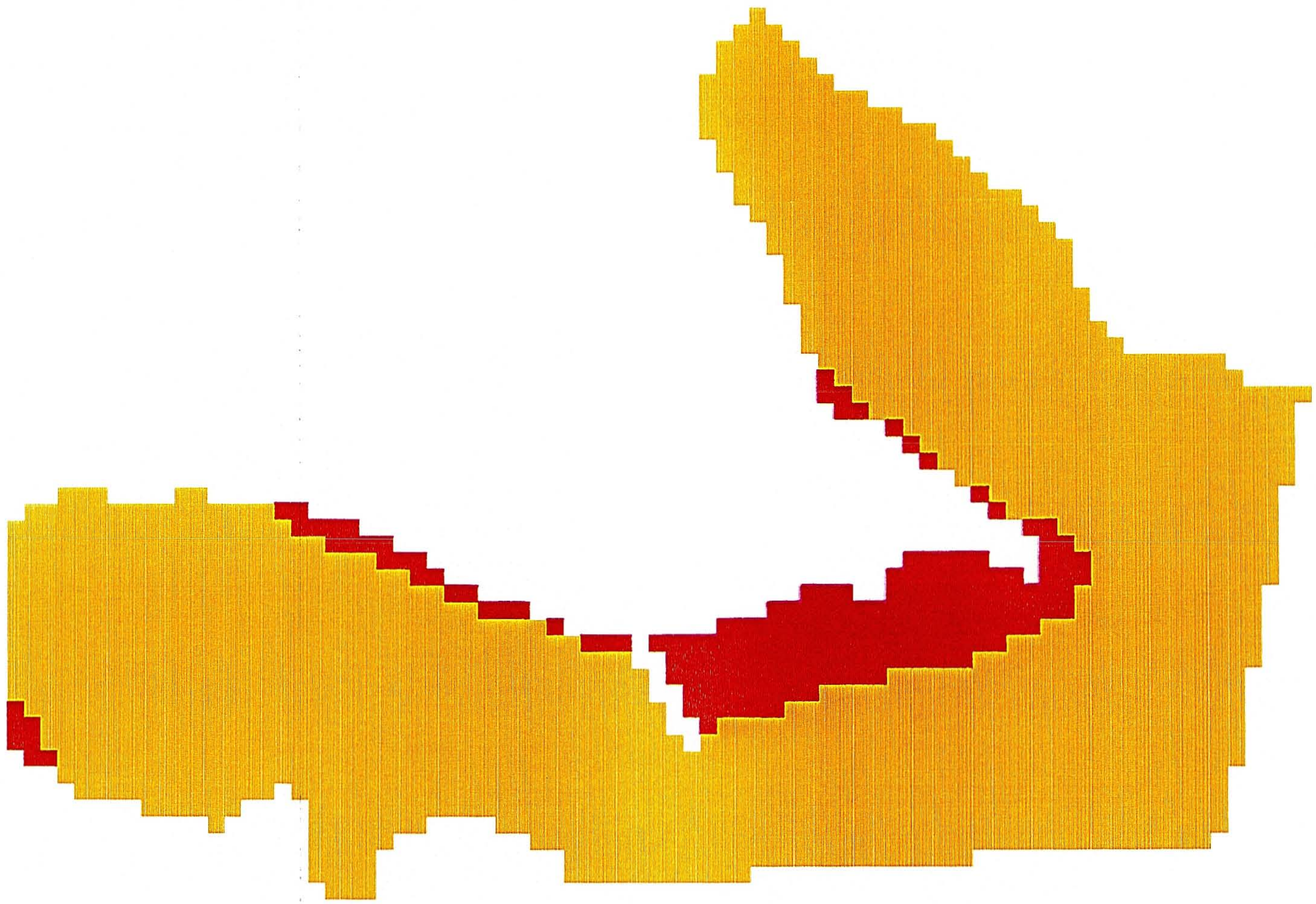


Figure A12.27
Map of soil types for the sugarcane enterprise on Greenhill at Richmond

