DEVELOPMENT AND EVALUATION OF A SUGARCANE YIELD FORECASTING SYSTEM

TREVOR GRAEME LUMSDEN

Submitted in partial fulfilment of the
requirements for the degree of
Master of Science

School of Bioresources Engineering and Environmental Hydrology
University of Natal
Pietermaritzburg

2000

ABSTRACT

There is a need in the South African sugar industry to investigate improved techniques for forecasting seasonal sugarcane yields. An accurate and timely forecast of seasonal cane yield is of great value to the industry, and could potentially allow for substantial economic savings to be made. Advances by climatologists have resulted in increasingly accurate and timely seasonal climate forecasts. These advances, coupled with the ongoing advances made in the field of crop yield simulation modelling, present the sugar industry with the possibility of obtaining improved cane yield forecasts. In particular, the lead time of these forecasts would be improved relative to traditional techniques. Other factors, such as the flexibility offered by simulation modelling in the representation of a variety of seasonal scenarios, would also contribute to the possibility of obtaining improved cane yield forecasts.

The potential of applying crop yield simulation models and seasonal rainfall forecasts in cane yield forecasting was assessed in this research project. The project was conducted in the form of a case study in the Eston Mill Supply Area. Two daily time step cane yield simulation models, namely the *ACRU*-Thompson and CANEGRO-DSSAT models, were initially evaluated to test their ability to accurately simulate historical yields given an observed rainfall record. The model found to be the more appropriate for yield forecasting at Eston, the *ACRU*-Thompson model, was then used to generate yield forecasts for a number of seasons, through the application of seasonal rainfall forecasts in the model. These rainfall forecasts had previously been translated into daily rainfall values for input into the model. The sugarcane yield forecasts were then evaluated against observed yields, as well as against forecasts generated by more traditional methods, these methods being represented by a simple rainfall model and Mill Group Board estimates.

Although the seasonal rainfall forecasts used in yield forecasting were found not to be particularly accurate, the proposed method provided more reliable cane yield forecasts, on average, than those using the traditional forecasting methods. A simple cost-benefit analysis indicated that the proposed method could potentially give rise to the greatest net economic benefits compared to the other methods. Recommendations are made for the practical implementation of such a method. Future areas of research are also identified.

PREFACE

I hereby certify that the research presented in this dissertation is my own original and unaided work except where specific acknowledgement is made.

Signed: ______________

T.G. Lumsden

ACKNOWLEDGEMENTS

The author gratefully acknowledges the following individuals and organisations for their assistance in this research:

- Prof R.E. Schulze, Professor of Hydrology, School of Bioresources Engineering and Environmental Hydrology, University of Natal, for his enthusiastic supervision of this research, and for providing encouragement while the work was being completed.

- Mr N.L. Lecler, formerly of the School of Bioresources Engineering and Environmental Hydrology, University of Natal, for his active and helpful co-supervision of this research. He is also acknowledged for developing the methodology evaluated in the research relating to the translation of categorical rainfall forecasts into daily model input, along with the computer programs required to implement it.

- The SA Sugar Association Experiment Station, who initiated and funded a collaborative research project with the University of Natal. This dissertation is a product of the research project. A variety of data relating to the Eston Mill Supply Area including climate, soils, and observed cane yield data were supplied by SASEX.

- Mr E.J. Schmidt, Head of the Department of Agricultural Engineering, SASEX, for playing an active role in ensuring the applicability of the research to industry.

- Mr D. Hellmann, Midlands South Extension Officer, SASEX, for making available his local knowledge of sugarcane agriculture in the Eston Mill Supply Area.

- The steering committee of the collaborative research project for their involvement in directing the research. The committee comprised of:
  - Mr E.J. Schmidt (chairman), Dr A. Singels and Mr D. Hellmann of SASEX
  - Mr D. Cousens and Mr R. Thompson of Illovo Sugar
  - Mr G. Narciso of the Institute of Soil, Climate and Water, ARC
  - Mr R. Crausaz of the Midlands South Planters' Association
  - Prof P.W.L. Lyne of the University of Natal.

- Dr G.A. Kiker of the School of Bioresources Engineering and Environmental Hydrology, University of Natal, for guidance in the application of the CANEGRO-DSSAT model.

- The National Research Foundation for partial funding of the research.

- The Water Research Commission for partial funding of the research.

- Mr W. Landman and Mrs A. Bartman, South African Weather Bureau, for supplying seasonal rainfall forecasts specific to the area of study.

- The Institute of Soil, Climate and Water, ARC, for providing soil and climate information.

- The Computing Centre for Water Research for providing climate data and computing facilities.

- Illovo Sugar for making available a copy of the Simple Rainfall Model and for discussions relating to the Mill Group Board forecasts.

- Messrs Q. Mann, J. Meyer and D. Hellmann of SASEX, for developing the working rules used to predict soil type based on soil parent material and other information.

- Mr Q. Hildebrand, SA Canegrowers Association, for providing output from the LOMS model used in the cost-benefit analysis.

- Messrs M. Horan and B. Howe and Mrs L. Hallowes of the School of Bioresources Engineering and Environmental Hydrology, University of Natal, for assistance with GIS.

- Mrs M. Maharaj of the School of Bioresources Engineering and Environmental Hydrology, University of Natal, for performing temperature patching.

- My parents, Mr and Mrs I. Lumsden, for their support and encouragement throughout my years of studying.

- My wife, Zena, for her patience and support while I completed this dissertation.

It is the glory of God to conceal a matter;
to search out a matter is the glory of kings.

Proverbs 25 vs 2

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF APPENDICES

1 INTRODUCTION

There is a need in the South African sugar industry to investigate improved techniques for forecasting seasonal sugarcane yields. An accurate and timely forecast of seasonal cane yield is of great value to the industry, and has the potential to result in savings of many millions of Rands annually. Traditional forecasting techniques are successful to a degree, but are generally simple in nature, with their success frequently depending on the experience of those involved. Advances by climatologists have resulted in increasingly accurate and timely seasonal climate forecasts (Hammer, Holzworth and Stone, 1996a). These advances, coupled with the ongoing advances made in the field of crop yield simulation modelling, present the sugar industry with the possibility of obtaining improved cane yield forecasts. In particular, the lead time of these forecasts would be improved relative to traditional techniques. Other factors, such as the flexibility offered by simulation modelling in the representation of a variety of seasonal scenarios, would also contribute to the possibility of obtaining improved cane yield forecasts.

The South African Sugar Association Experiment Station (SASEX) funded a research project to investigate whether improved sugarcane yield forecasts could be derived from seasonal climate forecasts and crop yield simulation modelling. This dissertation is based upon the research conducted in this project. It was decided that the project would focus on a case study conducted at the scale of a mill supply area. The Eston Mill Supply Area was selected for this purpose because of the availability of good observed cane yield, soils and climate data. The main objective of the project can be stated as:

The development and evaluation of a sugarcane yield forecasting system for a mill supply area, using crop yield simulation modelling and seasonal climate forecasts.

In order to achieve this objective, elements of research focussed on:

- an evaluation of two sugarcane yield simulation models of differing complexity, making use of historical climate and yield data, in order that the ability of the models to accurately predict historical yields be verified, thus leading to the identification of a model suitable for application in cane yield forecasting;

- the assembling and assessment of the inputs required by the above models at their scale of application, including inputs related to climate and soils, with a number of available sources of information being investigated in order that the most appropriate source be identified;
- an investigation of the sugarcane growth cycles practised in the selected mill supply area and the formation of a strategy to represent these in the models, given that these cycles are an important and influential form of management;
- an evaluation of the seasonal climate forecasts (relating to rainfall) used in yield forecasting to assess their accuracy;
- the application and evaluation of a methodology developed by Lecler (described in Lumsden, Schulze, Lecler and Schmidt, 1999) to translate the seasonal rainfall forecasts into a form suitable for application in the selected yield simulation model;
- a comparison of the historical yield forecasts generated by the proposed yield forecasting system against observed yield data, and against forecasts derived from traditionally employed methods;
- a simple benefit analysis of the yield forecasting system to determine whether benefits can be derived from its use; and,
- an assessment of the range of possible outcomes associated with a forecast generated by this system, thus indicating the level of risk associated with decisions made on the basis of the forecasts.

Based on the findings of the research, recommendations regarding practical application of the system in the sugar industry and the direction of future research, were also made.

The yield simulation models evaluated for possible application in yield forecasting, included the *ACRU*-Thompson model (Schulze, Domleo, Furniss and Lecler, 1995) as modified by Lumsden, Lecler and Schulze (1998) and the CANEGRO-DSSAT model (Inman-Bamber and Kiker, 1997). The traditional yield forecasting methods against which the selected yield simulation model was compared, included a simple rainfall model and the estimates of the local Mill Group Board.

The Simple Rainfall Model (SRM) was applied in the initial model verification phase of the research to assess its ability (along with the *ACRU*-Thompson and CANEGRO-DSSAT models) to predict historical yields given an observed rainfall record. The SRM was developed by Illovo

Sugar Limited (Cousens, 1998) for application in the midlands of KwaZulu-Natal. It is representative of the level of model applied in the South African sugar industry for practical yield prediction at Mill Supply Area (MSA) scale, and represents a base against which more complex yield simulation models can be compared. The model is applied in this project in a manner consistent with that applied in practice.

Mill Group Board (MGB) yield forecasts are based on surveys of growers' expectations for a season. The yields forecasted by the growers are reviewed by the MGB. If deemed necessary, the forecasts are adjusted to be more representative of the yields anticipated by the MGB for the prevailing conditions.

The *ACRU*-Thompson and CANEGRO-DSSAT yield models represent, respectively, intermediate and higher levels of model complexity. The range in model complexities has implications in terms of the effort required to set up and operate the models, as well as in their potential to provide other useful information, such as the response of a crop to varying management strategies. The development of spatial databases of model inputs was considered a necessary and valuable process in the setting up and operating of the models for prediction and interpretation of the MSA yields.

The rainfall forecasts used in this research were obtained from the South African Weather Bureau (SAWB). These rainfall forecasts were categorical in nature, in that they indicated whether the forthcoming months' rainfall was forecasted to be above normal, near normal or below normal. This type of forecast is typical of seasonal rainfall forecasts, and requires a methodology to be developed to "translate" the forecasts into a form suitable for input into the yield models.

In this dissertation cane yields are defined as the mass of stalks (at field moisture content) per hectare. This is distinct from sucrose yield, which is mass of extracted sucrose per hectare. Although the end value of a crop is based on the sucrose yield, the cane yield is important in terms of many functions such as mill operations planning, harvest and haulage scheduling and crop management, all of which would benefit from advance estimation of the seasonal yield.

The potential benefits of forecasting sugarcane yields are discussed in Chapter 2 of this dissertation. In Chapter 3 a literature review is presented of the techniques available for forecasting crop yields. Descriptions of the models applied in the research and the area selected for study are given in Chapters 4 and 5 respectively. This is followed in Chapter 6 by the verification of outputs from the sugarcane yield models using observed climate and yield data. The application of the yield forecasting methodology and the results thereof are presented in Chapter 7. These results are compared with those of the traditional yield forecasting methods. Thereafter, Chapters 8, 9 and 10 relate to discussion, conclusions and recommendations.

2 POTENTIAL BENEFITS OF FORECASTING SUGARCANE YIELDS

In the sugar industry there are numerous benefits that can be derived from having accurate and timely forecasts of seasonal sugarcane yield. Such forecasts can potentially be applied in decision-making at national, MSA and individual grower scale. At the national scale, forecasts could be used in the development of marketing and pricing strategies, in the early signing of export contracts and in the provision of forward cover for exchange rate fluctuations. At the MSA scale forecasts could be applied in the planning of mill operations such as the determination of opening and closing dates, haulage scheduling and in the determination of crushing and extraction rates. At grower scale, crop forecasts could be used in decisions relating to cash flows, in the planning of harvest and haulage scheduling and in crop husbandry decisions such as fertilizer applications and irrigation scheduling (Schmidt, 1998).

As an illustration of the potential benefits of having accurate and timely cane yield forecasts, an analysis of the economic implications of selecting the length of the milling season was obtained from Hildebrand (1998a), who applied the Length of Milling Season (LOMS) model (Hildebrand, 1998b). The analysis, which was conducted for the Noodsberg Mill in the KwaZulu-Natal midlands, considered the effect of varying the mill opening date when crushing a 1.5 million ton cane crop. The LOMS model predicted that the optimum length of the milling season was between 12 April and 24 December. If the crop was overestimated by 2.8% and the mill was opened a week early, the model predicted a reduction in profit of R128 000 for the area. This reduction in profits increased to R807 000 for an opening date four weeks early (crop overestimated by 11.2%). If the crop was underestimated by 2.8% and the mill was opened a week late, reductions in profit of R566 000 were predicted, with this loss increasing to R1 937 000 for an opening date four weeks late (crop underestimated by 11.2%). Reductions in profit were ascribed to poorer cane quality, less favourable ratooning and increased growing and milling costs at certain times of the year. The above analysis did not account for benefits such as those that could have been derived from improved marketing, pricing and export strategies, were accurate crop forecasts available.

Potential benefits of forecasting sugarcane yields have been discussed in this chapter. A literature review of the techniques that are available for forecasting crop yields, follows in the next chapter.

3 TECHNIQUES AVAILABLE FOR CROP YIELD FORECASTING

A review was conducted of the techniques available for crop yield forecasting. This review was conducted in order to gain an understanding of these techniques, including an appreciation of their advantages and disadvantages, so that the selected technique, namely simulation modelling, could be viewed within the context of the field of crop yield forecasting. The review was not restricted to techniques used to forecast sugarcane yields; however, all the reviewed techniques were considered to have potential for sugarcane crop forecasting.

Techniques that can potentially be used for forecasting yields include rules of thumb, neural networks, statistical modelling, remote sensing and simulation modelling (Wisiol, 1987; Uhrig, Engel and Baker, 1992). Statistical modelling techniques include stochastic Markov chain models as well as time series analysis models and uni- or multivariate regression models. Rules of thumb, although widely used, are simplistic and specific to their location of development. They furthermore require time and local experience to develop (Wisiol, 1987). Neural network and Markov chain models are possibly useful yield forecasting techniques, however, there are few examples at present in literature of their application in this context. Time series analysis models require large data sets in order to successfully capture trends useful for yield forecasting (Wisiol, 1987).

Regression modelling is the most extensively applied technique up to the present and contains many attractive features to the crop forecaster (Horie, Yajima and Nakagawa, 1992). Remote sensing has unique advantages such as direct and crop specific monitoring, which is possible over large areas and at good spatial and temporal resolution, thus making it a very promising technique (Dubey, Ajwani, Kalubarme, Sridhar, Navalgund, Mahey, Sidhu, Jhorar, Cheema and Narang, 1994). Simulation modelling has the ability to incorporate management practices, climate, soils and crop type. This ability, combined with the advent of increasingly accurate mid to long range seasonal climate forecasts, make simulation modelling an attractive forecasting option (Hammer *et al.*, 1996a).

For reasons discussed above, the techniques of regression modelling, remote sensing and simulation modelling were selected for review in this chapter. An emphasis was placed on simulation modelling, as this technique was utilized in the research presented in this dissertation.

3.1 Regression Modelling

Regression modelling is a simple, yet often effective technique, of relating one or more determining factors to the final yield of a crop (Wisiol, 1987). Data relating to observed yields and yield determining factors are collected, and relationships between the two determined. This allows for predictions of yields to be made when no observed yields are available. For the purposes of real-time forecasting, the variables chosen are those that can easily be obtained and related to the final yield as early as possible in the season, and with the highest level of predictive ability. Regression analysis is a widely used technique in modelling, but requires long, good quality data sets for the greatest success. Since regression models are based on collected data, they reflect the response of the crop occurring in that specific area. As a result the application of a model in a different area need not necessarily give good results (Horie *et al.*, 1992). A range of variable types has been used to forecast crop yields and they are generally based either on climate (Stephens, Walker and Lyons, 1994), predicted climate indicators (Rimmington and Nicholls, 1993) or on observations of the crop during the growing season (Horie *et al.*, 1992).

Models based on observations of the crop are effective and reliable as they reflect the influences of the growing environment on the crop up to the point of forecast (Horie *et al.*, 1992). Periodic measurements are made of crop characteristics and evidence of disease, insects or environmental stresses noted. These measurements are made on random samples. The relationship that is established between the attributes of a crop and the final yield can then be used in subsequent years to forecast yield. This approach can be problematic, however, as the data collection is labour intensive and must be carried out meticulously, rendering this approach unsuitable for large areas (Singh and Bapat, 1988).

Yield regression models that are based on climate will typically relate climate at various stages within a crop's growth to the final yield. If the growth stage of a crop is particularly sensitive to the climate prevailing at that time, then weighting factors can be used to make the model more sensitive to climate during that stage, thus rendering potentially greater accuracy in results. Minimum thresholds can also be used, for instance, to eliminate rainfall or temperature that does not contribute to growth (Stephens *et al.*, 1994; Durling, Hesterman and Rotz 1995). If a regression model requires the entire season's climate to be known, then average climate can be assumed for the period between the forecast and harvest dates. This implies that as the season progresses, yield forecasts should become more and more accurate as measurements of rainfall become available.

In addition to using observed climate data, yield regressions have also been developed using pre-season indicators of climate. The Southern Oscillation Index (SOI) is an example of such an indicator. This indicator is related to the El Niño and Southern Oscillation (ENSO) phenomenon which is associated with changes in ocean temperature conditions across the Pacific Ocean and the consequent atmospheric circulation (McBride and Nicholls, 1983). The extent of influence of the ENSO phenomenon has been shown to extend to various regions around the world through climate links known as teleconnections. There are two extremes of the ENSO phenomenon which are generally associated with either favourable or unfavourable rainfall conditions depending on the region concerned. Normal rainfall conditions generally exist if the ENSO signal is weak, providing no other phenomena influence the rainfall. Another indicator of the ENSO phenomenon that can be used in yield regressions is sea surface temperature (SST) in the eastern equatorial Pacific Ocean, as well as the SST of other oceans (Cane, Eshel and Buckland, 1994). Rimmington and Nicholls (1993) found that trends of the seasonal and interannual SOI were more useful than the actual values. Complex ocean-atmospheric circulation models can be used to predict the behaviour of ENSO in advance, thus allowing for longer lead times of yield regression forecasts (Cane *et al.,* 1994). Other indicators of the ENSO phenomenon are also available for yield regressions, such as those relating to wind circulations and sunspot cycles (Kuhnel, 1993).

Regression models are relatively simple models that can yield good results. They are, however, not without limitations, particularly in regard to the use of observed historical yields. Historical crop yields are often unreliable especially within the small farm sector (Cane *et al.,* 1994). They are also subject to socio-economic, technological and management influences which distort the relationship between yields and climate (Martin, Washington and Downing, 2000). There is also the difficulty of applying regression models in circumstances not consistent with their development, as they are static in nature and cannot adapt to a new environment.

3.2 Remote Sensing

Remote sensing (RS), in the context of crop monitoring, involves the remote measurement of spectral reflectances, thermal radiations or other electromagnetic spectra from crop fields. The sensors used for these measurements are mounted on craft such as satellites or aeroplanes (Wisiol, 1987; Horie *et al.*, 1992). Early application of remote sensing focussed on the determination of area under crops (MacDonald and Hall, 1980), but since then has expanded to the forecasting of yields (Wisiol, 1987; Horie *et al.*, 1992). RS approaches to yield forecasting allow for direct, albeit remote, measurements of the condition of a crop and reflect aspects of growth such as leaf area development, photosynthesis processes and plant stress owing to water deficits, pests or disease (Wisiol, 1987). They also reflect the effect of crop management on growth (Maselli, Conese, Petkov, and Gilabert, 1993). Measurements of crops are often used to calculate vegetation indices such as the Normalized Difference Vegetation Index (NDVI) which are then related to yield through regression relationships (King and Meyer-Roux, 1990; Sridhar, Dadhwal, Chaudhari, Sharma, Bairagi and Sharma, 1994; Smith, Adams, Stephens and Hick, 1995). Vegetation indices have also been related to yield on a more physiologically explicit basis through relationships that account for solar radiation interception by the canopy (Jaggard and Clark, 1990; Wiegand and Richardson, 1990). Measurements of canopy temperature using RS have been used to calculate stress indices, which are then related to yield (Gardner, Blad, Garrity and Watts, 1981; Jackson, Idso, Reginato and Pinter, 1981).

For the purposes of crop yield forecasting, most RS is carried out from satellites. Two varieties of satellites are available for crop monitoring, these being those that produce high resolution images with low temporal frequency and those producing low resolution images with high temporal frequency (King and Meyer-Roux, 1990). The more frequently available, lower resolution images are more popular for yield forecasting, as cloud cover often obscures remotely sensed images, resulting in the need for frequent capture of images to ensure representation of critical growth stages (Smith *et al.*, 1995). This has implications for ground referencing and verification, as the coarser spatial resolution makes it difficult to identify features on the ground. A number of other image processing procedures are necessary before yield forecasting can commence (Maselli *et al.*, 1993; Dubey *et al.*, 1994; Smith *et al.*, 1995). These procedures can become prohibitively costly in crop yield forecasting where data must be readily available to facilitate real-time estimates of yield.