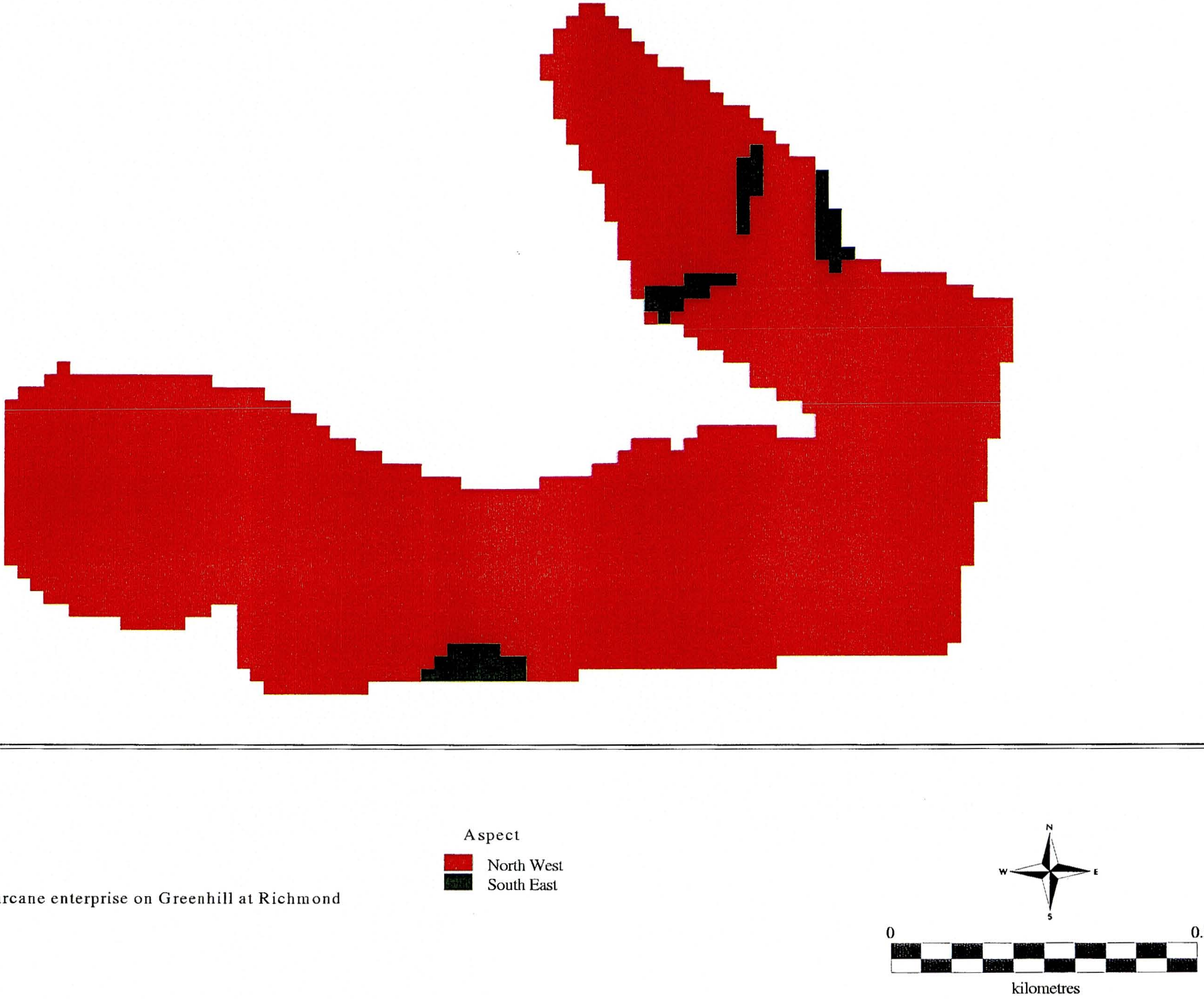


Figure A12.28
Map of aspect for the sugarcane enterprise on Greenhill at Richmond