Three techniques are commonly used to relate RS acquired data to crop yields. These techniques include vegetation index models, radiation interception models and canopy temperature models.

3.2.1 Vegetation index models

Many vegetation index models are functionally equivalent, with the NDVI generally being the most used index when analysing RS data (King and Meyer-Roux, 1990). The index is based on measurements of solar irradiance in the red and near infrared wavebands, where the red waveband is strongly absorbed by chlorophyll and the near infrared band is scattered by leaf tissue. The index is thus strongly correlated with green biomass, intercepted radiation and water use (Smith *et al.*, 1995). The NDVI follows a trend during the season with the index increasing in value as leaf area develops and then decreasing as the crop passes through senescence. When a regression relationship is formed between the NDVI and observed yields, several forms of the index may be used, including the value of the index at specific points in the season, or averages or increments of the index over certain periods within the season.

3.2.2 Radiation interception models

Vegetation indices have been found to be good predictors of final yield when used in regression relationships. However, they are location specific and could possibly be used more effectively if more physiologically based relationships with yield were to be found (Horie *et al.*, 1992). An example of how remotely sensed vegetation indices can be applied in more physiologically based relationships to predict crop yields is through radiation interception models. Vegetation indices are directly related to radiation interception as they reflect the area of foliage available for interception (Wiegand and Richardson, 1990). Since crop biomass at a given point in the growth cycle of a crop is proportional to crop intercepted radiation accumulated up to that time, crop models can be developed which relate crop yield to intercepted (photosynthetically active) radiation based on inputs from a vegetation index and measurements of radiation from meteorological stations (Horie *et al.*, 1992). Radiation interception models are generally based on the assumption of a constant radiation to plant material conversion factor. Another assumption that is made is that the relationship between total aboveground biomass and crop yield is constant (Horie *et al.*, 1992). The above assumptions require that there be no plant stress owing to water or temperature influences or as a result of nutrient deficiencies. These factors also complicate the interpretation of vegetation index models, and as a result the applicability of vegetation index and radiation interception models can be somewhat limited.

3.2.3 Canopy temperature models

The limitations of vegetation index and radiation interception models are, to some extent, overcome by canopy temperature models. These models are based on the principle that a freely transpiring crop not subjected to water stress will have a canopy temperature a few degrees cooler than the surrounding air, while a stressed crop will have a canopy temperature a few degrees warmer (Idso, Jackson and Reginato, 1977). Several crop stress indices incorporating remotely sensed canopy temperature have been developed to reflect the water status of a crop (Idso *et al.*, 1977; Jackson *et al.*, 1981; Gardner *et al.*, 1981). Crop water status is known to be related to yield and thus remotely sensed canopy temperatures can be used for yield prediction (Idso *et al.*, 1977). Early canopy temperature models relied on vapour pressure deficits being relatively constant, however this limitation was later overcome (Jackson *et al.,* 1981). A further development of the models was in the simplification of input, through consideration of the relative canopy temperatures of the field of interest and an unstressed field nearby. It is believed that stress indices can be applied to yield forecasting, although it is recognized that they only become useful once the canopy is closed, as non-crop temperature effects influence remotely sensed temperature readings. This results in shorter lead times (Horie *et al.,* 1992). If final yield is limited by slow leaf area development in the early phases of growth, then stress indices will not reflect this in yield forecasts (Horie *et al.,* 1992).

3.3 Simulation Modelling

Crop simulation models can be broadly classified as being either crop yield models or crop growth models (Schulze *et al.,* 1995). Yield models only simulate the final yield of a crop whereas growth models simulate the development of the crop and give outputs that relate to various aspects of growth, including the final yield. Growth models are generally complex models while yield models are more simplified. The two forms of simulation models and their application in crop yield forecasting will be reviewed in the following sections.

3.3.1 Yield models

Agroclimatological Yield (ACY) models, also known as crop weather analysis models, are a form of yield model which has been identified as being appropriate for application in crop yield forecasting (Baier, 1979). Most examples found in available literature of the application of yield simulation models in crop forecasting, have involved the use of ACY models. Agrohydrological Yield (AHY) models are another form of yield simulation model which can be used in crop forecasting. An example of the application of an AHY model in yield forecasting, is that of the *ACRU* maize model (Lecler, Schulze and Pike, 1996). The yield simulation model selected for yield forecasting in this project is an AHY model, and is a modification of the sugarcane yield model presented in Schulze *et al.* (1995). ACY and AHY models differ in that yield is related more to climate in the case of ACY models than AHY models, in which a soil water budget forms the heart of the model. ACY models make use of a simple index of soil moisture or, alternatively, a simplified water budget of coarse temporal resolution. In contrast, the *ACRU* AHY models (maize and sugarcane) operate using a detailed water budget operating at a daily time scale. The review of yield simulation models in this chapter will cover both ACY and AHY models.

3.3.1.1 Agroclimatological yield models

ACY models simulate the accumulated responses of a crop to prevailing agroclimatological conditions during the growing season. Crop response is represented through the use of an agroclimatological index which is based on climate data and a simple representation of soil moisture (Motha and Heddinghaus, 1986; Kalma, Lyons, Nunez and Pitman, 1991). Agroclimatological indices are usually calculated at intervals of ten days, and changes in value of the indices from one interval to the next reflects the crop response to growing conditions during that interval. The final accumulated index value is then used to make a yield prediction for the season, typically through the use of a previously developed regression relationship between the index and observed historical yields. In order to use ACY models for real-time forecasts, assumptions regarding the climate must be made for the period between the forecast date and the end of the season. This is so in order for the model to be run using a complete climate data set. As the season progresses, more and more observed data become available, thus reducing the period over which assumed data must be used. As a result the yield forecasts become more and more accurate (Walker, 1989).

ACY models are effective in areas where climate and interannual climate variability dominate crop yield response. This is as a result of their emphasis on crop-climate relationships. Ritchie (1983) states that simple models can be powerful predictors if one or two major factors dominate the crop environment. The agroclimatological indices have been designed to reflect conditions such as water stress (Stephens, Lyons and Lamond, 1989) and drought (Walker, 1989). Examples of ACY models are those of Stephens *et al.* (1989) and Walker (1989) as well as those of Motha and Heddinghaus (1986) and Meyer, Hubbard and Wilhite (1993). High correlations have been found between the final index value of ACY models and the corresponding observed yields (Stephens *et al.,* 1989; Walker, 1989; Hammer, Stephens and Butler, 1996b). A feature of ACY models is their relative simplicity, which allows for ease of use in the operational environment. It is ensured that the input data required are kept to a minimum and are readily available (Walker, 1989). The phenology of crops is taken into account through the use of concepts such as crop coefficients and weighting factors, as well as through the computation of growing degree days which are used to determine the timing of growth stages and maturity dates (Motha and Heddinghaus, 1986).

Various techniques are employed to complete the seasonal climate data record for the period between date of forecast and harvest. A conservative technique is to use monthly medians or averages of historical data (Walker, 1989). Another technique involves using historical climate data from as many years as are available. Each year of climate data is used to complete the data record, thus giving rise to multiple model outputs which can then be used to create a probability distribution of likely outcomes (Meyer *et al.*, 1993). A level of probability can be selected and the corresponding yield determined. A variation of this approach is to select a good, average and poor year in the historical climate data record and then use data from those years to complete the seasonal record. This would give an idea of the extremes of yield that are possible. A third technique available is to use seasonal climate forecasts (Stephens *et al.*, 1989). These climate forecasts are commonly based on the ENSO phenomenon and usually include predictions of rainfall and temperature. The forecasts are typically categorical in nature, the categories often being given as above normal, near normal or below normal. When applying climate forecasts, techniques must be sought to translate the categories into appropriate model data sets. The use of seasonal climate forecasts will be discussed further in the sections on AHY models and growth models.

The use of monthly averages to complete the climate data record, as was done in the study by Walker (1989), gives rise to accurate yield forecasts towards the end of the season. This is so because the remaining seasonal climate has little effect on the crop and reliance on the assumption of average climate is no longer important. Early season forecasts are inaccurate in years of extreme climate as average climate data cannot represent these conditions (Walker, 1989). The use of multiple years of historical climate data to produce a probability distribution of yields can be very useful. The range of possible yield outcomes allows for risk to be associated with any management decisions that are made. Stephens *et al.* (1989) made use of seasonal climate forecasts to complete the seasonal climate data record in their ACY model. They translated the three-monthly categorical rainfall forecasts into suitable model data sets, by assuming that below normal, normal and above normal rainfall corresponded with the 15th, 50th and 85th percentiles of the rainfall probability distribution of the area. Forecasts of wheat yield were made at various times during the 1984, 1985 and 1986 growing seasons, and were compared with the ten year average yield and the actual final yield. Throughout all three seasons, forecasted yields were closer to the actual yield than the ten year average. The forecasts improved as the seasons progressed, owing to the incorporation of observed climate data. For early season forecasts, the accuracy of the forecasted yield was dependent mostly upon the accuracy of the climate forecast used (Stephens *et al.*, 1989). It is important to consider that the value of a yield forecast is usually not in the actual yield predicted, but in its relative magnitude to the previous year (Stephens *et al.*, 1989).

ACY models are ideally suited to real time regional yield forecasting (Kalma *et al.,* 1991). They have been successfully implemented at locations in Australia (Stephens *et al.*, 1989), North America (Meyer *et al.*, 1993; Walker, 1989), Africa and in South Asia (Popov, 1990). Their success is ascribed to their use of a minimum amount of actual data during the season and the fact that strong relationships between yield and the agroclimatological indices are possible. ACY models are easily transferrable between locations and crops, with the only calibration needed being that between the index and yield (Stephens *et al.*, 1989; Walker, 1989). In areas where other factors beside the climate play an important role, a framework should be provided to accommodate these factors in the forecasting procedure (Motha and Heddinghaus, 1986). Improvements in the accuracy of climate forecasts are likely to improve the results of ACY models more than improvements in the models themselves (Walker, 1989).

3.3.1.2 Agrohydrological yield models

Lecler *et al.* (1996) applied the *ACRU* maize AHY model in maize yield forecasting in a pilot study of the application of seasonal rainfall forecasts in an AHY model to produce crop yield forecasts. The methodology adopted in this project for sugarcane yield forecasting is very similar to that used in the pilot study of Lecler *et al.* (1996). The maize yield forecasting study is reviewed below, as it forms a strong basis of the cane yield forecasts.