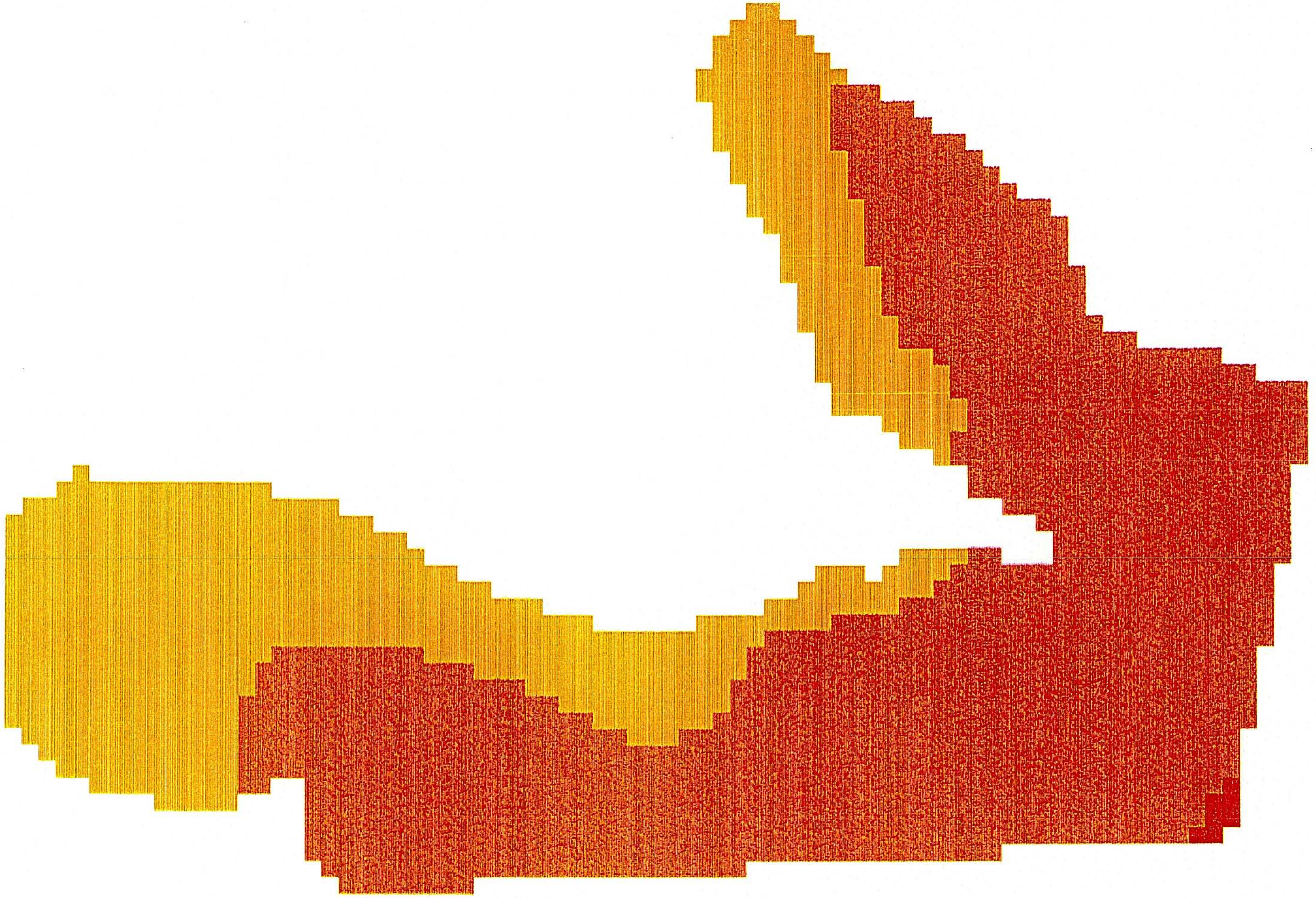


Figure A12.29
Map of altitude for the sugarcane enterprise on Greenhill at Richmond

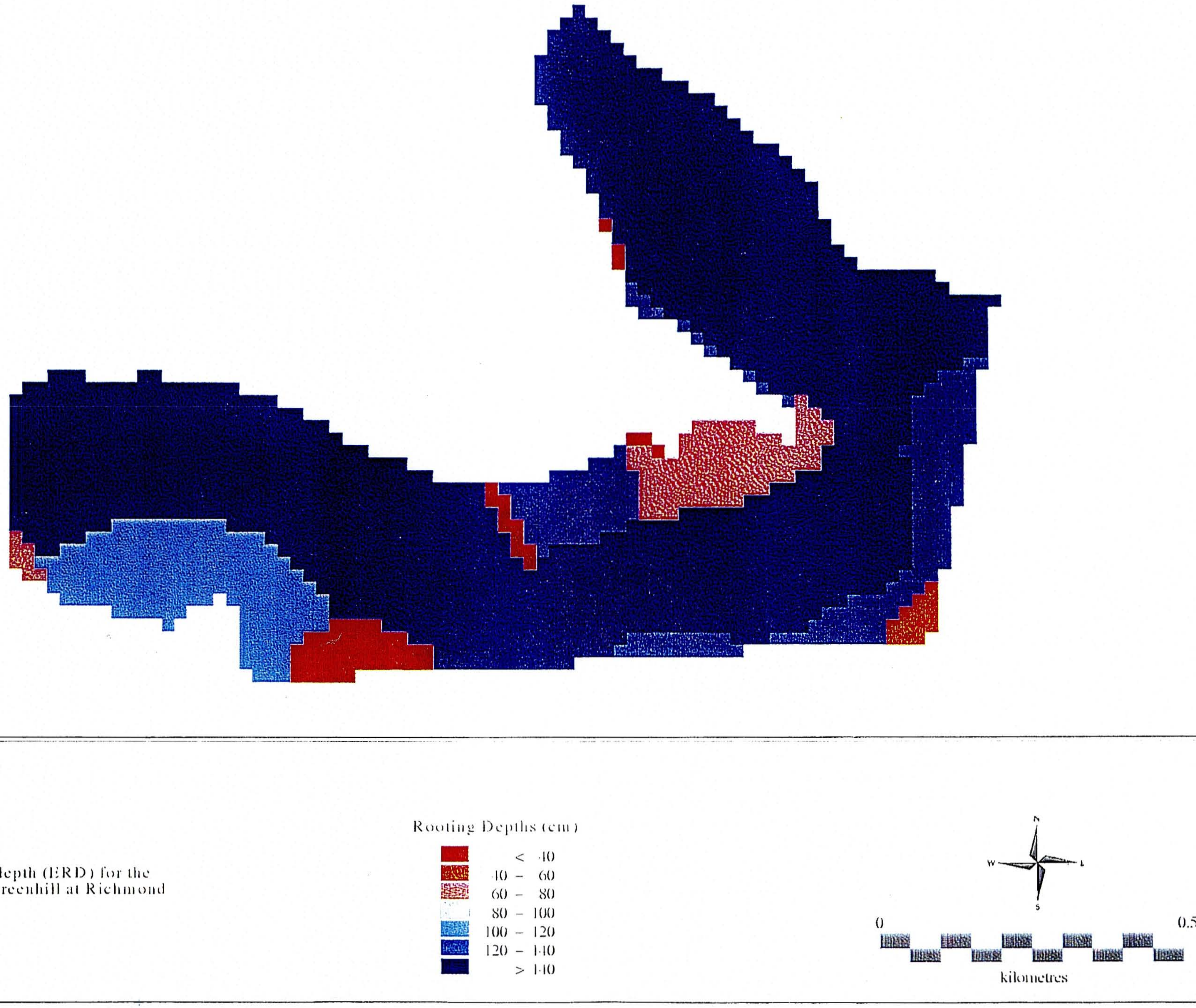


Figure A12.30
Map of effective rooting depth (ERD) for the sugarcane enterprise on Greenhill at Richmond

APPENDIX 13

FIELD RECORDS

TABLE OF CONTENTS

NOTE: These data are provided in an electronic format on the 3½″ diskette inserted in the back-cover. The files are saved in space-delimited ASCII format (.txt) which can be converted for use in most spreadsheet and database application programs. The data columns for the two files are arranged in the same order, as detailed in the LIST OF DATA below. Missing values are indicated by an asterisk (*).

LIST OF DATA

Column 1:	Observation number
Column 2:	Exclusion criterion (Section 4.1.1)
Column 3:	Checked data (Table 10)
Column 4:	GIS field number
Column 5:	District
Column 6:	Estate
Column 7:	Latitude (°S)
Column 8:	Longitude (°E)
Column 9:	TCH
Column 10:	TCHM
Column 11:	TSH (relative)
Column 12:	TSHM (relative)
Column 13:	Relative sucrose (%)
Column 14:	Harvest season

Column 15: Planting month
Column 16: Crop cycle starting month
Column 17: Crop cycle harvest month
Column 18: Age at harvest (months)
Column 19: Fallow period (months)
Column 20: Summer months growth
Column 21: Winter months growth
Column 22: Plant / ratoon status
Column 23: Crop number
Column 24: Crop group
Column 25: Crop cycle
Column 26: Variety
Column 27: Field area (ha)
Column 28: Field altitude (m a.s.l.)
Column 29: Field aspect (N / S)
Column 30: Field aspect (N / E / S / W)
Column 31: Mean field slope
Column 32: Field terrain unit
Column 33: Lithology
Column 34: Topsoil texture
Column 35: Dominant soil type
Column 36: Soil colour
Column 37: Soil humus (orthic / humic)
Column 38: Effective rooting depth (cm)
Column 39: Topsoil clay content (%)
Column 40: Subsoil clay content (%)
Column 41: Total available water (mm)
Column 42: A-horizon carbon content (1 = 1.4 – 1.8%; 2 = 1.8 – 2.0%)
Column 43: Profile gravel (%)
Column 44: Profile rocks (%)
Column 45: Trees along a field boundary