The *ACRU* maize model calculates seasonal yield from the potential seasonal maize yield of a district, and the ratio of actual to potential transpiration occurring for the season (Schulze *et al.*, 1995). The ratios of transpiration are determined for the various growth stages occurring in maize, with each stage being given an appropriate stress weighting. The timing of growth stages is determined according to growing degree days (a thermal time concept).

Lecler *et al.* (1996) translated seasonal rainfall forecasts into representative local daily rainfall data sets that could be input into the *ACRU* maize model to forecast yield for their study area in a prominent maize growing region of South Africa. The rainfall forecasts used were those issued by the South African Weather Bureau (Landman, 1995). These forecasts were categorical in nature and had a lead time of up to 6 months, with categories of rainfall being predicted for each month. The categories corresponded to being either above normal, near normal or below normal rainfall. A methodology was developed by Lecler *et al.* (1996) to downscale the seasonal rainfall forecasts, both spatially and temporarily, for input into the *ACRU* model. This methodology was as follows:

Step 1

Daily rainfall data for a selected representative station in the forecast area were totalled to give monthly values of rainfall.

Step 2

The monthly rainfall totals were subjected to a frequency analysis to determine monthly percentiles of rainfall.

Step 3

The three possible categories of rainfall were defined in terms of probabilities. A monthly total of rainfall in excess of the 65th percentile value was defined as an above normal (A) rainfall. A monthly rainfall of less than or equal to the 65th percentile value and greater than or equal to the 35th percentile value, was defined as a near normal (N) rainfall. Monthly totals of rainfall below the 35th percentile value were defined as below normal (B) rainfalls.

Step 4

Rainfall totals for each of the months October through to March (i.e. the maize growing season in South Africa) for the period from 1931 to the year before the season being forecast, were examined, to determine which years had months with rainfall totals within the given forecast category, *viz.* A, N or B. If a month's rainfall total was within the given forecast category, the preceding month's monthly rainfall total was then also examined to determine whether it was within its given forecast category. The preceding month's rainfall totals were also considered because some continuity between monthly rainfall is likely (Berri, 1995 cited by Lecler *et al.* 1996).

Step 5

The selected years having monthly totals (October to March) corresponding to the forecast categories were arranged in all possible combinations, but in the correct monthly sequence and written to an ASCII file. If, for example, for each of the months October to March there were six years from the total data record for which the relevant monthly rainfall totals satisfied the prediction category criteria, there would be a total of 6^6, or 46656 possible arrangements of composite monthly sequences.

Step 6

The large number of possible monthly sequences generated in Step 5 would have been impractical for modelling purposes. Hence a random sample of 100 sequences was drawn from the total number of possible sequences.

Step 7

The daily rainfall data corresponding to the months and years selected in Step 6, were written into the rainfall file format required by the *ACRU* model. Observed daily data were written into the files for the period prior to the growing season. Arbitrary data were written into the files for the period at the end of the season where rainfall forecasts were not available. The simulation period prior to the growing season served to initialize the water budget before planting of the crop.

Following the generation of each of the rainfall files, the maize yield simulation model was executed and the results appended to a file. Maize yields were simulated for the seasons 1982/1983 to 1992/1993 and the mean determined of the 100 yields simulated for each season.

The technique for downscaling rainfall forecasts was of primary interest in this pilot study and in order to test the technique, forecast categories were selected based on the assumption of a perfect forecast, ie. the actual categories that occurred. This eliminated uncertainty in the accuracy of the rainfall forecasts, and allowed the downscaling technique to be evaluated. Yields were also simulated for each season using entirely observed rainfall data. These simulations served as the basis against which forecasted yields were compared, and were assumed to represent yields obtained in the region.

The forecasted yields and yields simulated using observed rainfall data were plotted in the form of a time series plot and a scatter plot. Both figures indicated that the forecasted yields corresponded closely to the yields simulated using observed rainfall data, with the largest difference in yield of 0.5 t/ha occurring in the 1986/1987 season. The scatter of yields was random, with the data points being distributed closely to the 1:1 line. The correlation coefficient (Pearson's r) between the two sets of yields was 0.94, and the Index of Agreement (Willmot, 1981) was 0.97 (Lecler *et al.*, 1996).

The methodology to translate regional categorical forecasts of rainfall into locally representative, daily rainfall data sets for use in the *ACRU* maize model was concluded to be successful. Lecler *et al.* (1996) also concluded that the rainfall forecast downscaling methodology, when integrated with a yield simulation model and rainfall forecasts, was a potentially powerful tool for crop production planning on a season by season basis. In the same study a methodology was also developed to downscale six month aggregate forecasts of rainfall (obtained from Mason, 1996, cited in Lecler *et al.*, 1996) into locally representative daily data sets for application in the *ACRU* model for streamflow forecasting in the Bivane River in KwaZulu-Natal, South Africa. Based on the use of perfect rainfall forecasts, forecasts of streamflow compared favourably with streamflow simulated using observed rainfall. However, when a set of actual rainfall forecasts for the period concerned was used to forecast streamflow using the *ACRU* model, the accuracy of the streamflow forecasts was found to deteriorate. It was concluded that improvements in rainfall forecasting were still required before accurate streamflow forecasts could be obtained. This finding could also apply to crop yield forecasts. Lecler *et al.* (1996) recommended that the forecasting methodologies developed, be assessed in other regions of the country to identify regional patterns in the accuracy of forecasting. It was further recommended that the forecasting methodologies be re-assessed using three month aggregate forecasts of rainfall (both perfect and observed) which have, more recently, been issued by the South African Weather Bureau. The use of updated rainfall forecasts during a season was also suggested, with observed rainfall records being updated simultaneously in near real-time. This would result in crop yield forecasts being adjusted as the season progresses, with the potential for improvements in the accuracy of forecasts. The latter two recommendations made by Lecler *et al.* (1996) are applied in the methodology evaluated in this project for the translation of seasonal rainfall forecasts into probabilistic forecasts of sugarcane yield.

The water budgets of AHY models such as *ACRU* are more complicated than those of ACY models, and this has implications for preparation of model inputs. Their greater hydrological emphasis and the non-use of agroclimatological indices, distinguishes AHY models from ACY models.

3.3.2 Growth models

Crop growth models attempt to simulate the development and growth of a crop through its various growth stages until maturity (Hammer *et al.*, 1996b). They will typically have subroutines for water budgeting, crop growth and phenological development. Crop development refers to the advance of the crop from one growth stage to the next. Growth refers to the accumulation of carbohydrate through the process of photosynthesis and the partitioning of this carbohydrate into the various parts of the plant. Growth models run on a daily time-step and usually assess the current phenological stage of the crop through the calculation of growing degree days, which are determined from temperature data. The growth for a day, depends on the response of the crop to the growing environment for that day. This response is a function of phenological stage and cultivar variety (Hammer *et al.*, 1996b). Input data required by growth models commonly relate to climate, soil, genetics and management. Climate data usually include daily maximum and minimum temperatures, precipitation and solar radiation (Duchon, 1986; Lourens, 1995).

Before a crop growth model can be used in a crop forecasting system, its output must be verified against historical data from the area (du Pisani, 1987; Wilkens, Thornton and Bowen, 1994). If a model cannot produce good simulations based on observed climate data, then it is unlikely to do so in a forecasting environment where there is uncertainty about future climate. Verification must be carried out at the same scale that the model is to be applied for forecasting. If forecasting is to be carried out at a scale larger than the farm level, such as at regional or national level, then a spatial modelling framework must be developed with the preparation of spatially distributed model inputs (du Pisani, 1987; Lourens, 1995). This should be done bearing in mind the sensitivity of model simulations to the various input parameters (du Pisani, 1987). For example, du Pisani (1987) found that under South African conditions the CERES-maize model (Ritchie, 1985) was more sensitive to plant available water capacity than it was to planting date (provided that date was within two weeks of the actual), and that care should be taken accordingly when preparing input parameters. The need for verification and a spatial modelling capacity apply also to yield models, but this need is less critical as data requirements are less demanding (Hammer *et al.*, 1996b). When developing spatially distributed climate data for input into models, interpolation techniques are often employed (Lourens 1995). In his development of an agricultural drought monitoring system for South African conditions, Lourens (1995) made use of existing interpolation techniques to develop observed rainfall and temperature databases, while for solar radiation he modified a satellite based technique developed by Japanese researchers. This satellite based technique which made use of METEOSAT data, was recommended for further research to determine if temperatures could also be estimated. A satellite based technique yields truly spatially distributed data, as opposed to interpolation which makes use of existing data to make inferences about areas where data is not available. The use of a GIS can greatly aid in the development and running of a spatially distributed modelling system (Lourens, 1995; Hammer *et al.*, 1996b). This applies to the generation (in some cases), storing, manipulation and displaying of model input and results (Lourens, 1995).

The use of growth models in crop forecasting also requires assumptions to be made regarding climate between the forecast date and the end of the season. The techniques used for yield models are also used for growth models. Median climate data were used by du Pisani (1987) to complete the seasonal record. Monthly probability distributions were developed for each climate element and for each location in the study and the median values determined. Those months in the historical record whose medians were closest to those of the probability distributions, were selected and the data from these months used to complete the data record. This procedure was adopted to avoid using daily averages on a day by day basis which do not reflect realistic sequences of climate. In the case of rainfall, daily averages are very small and would not contribute much to growth. Lourens (1995) used a similar approach except above average, average and below average scenarios were considered. These scenarios corresponded with the

90th, 50th and 10th percentiles of the probability distributions. Duchon (1986) used all the years available in an historical data record to complete the seasonal data record. A procedure was presented to "splice-in" in the years so that a smooth transition was ensured from the current year to the year in question. The procedure was carried out for maximum and minimum temperature and was performed at each forecast date during the season.

Another approach to completing a seasonal data record could be to analyse the observed data up until the time of forecast, and then to find similar (analogue) years in the historical record according to pre-determined criteria (Sakamoto, 1989). These analogue years of record would then yield a very much smaller range in final yields. In situations where there is a long delay between when rainfall measurements are made and when they are available for crop forecasting, rainfall may be estimated in real time through the use of satellite images of cold cloud duration and appropriate interpretation algorithms (Wilkens *et al.,* 1994)

Seasonal climate forecasts have been used to complete the climate data record of growth models used in crop yield forecasting. Examples of studies where this has been done are those of Carter and Brook (1996), cited by Hammer and Nicholls (1996), and Hammer *et al.* (1996a). Carter and Brook (1996) developed a prototype national drought alert system which emphasised Australia's rangelands and was based on a spatial model of pasture growth coupled with the seasonal climate forecast technique of Stone and Auliciems (1992). The forecast technique of Stone and Auliciems (1992) involves classifying the current season into one of five phases depending on the value and rate of change of the Southern Oscillation Index. Hammer *et al.* (1996a) tested three climate forecasting techniques in relation to their usefulness for management of wheat. The average profit and risk of making a loss were calculated for the possible range of fixed and tactical management strategies, based on the non-use or use of climate forecasts. The potential value of improved forecast quality was also considered. Significant increases in profit of up to 20 percent and/or reduction in risk of up to 35 percent were associated with tactical adjustment of nitrogen fertilizer or cultivar maturity. When tactical management decisions based on climate forecasts were compared with decisions based on perfect prior knowledge of the season, it was concluded that current skill in seasonal forecasting is sufficient to justify the use of forecasts in decision making. They were quick to caution, however, that the success or failure of adopting a tactical strategy in any one year was not certain and that the success achieved related to average performances over a number of years. Of the three climate forecasting techniques that were tested, the method of

Stone and Auliciems (1992) was found to give the greatest value in management decisions. Significant skill was obtained for predicting seasonal rainfall and timing of frost.

Varying degrees of success have been achieved in forecasting crop yields using growth models in combination with assumptions relating to seasonal climate. Du Pisani (1987) noted that the system he developed lent itself most promisingly to assessments of drought impacts on South African maize production and that these impacts should be predicted to within acceptable limits up to four months prior to harvest. The drought monitoring system of Lourens (1995) accurately portrayed general maize production trends during the severe drought of 1991/1992. Wilkens *et al.* (1994) found good agreement between observed and forecasted millet yields following a preliminary analysis for the year of 1986. Projected yields were found to be within the final estimates after only 50 days of simulation.

Duchon (1986) and Carter and Brook (1996) were less positive regarding their results. Duchon (1986) stated that his approach represented a first step toward the ultimate goal of developing a real-time maize production forecasting system. Carter and Brook (1996) calculated correlations between simulations based on actual and forecasted climate data. Caution was given to the use of the Stone and Auliciems (1992) climate forecast technique for forecast periods longer than 90 days, as correlations became poor.

De Jager, Potgieter and van den Berg (1998) forecasted maize yields as part of a drought monitoring system using a growth model and climate forecasts based on the method of Stone and Auliciems (1992). Although tests of accuracy were still being performed when they reported in 1998, the system was being applied operationally in the Free State Province of South Africa and enjoyed wide acceptance and credibility amongst users. Stochasticity was maintained in the system by comparing the forecasted yields with the long term cumulative probability distribution and determining the relevant level of probability. A number of analogue years corresponding to the current phase of the SOI were used to forecast yields and the average of the resulting probabilities was given as the final result. The stochasticity of the forecasts was considered important as it provides a perspective of the likely outcomes for decision-makers.

A facility is currently available on the SASEX internet site (http://www.sasa.org.za/sasex/irricane/) to enable simple cane yield forecasts to be made (Singels, Kennedy and Bezuidenhout, 1999). The facility makes use of the CANESIM model to simulate yields, where climate inputs to this model may be derived from a number of automatic weather stations situated in the sugarcane growing regions of South Africa. Yields may be simulated for past seasons for benchmarking purposes, or for the current season where a forecast of the seasonal yield is required. When forecasting yields, observed climate data are used to fill the seasonal record up to the date of forecast. The period following this until harvest, is then filled with historical data from selected years which resemble the current phase of the SOI. Soil total available moisture, crop cycles and irrigation status are considered in the CANESIM simulations.

The general indication with respect to the use of growth models in crop yield forecasting is that the models hold promise as a tool for the future. The accuracy of the method depends partly on the adequacy of the crop model and partly on the technique used to complete the seasonal data record (Duchon, 1986). Model outputs should be verified during the development phase of a forecasting system using accurate data from the area of interest. The accuracy of seasonal climate forecasts is currently improving and this, in turn, improves the prospects for the use of growth models in crop forecasting. The advent of climate forecasts allows for crop predictions to be made on the basis of more than just what is known up to the time of forecast (Hammer and Nicholls, 1996), although some success is possible if observed data are used (du Pisani, 1987; Wilkens *et al.,* 1994; Lourens, 1995). There are, however, considerations that must be resolved with regard to the use of climate forecasts and at present they are still far from perfect (Hammer and Nicholls, 1996). An example of an area that requires further research is the scaling down of regional climate forecasts to the local level, and the subsequent extrapolation of simulated yield and farm management implications back up to the regional scale (Phillips, Rosenzweig and Cane, 1996).

3.4 Comparison of Techniques

Hammer *et al.* (1996b) evaluated the regression, yield simulation and growth simulation techniques for Australian wheat at the county, state and national scale. All techniques were found to forecast yield satisfactorily at the county scale, although the regression and yield simulation techniques displayed greater predictive ability. At the state scale the regression technique performed consistently and was marginally superior at the national scale to yield simulation. The performance of the growth simulation models used was said to have been affected by a poor knowledge of the spatial distribution of cropping history and management, to which growth models are sensitive. If improvements to the data input were made, these would have to be weighed up in terms of the precision gained. The yield simulation models were identified as having the most potential for use in the operational environment owing to their robust nature and their sufficient representation of biophysical processes, thus overcoming concerns relating to extrapolation to different seasons and years. The additional data requirements of yield models when compared to regression models, were considered not to be restrictive.

The choice of yield simulation models as the best overall technique is supported by Mottha and Heddinghaus (1986), Walker (1989) and Stephens *et al.* (1994). These models which make use of simple data and which include some account for phenological growth stages and the balance between water supply and crop demand, were said to combine the strengths of regression and growth simulation. Regression models cannot explain cause and effect relationships which incorporate complex, non-linear interactions between the independent variables (Stephens *et al.*, 1989). Some cannot account for antecedent conditions prior to a season (Walker, 1989), or may be biased by outliers in the yield data set (Stephens *et al.*, 1989). Yield simulation models usually include regression relationships. However, the combination with simple water budgets and transformation of measured variables into index form, allows for many of the regression associated problems, including collinearity of variables, to be overcome (Sakamoto, 1989; Stephens *et al.*, 1994). The transformation of RS data to an index also overcomes many regression related problems.

If the spatial data input of growth models can be obtained at a sufficient level of accuracy, then growth models can be very useful for evaluating potential changes in management and the resultant effect on yields. This is particularly so with the advent of improved accuracy in seasonal climate forecasts (Hammer *et al.*, 1996a). In addition to difficulties relating to spatial data input, Wiegand and Richardson (1990) note the difficulties associated with simulating simultaneous multiple stresses in growth models. RS approaches are suggested as an alternative or perhaps as an aid to growth modelling as the direct canopy measurements allow for large area responses to be evaluated with all stress effects inherently accounted for (Horie *et al.*, 1992).

Research has been conducted into combining the results of current forecast methods to evaluate whether an improved result can be achieved. A measure of success has been obtained in this regard (Brodskil and Kostyukov, 1992; Pandey, Dadhwal, Sahai and Kale, 1992; Pasov and

Yatsalo, 1992).

3.5 Discussion

The techniques of regression modelling, remote sensing, yield simulation and growth simulation have been reviewed in relation to their use and suitability for forecasting crop yields. Regression modelling is a simple, but location specific, technique. It is applicable at all spatial scales, but especially at larger scales where techniques requiring minimal data are preferred. If the area of interest is influenced by the ENSO phenomenon then these effects can and should be incorporated into the modelling framework. Future weather indicators and current season weather variables are both preferable to predictors based on crop observations, as gathering these observations is a timely and costly exercise. Regression modelling is subject to a number of problems associated with using observed yield data during model development.

Remote sensing techniques are also based on crop observations, however the spatial coverage is considerably better than that of traditional field observations. Regressions of vegetation indices against yield offer reliable results, but are still to a degree location specific. Radiation interception and canopy temperature models assist in overcoming this limitation, but tend to be more restricted in their application. The influence of crop stress, be it through water or temperature stress or lack of nutrients, complicates the application of vegetation index and radiation interception models. Canopy temperature models tend only to be applicable once the canopy has formed. RS approaches require a degree of data processing that should not be overlooked. Classification of pixels can be problematic and depends to an extent on the skill of the individual concerned.

Yield simulation models in the form of ACY and AHY models are sensitive to climate, and are useful tools for yield forecasting in situations where yields are dominated by climate variability. Many important grain producing areas are dominated by such variability. Yield models are relatively simple simulation models that require minimum amounts of data, and are generally robust in nature. Growth simulation models have much greater input data requirements and consequently demand more time and expense for the collection and preparation of these inputs. If a few factors dominate interannual yield variability, it is likely that growth models, once verified, become simpler to apply in yield forecasting. Year to year running of the models need only focus on those few dominant factors, thus making the technique less demanding. The use of RS to assist in the preparation of spatial data inputs, could receive more attention, as this has potential to provide more accurate input data and thus improved results. The benefits of using growth models revolve around their ability to simulate a wide range of conditions, including management scenarios. Differing management scenarios influence yield and therefore the crop expected for a season.

Based on the review of forecasting techniques, it appears that all of the techniques reviewed could potentially be applied for small areas where all the necessary data are available. For large areas data availability becomes critical, and the techniques of regression modelling and remote sensing are favoured over growth simulation modelling. Yield simulation combines the advantages of regression and growth simulation and is a robust, dynamic technique that does not have prohibitively large data requirements. The choice of yield forecasting technique for a particular application, should include considerations relating to availability of data, the time and expertise required to implement the system, the scale of operation, the accuracy required and the factors which necessitated the development of the system. It is believed that all of the yield forecasting techniques reviewed could potentially be applied at a mill supply area scale in the South African sugar industry. In this project the techniques of growth simulation and yield simulation were evaluated for application, as well as a simple rainfall model representative of traditionally employed techniques in industry. This model is not a formally developed regression model, but is similar in nature in that it is empirically derived from observed yield and rainfall data.

The area selected for investigation in this project, namely the Eston Mill Supply Area, has good climate and soil data records, which can be interpolated spatially to satisfy the input requirements of the relevant yield and growth models proposed for use in the research. The potential benefits that could be derived from improved accuracy and timeliness in sugarcane yield forecasts may compensate for the time and effort expended in developing and evaluating a simulation modelling based yield forecasting system.

In this chapter, a review of techniques available for forecasting crop yields has been presented. In the following chapter, a description of the models actually applied in this study, is given. It is important to gain an understanding of these models, in order that their application, and the interpretation of the resulting output, be carried out correctly.

4 SUGARCANE YIELD MODELS EVALUATED IN THIS STUDY

The sugarcane models evaluated in this study are the Simple Rainfall Model, the *ACRU*-Thompson yield model and the CANEGRO-DSSAT growth model. These models are representative of simple, intermediate and higher levels, respectively, of model complexity.