Column 46: Trees along lower field boundary

Column 47: Frost class

Column 48: N fertilizer (kg ha^{-1})

Column 49: P fertilizer (kg ha^{-1})

Column 50: K fertilizer (kg ha^{-1})

Column 51: Zn fertilizer (kg ha^{-1})

Column 52: Calmag (kg ha^{-1})

Column 53: Lime (t ha^{-1})

Column 54: Gypsum (t ha^{-1})

Column 55: Rain for 2 months before crop cycle start (mm)

Column 56: Rain for 12 months’ growth (mm)

Column 57: Rain total accumulated for crop cycle (mm)

Column 58: Rain for 1 month before harvest (mm)

Column 59: Leaf sampling month

Column 60: Age at leaf sampling

Column 61: Leaf N (%)

Column 62: Leaf N (% SASEX threshold)

Column 63: Leaf P (%)

Column 64: Leaf P (% SASEX threshold)

Column 65: Leaf K (%)

Column 66: Leaf K (% SASEX threshold)

Column 67: Leaf Ca (%)

Column 68: Leaf Ca (% SASEX threshold)

Column 69: Leaf Mg (%)

Column 70: Leaf Mg (% SASEX threshold)

Column 71: Leaf S (%)

Column 72: Leaf S (% SASEX threshold)

Column 73: Leaf Zn (%)

Column 74: Leaf Zn (% SASEX threshold)

Column 75: Leaf N:P (calculated from %)

Column 76: Leaf K:Mg (calculated from %)

Column 77: Leaf Ca:Mg (calculated from %)

Column 78: Topsoil analysis laboratory

Column 79: Topsoil volume density (g mL^{-1})

Column 80: Topsoil organic carbon (%)

Column 81: Topsoil mineralizing category

Column 82: Topsoil phosphorus desorption index

Column 83: Topsoil P (kg ha^{-1})

Column 84: Topsoil K (kg ha^{-1})

Column 85: Topsoil Ca (kg ha^{-1})

Column 86: Topsoil Mg (kg ha^{-1})

Column 87: Topsoil Zn (ppm)

Column 88: Topsoil Aℓ (ppm)

Column 89: Topsoil aluminium saturation index

Column 90: Topsoil cations (c mol$_c$ kg^{-1})

Column 91: Topsoil total cations (c mol$_c$ L^{-1})

Column 92: Topsoil exchangeable acidity (c mol$_c$ L^{-1})

Column 93: Topsoil acid saturation (%)

Column 94: Topsoil pH$_{(Water)}$

Column 95: Topsoil pH$_{(KCl)}$

Column 96: Topsoil K:Mg (calculated from kg ha^{-1})

Column 97: Topsoil K:Mg (calculated from c mol$_c$ kg^{-1})

Column 98: Topsoil Ca:Mg (calculated from kg ha^{-1})

Column 99: Topsoil Ca:Mg (calculated from c mol$_c$ kg^{-1})

Column 100: Total P (kg ha^{-1})

Column 101: Total K (kg ha^{-1})

Column 102: Total Zn (kg ha^{-1})

Column 103: Subsoil analysis laboratory

Column 104: Subsoil P (kg ha^{-1})

Column 105: Subsoil volume density (g mL^{-1})

Column 106: Subsoil K (kg ha^{-1})

Column 107: Subsoil Ca (kg ha^{-1})

Column 108: Subsoil Mg (kg ha^{-1})

Column 109: Subsoil Zn (ppm)

Column 110: Subsoil exchangeable acidity (c mol_c L^{-1})

Column 111: Subsoil total cations (c mol_c L^{-1})

Column 112: Subsoil acid saturation (%)

Column 113: Subsoil $pH_{(Water)}$

Column 114: Subsoil $pH_{(KCl)}$

Column 115: Diuron ® (kg ha^{-1})

Column 116: Paraquat ® (kg ha^{-1})

Column 117: Hexazinone ® (kg ha^{-1})