4.1 Simple Rainfall Model

The SRM (Cousens, 1998) relates sugarcane yield to rainfall through a relationship developed from experience. The average rate of yield accumulation per 100mm of rainfall is calculated, based on observed means of yield and rainfall. This rate is assumed to be constant and is applied to determine the yields of individual seasons. The model is applied in the form of a spreadsheet and an example of the calculation procedures is given in Tables 1 and 2, for a hypothetical 14 month crop harvested in 1995. Aggregate statistics relating to the region of application, including the calculated average rate of yield accumulation, are given in Table 1, while the calculation of the seasonal yield is given in Table 2. In order to calculate the mean rate of yield accumulation, the average long term yield is first annualized (by multiplying by the ratio of 12/14), and then divided by the mean annual precipitation. A multiplication factor of 100 is then applied to obtain a per 100mm rate of yield accumulation.

Table 1 Rainfall and yield statistics relating to a hypothetical 14 month crop harvested in 1995

Average Length of Crop Cycle (months)	14
Average Long Term Yield (t/ha)	80
Mean Annual Precipitation (mm)	729
Mean Rate of Yield Accumulation (t/ha/100mm)	9.41

The length of the growth cycle used in calculations would have been selected to be representative of the area under consideration. For each month of the growing cycle, the yield increment corresponding to that month is calculated from the average yield accumulation rate, and the rainfall occurring in that month. A crop harvest is assumed to occur for each month of the harvest season, which in this case extends from April to December. For each of the months of harvest, the

Table 2 Calculation of the seasonal yield of a 14 month hypothetical crop harvested in 1995 using the Simple Rainfall Model

Year	Month	Rainfall (mm)	Yield Increment (t/ha)	Yield (t/ha)
1994	Feb	24.1	2.27	
1994	Mar	203.0	19.11	
1994	Apr	14.5	1.36	
1994	May	8.7	0.82	
1994	Jun	13.0	1.22	
1994	Jul	59.2	5.57	
1994	Aug	43.4	4.08	
1994	Sep	1.6	0.15	
1994	Oct	74.6	7.02	
1994	Nov	43.2	4.07	
1994	Dec	49.7	4.68	
1995	Jan	23.4	2.20	
1995	Feb	45.6	4.29	
1995	Mar	145.3	13.68	
1995	Apr	76.0	7.15	70.5
1995	May	13.8	1.30	75.4
1995	Jun	72.7	6.84	57.6
1995	Jul	5.6	0.53	63.1
1995	Aug	6.2	0.58	62.8
1995	Sep	16.8	1.58	62.1
1995	Oct	102.1	9.61	58.2
1995	Nov	142.2	13.38	63.7
1995	Dec	331.1	31.16	76.9
			Mean	65.6

preceding 14 monthly yield increments are summed separately to give a seasonal yield. The mean of the yields calculated for each month of harvest is then determined, and given as the final model estimate. The final yield estimate in this case is 65.6 t/ha.

The SRM is designed to calculate mill supply area yields. It would typically be applied in practice by using the rainfall data from a representative rainfall station in the MSA, and observed yields from the mill. The rainfall data would be used to determine the mean annual precipitation, as well as the monthly rainfalls occurring during the growth cycles. When running the model in real-time during a growing season, assumptions would have to be made regarding the monthly rainfalls occurring between the date of forecast and the harvest date. The observed yields from the mill are used to calculate the long term average yield of the MSA, which is used in calculating the mean rate of yield accumulation.

The model is based only on rainfall and does not take into account directly other factors affecting the growth of a crop such as solar radiation, temperature, soils, soil water status and management. However, the model is a calibration model, as observed yields are used in calculating the mean rate of yield accumulation. Other factors affecting growth are thus accounted for indirectly. The relationship between yield and rainfall is constant, however, and changes in the growth environment of the crop over time are thus not reflected. The SRM, like other calibration models, requires a good quality observed data set of the variable being predicted (in this case cane yield) in order for accurate model predictions to be made.

4.2 *ACRU*-Thompson Model

The *ACRU*-Thompson model used in this study (Lumsden *et al.*, 1998) is an enhancement of the version described in Schulze *et al.* (1995), and comprises of the Thompson sugarcane yield model (Thompson, 1976) imbedded within the *ACRU* agrohydrological modelling system (Schulze, 1995; Smithers and Schulze, 1995). The following review of the *ACRU*-Thompson model comprises of sections relating to processes represented within the *ACRU* model, *ACRU* model input and output, modifications to the *ACRU* model and background to the Thompson sugarcane yield model.

4.2.1 Processes represented within the *ACRU* model

The *ACRU* model is a two layer (top and subsoil) soil water budgeting model that operates at a daily step. A number of processes are represented within the structure of the model, as illustrated in Figure 1. Water may enter the soil profile in the form of rainfall or irrigation, and may exit in the form of runoff (quickflow and delayed subsurface flow), total evaporation (transpiration, evaporation from the soil and canopy surfaces) and deep percolation into the groundwater store.

Rainfall and/or irrigation water which is not abstracted as interception or as stormflow (either rapid response or delayed), is infiltrated through the soil surface and is stored in the topsoil horizon. When the topsoil is "filled" to beyond its drained upper limit (field capacity), the "excess" water percolates into the subsoil horizon(s) as saturated drainage, at a rate dependent on respective horizon soil textural characteristics, wetness and other drainage related properties. Should the soil water content of the subsoil horizon of the plant root zone exceed the drained upper limit, saturated vertical drainage or recharge into the intermediate and eventually groundwater stores occurs, from which baseflow may be generated at an exponential decay rate dependent on geological/ aquifer characteristics and the groundwater store. Unsaturated soil water redistribution, both upwards and downwards, also occurs, but at a rate considerably slower

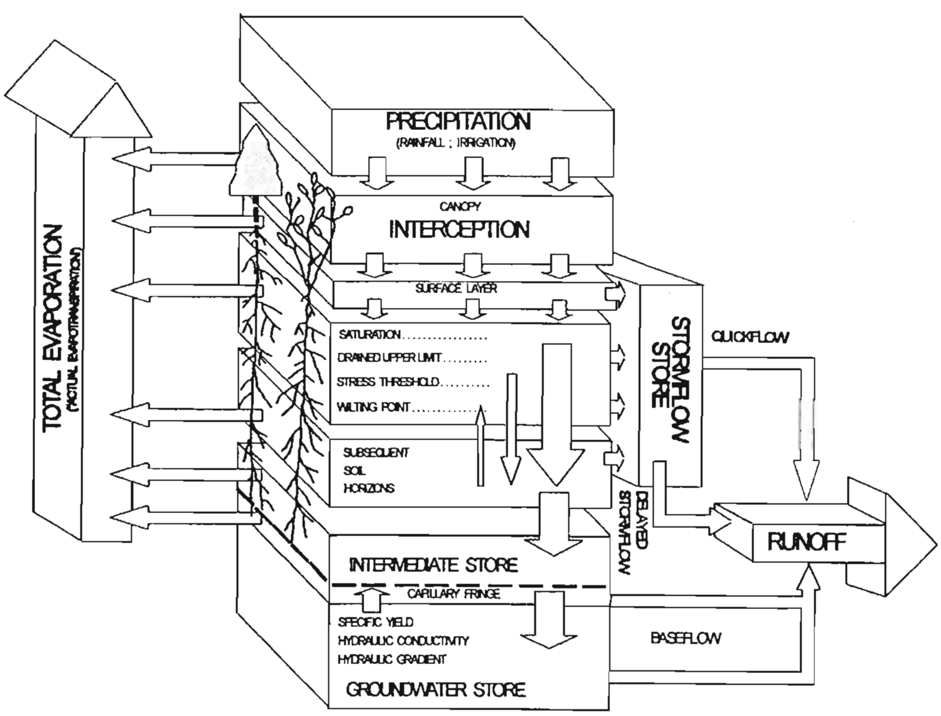


Figure 1 Structure of the *ACRU* modelling system (Schulze, 1995)

than the water movement under saturated conditions, and is dependent, *inter alia*, on the relative wetnesses of adjacent soil horizons in the root zone.

Evaporation takes place from previously intercepted water and from the various soil horizons. Evaporation from soil horizons is either split into separate components of soil water evaporation and plant transpiration, or combined, as total evaporation. Evaporation is estimated by considering the atmospheric demand, the water use characteristics of the crop and the moisture available in the soil. Atmospheric demand is represented through a reference potential evaporation (A-pan or A-pan equivalent) and the water use characteristics of the crop through water use coefficients. These coefficients, which are ratios of maximum evaporation to reference evaporation, depend on climate and the crop's stage of growth, and are usually input into the model as 12 monthly values. Maximum evaporation is determined by multiplying the relevant water use coefficient by the reference evaporation. The roots of the crop absorb soil water in proportion to the distributions of root mass density within the respective horizons, except when conditions of low soil water content prevail, in which case the relatively wetter horizons provide higher proportions of soil water to the plant in order to obviate plant stress as long as possible.

It is vital in crop yield modelling to determine at which point in the depletion of the plant available water reservoir plant stress actually sets in, since stress implies a soil water extraction below optimum, the necessity to irrigate (if irrigation is applied) and also implies a reduction in crop yield. In modelling terms, this problem may be expressed as the critical soil water content at which total evaporation, E, is reduced to below the vegetation's maximum evaporation, E_m (formerly termed "potential evapotranspiration"). E equals E_m until a certain fraction of maximum (profile) available soil water to the plant is exhausted. Research shows that the critical soil water fraction at which stress commences, varies according to atmospheric demand and the critical leaf water potential of the respective vegetation, the latter being an index of the resilience of the vegetation to stress situations. The implications of stress setting in at such different levels of soil water content are significant in terms of total crop evaporation and crop yield modelling.

4.2.2 Model input and output

In order to simulate crop yields, the *ACRU* model requires inputs of known, measurable, or derivable factors including data or information on:

- climate (e.g. daily rainfall, maximum and minimum temperature, potential evaporation)
- soils (e.g. horizon thicknesses, soil water retention, saturated drainage rates)
- land uses (e.g. crop types, levels of management, planting dates, growth rates, above-and-below-ground vegetation attributes at different growth stages)
- soil water budgeting (e.g. onset of plant stress, degrees of stress, capillary movement)
- runoff producing mechanisms (e.g. stormflow generation, baseflow rates)
- irrigation practices (if relevant, e.g. crop type, above-and-below-ground attributes at different growth stages, modes of scheduling and their controls, source of water, application efficiencies) and
- dams (if present, e.g. inflows, full supply capacities, surface areas, evaporation rates, releases, abstractions and inter-basin transfers).

This information is transformed in the model by considering

- the climate, soil, vegetative, hydrological and management *subsystems*
- how they *interact* with one another
- what *thresholds* are required for responses to take place
- how the various responses are *lagged* at different rates and
- whether there are *feedforwards* and *feedbacks* which allow the system to respond in a positive or reverse direction.

The model then produces output of the unmeasured variables to be assessed, such as stormflow, baseflow, sediment yield or reservoir status, but particularly relevant to this study

- crop yield (e.g. per season, annum or growth cycle; dryland or irrigated; and where relevant, with economic analysis) and
- irrigation water requirements (gross or net requirements, associated crop yields, deep percolation and stormflow from irrigated areas ; water use efficiencies under different modes of scheduling irrigating water ; benefit of irrigated vs dryland farming).

Risk analysis of the above outputs may be performed by the model (e.g. month-by-month, annual or seasonal statistics, extreme value analysis).

4.2.3 Modifications to the *ACRU* model

In the existing *ACRU* model (Schulze *et al.*, 1995) sugarcane yields are estimated assuming an annual crop (July to June). Water use by the crop is estimated through 12 monthly water use coefficients, the values of which may be set to 0.8 for average on-farm conditions (Schulze *et al.*, 1995). Given the effect of growth cycles on sugarcane yields (Hellmann, 1993), it was considered important that the various growth cycles occurring in the Eston area be represented in the modelling framework. In order to cater for a variety of growth cycle lengths and harvest dates, Lecler modified the *ACRU* model through the introduction of dynamic equations relating crop water use to daily temperature, as reported in Lumsden *et al.* (1998). These equations, taken from the research of Hughes (1992), allow for the calculation of daily water use coefficients. The equations are as follows:

$$K_c = 0.297 + (1.32 \times 10^{-6} \times GD_a^2) - (6.83 \times 10^{-10} \times GD_a^3) - K_{red}$$

$$K_{red} = 0.05 + (1.32 \times 10^{-6} \times GD_r^2) - (6.83 \times 10^{-10} \times GD_r^3)$$

where

K_c = sugarcane water use coefficient

GD_a = accumulated degree days since planting and up to initiation of ripening at 1300 °C day (°C day)

GD_r = accumulated degree days after initiation of ripening (°C day)

K_{red} = reduction in water use coefficient after ripening

Degree day = $((T_{max}+T_{min}) / 2) - 12$ (°C day)

T_{max} = daily maximum temperature (° C)

T_{min} = daily minimum temperature (° C) .

Limits to K_c, taken from Hughes (1992), are:

K_c ≤ 1 for plant crop

≤ 0.96 for first ratoon crop

≤ 0.92 for second and subsequent ratoons

≥ 0.5 after initiation of ripening .

Daily observed maximum and minimum temperatures are input into the equations to allow for the calculation of the water use coefficients. If these temperatures are not available, then monthly long term means of temperatures may be specified, with these temperatures then being translated internally in the model to daily values by Fourier Analysis.

As the water use coefficients are related to temperature, they reflect the climate regime experienced by the crop during its growth cycle, thus allowing for the representation of different harvest dates. The use of temperature based relationships also overcomes the limitation in the existing *ACRU* model, which restricts the length of growth cycles to 12 months. The influence of two different harvest dates on the seasonal water use coefficient curve of a 12 month crop are illustrated in Figure 2. The curves were derived from temperatures recorded in the Eston area.

The curve of the crop harvested in October rises rapidly after growth commencement, reflecting the warm temperatures experienced by this crop in its initial growth stages during the summer months. In contrast, the curve of the crop harvested in April rises slowly after growth commencement, reflecting the colder winter temperatures experienced in the early stages of this crop's growth cycle.

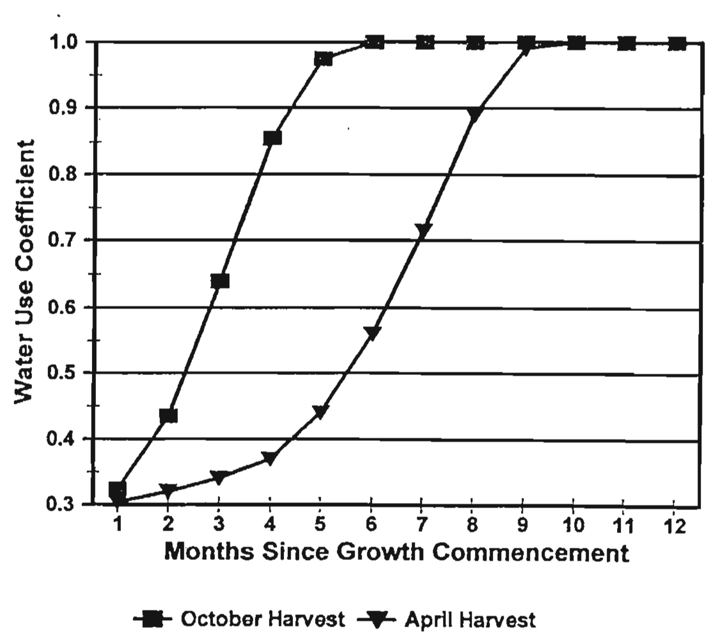


Figure 2 Seasonal water use coefficient curves of two 12 month crops harvested in October and April

4.2.4 Thompson sugarcane yield model

The Thompson (1976) model was derived from a collation and regression analysis of experimental yields and evaporation data from Hawaii, South Africa, Mauritius and Australia. The equation is as follows:

$$Y = 9.53\ (AET_{sum}/100) - 2.36$$

where Y = sugarcane yield (t/ha/growing season)

AET_{sum} = accumulated growing season total evaporation (actual evapotranspiration) in mm

The correlation coefficient (r) was 0.95 and the standard error of yield estimates 15.1 t/ha. The experiments were predominantly irrigated and thus the influence of water stress in the relationship would have been minimal. Under dryland conditions such as those occurring in the Eston area, water stress is inevitable, and needs to be accounted for in the estimation of cane yields. The *ACRU* model accounts for water stress by way of a threshold level of soil moisture deficit. If this threshold is exceeded, the ability of the plant to transpire at its potential rate is reduced.

In the *ACRU*-Thompson model, *ACRU* is used to estimate accumulated seasonal total evaporation. This seasonal total then forms the input into the Thompson equation, thus allowing for predictions of sugarcane yield.

4.3 CANEGRO-DSSAT Model

The CANEGRO model is a process level sugarcane growth model (Inman-Bamber and Thompson, 1989; Inman-Bamber, 1991a; Inman-Bamber, Culverwell and McGlinchey, 1993; McGlinchey, Inman-Bamber, Culverwell and Els, 1995). The model has recently been incorporated by Inman-Bamber and Kiker (1997) into the DSSAT shell, i.e. the Decision Support System for Agrotechnology Transfer (Tsuji, Jones, Hoogenboom, Hunt and Thornton, 1994). DSSAT contains a collection of verified crop growth models along with a database management system to facilitate preparation of model inputs, and a variety of tools for analysis and display of model output. The system is menu-driven, requires standardized input files and is designed to simplify the running of complex growth models. It is for this reason that the DSSAT version of

CANEGRO is applied in this research. This version is essentially unchanged from the original and gives identical results (Inman-Bamber and Kiker, 1997). Brüggemann (1998) conducted a comprehensive review of the CANEGRO model. The following review is adapted from that of Brüggemann's (1998).

4.3.1 Processes represented within the CANEGRO model

The CANEGRO model has been verified over a wide range of climate and soil conditions in South Africa. Crop growth is modelled only for the NCo376 sugarcane variety at the individual plant level. The temporal scale of modelling may be hourly or daily, depending on the modelling options chosen. The model comprises detailed balances for carbon, energy and water, with exchanges between these balances and the plant occurring at the root / soil water and canopy / atmosphere interfaces. A number of options are available for simulating the balances, depending on the detail of the available input data (Inman-Bamber *et al.*, 1993; McGlinchey *et al.*, 1995; Van Antwerpen, McGlinchey, Inman-Bamber and Bennie, 1996). Brüggemann (1998) presented a flow chart of the CANEGRO model adapted from Inman-Bamber *et al.* (1993) and McGlinchey *et al.* (1995). This flow chart is presented in Figure 3.

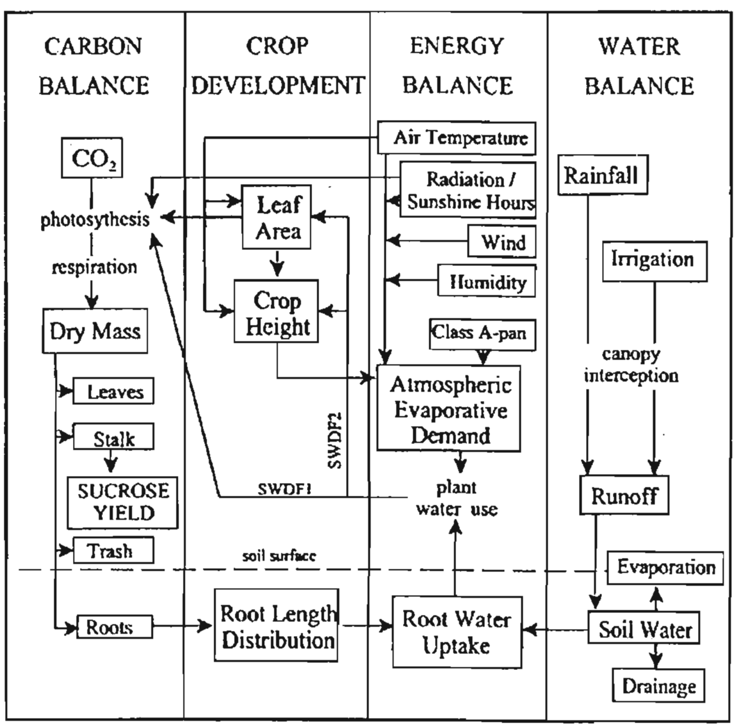


Figure 3 Flowchart of the CANEGRO model (Brüggemann, 1998, after Inman-Bamber *et al.*, 1993 and McGlinchey *et al.*, 1995)

Plant photosynthesis, respiration and partitioning of photosynthate are simulated in the carbon balance. 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, Fluck and Mishoe, 1984) and the work of McCree (1970) and Heskith, Baker and Duncan (1971), or using a modification by 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, 1991a).

An empirical approach is adopted in the determination of cane yield. It is calculated as the component of dry matter partitioned to the stalk, multiplied by the stalk population. For mature sugarcane, final stalk population is reasonably predictable, and in CANEGRO a value of 133 000 stalks ha^{-1} is assumed for variety NCo376 (Inman-Bamber, 1991a). The model predicts stalk dry mass, and this mass is then divided by the dry matter content of sugarcane stalks, about 29% (Glover, 1972), in order to derive cane yield (wet mass) as measured in the sugar industry. Partitioning of dry matter fractions within the sugarcane stalk to brix and juice purity is also described empirically in the CANEGRO model using multiple regression equations (Inman-Bamber, 1991b). 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 empirical yield equations were relatively small, 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.

Crop water demand is simulated in the energy balance by considering 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 Ångström equation (Inman-Bamber and Thompson, 1989). When daily relative humidity and / or wind speed data are not available, class A-pan evaporation is multiplied by a water use coefficient of 0.9 (for a fully canopied crop) in order to determine potential evaporation (Thompson, 1976; Inman-Bamber, 1995).

Crop water supply is simulated in the water balance of CANEGRO. This water balance (Inman-Bamber, 1991a) is a modified version of that from the CERES-Maize crop growth model (Jones and Kiniry, 1986). Available soil moisture is determined using the detailed soil water budget, where data defining the soil water holding and release characteristics are required for the soil profile in soil layer increments of 0.10 - 0.15m. Water may enter the soil in the form of rainfall and irrigation. Runoff and interception by the crop canopy are considered as losses, as is water removed from the profile by transpiration and surface soil water evaporation. Water may further be lost from the soil through drainage (Inman-Bamber, 1991a). The energy and water balances are closely linked in CANEGRO as plant water use is controlled by the energy balance when the soil water content is high, and by the water balance when the water supply is limited (Van Antwerpen, Meyer and Inman-Bamber, 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 (indicated as SWDF1 in Figure 3) reduces photosynthetic activity, thereby directly affecting biomass and sucrose accumulation. The CANEGRO model also gauges incipient water stress (indicated as SWDF2 in Figure 3), 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 reduced 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 at any one time about 90% of the sugarcane area in South Africa produces ratoon crops. No option is therefore available to specifically describe the development of sugarcane plant crops (Imnan-Bamber, 1994). If, however, plant crops must be simulated, then a constant period of 21 days may be allowed from planting to germination in order to represent plant conditions. 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 (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 days) is used to predict the rates of these processes (Inman-Bamber, 1994). The base temperature for tillering is 16 °C and stalk populations peak at about 500 °C days after ratooning. These populations stabilize at about half the peak stalk population after approximately 1200 °C days. 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, 1994). The phyllochron interval (interval between emergence of successive leaves) for each of the first 14 leaves is 109 °C days, and thereafter is 169 °C days per leaf. 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,1991a).

In CANEGRO, the total number of green leaves per stalk is allowed to vary between 3 and 11, this depending on the amount of available soil moisture. Leaf senescence is related to leaf emergence and is accelerated during periods of water stress (Inman-Bamber, 1994). The LAI for a fully canopied crop is allowed to vary between 2 and 4.5, depending on soil water content (Inman-Bamber,1991a). 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. The rate of extension is reduced according to water stress (SWDF2).

In the separation of maximum evaporation into transpiration and soil evaporation, the fraction of soil evaporation is calculated using soil water content and a simulated LAI that includes senesced leaves. These leaves are included because they continue to shade the soil for a significant period of time (Inman-Bamber, 1991a). 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 3).

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, 1991a, based on data of Van Dillewijn, 1952). Total rooting depths and root distributions vary from soil to soil and are defined empirically according to the roots' "affinity" for the defined soil layers of the profile (Inman-Bamber, 1991a). Root water use is calculated with the Richie equation of the CERES crop models (Jones and Kiniry, 1986). 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.

4.3.2 Model input and output

Inputs to the CANEGRO model include:

- crop start date (plant / ratoon), crop harvest date
- daily values of maximum and minimum temperatures, rainfall and irrigation (if applicable), daily values of either class A-pan evaporation or humidity, wind run and solar radiation/sunshine duration
- soil albedo, maximum soil water evaporation, a soil runoff category (according to three classes of cover), soil layers in 0.10 - 0.15m 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.

Where soil data are available, the soil physical properties are usually not described in adequate detail to form model inputs, and modal profiles for a number of well documented soils are usually used in model runs (Inman-Bamber *et al.*, 1993).

Numerous outputs are available from the CANEGRO model, where these outputs relate to various aspects of crop development and the carbon, energy and water balances.

4.3.3 Model applicability

The CANEGRO model is driven by solar radiation, 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, 1991b). The crop is assumed to be free of weeds and disease and is optimally supplied with nutrients. These assumptions seldom apply to commercial crops, and these and other factors 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 management, 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, Kennedy and Bezhuidenhout, 1998), for determining target production levels, i.e. for yield benchmarks (Inman-Bamber,1995; Inman-Bamber, Singels and Muchow, 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 of KwaZulu-Natal, is that CANEGRO applies only to ratoon crops of variety NCo376, which has been replaced by superior-yielding varieties, these being mainly N12, N16 and N22. While it is unlikely that the basic photosynthetic efficiencies of these varieties will be superior to those of 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,1994), which should account for some differences in varietal yield potential.

The various models applied in this study have been reviewed in this chapter. A description of the study area (Eston Mill Supply Area) is given in the following chapter.

5 THE STUDY AREA : ESTON MILL SUPPLY AREA

The Eston Mill Supply Area is situated in the Midlands of KwaZulu-Natal Province, South Africa, and is located around latitude 29°55'S and longitude 30°30'E. It comprises of farms which generally supply cane to the Eston Mill, and formerly, the Illovo Mill. Figure 4 shows the Eston MSA and indicates the boundaries of farms falling within the MSA. The roads and towns in the district are also shown. A small map is inserted to indicate the location of the Eston MSA within KwaZulu-Natal. Not all farms in the MSA were included in analyses, as a result of difficulties in obtaining good quality observed yield data. The farms that were included (numbering 85) constituted a large proportion of the total number of farms, and were believed to be a representative sample of the MSA (Hellmann, 1997). Mean annual precipitation (MAP) in the MSA ranges from approximately 600 to 1000 mm. A map showing the spatial trends in MAP is presented in Chapter 6, under the discussion relating to preparation of rainfall data sets (Figure 6). The annual means of daily maximum and minimum temperatures, as derived from two representative climate stations in the MSA, are 23.2°C and 13.2°C respectively. A map of altitude for the Eston MSA, derived from a 200m digital terrain model, appears in Figure 5. The range in altitude within the MSA is from approximately 400 to1000 m. An altitude gradient runs from north west to south east in the area. Soils originate mostly from Table Mountain Sandstone (ordinary) and Dwyka tillite parent materials, and to a lesser extent from the mist belt variant of Table Mountain Sandstone, dolerite and Lower Ecca shale parent materials (Hellmann, 1993). A map of these soil parent materials appears in Chapter 6 (Figure 9).

The study area selected in this project, namely the Eston MSA, has been described in this chapter. The following chapter relates to the verification of the sugarcane yield models applied in the study, these models being the SRM, *ACRU*-Thompson and CANEGRO-DSSAT.

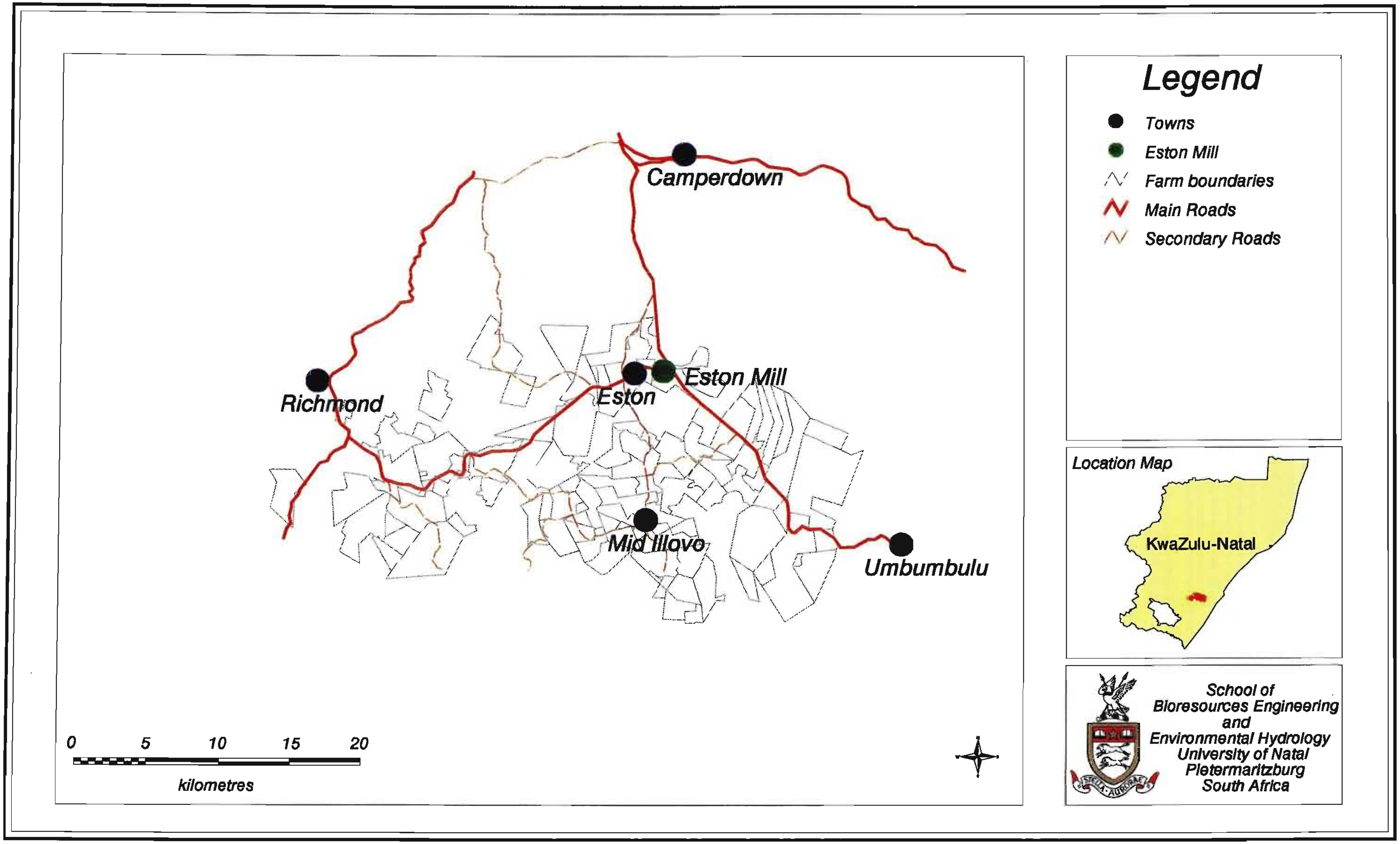


Figure 4 Eston mill supply area : Farm boundaries and other features

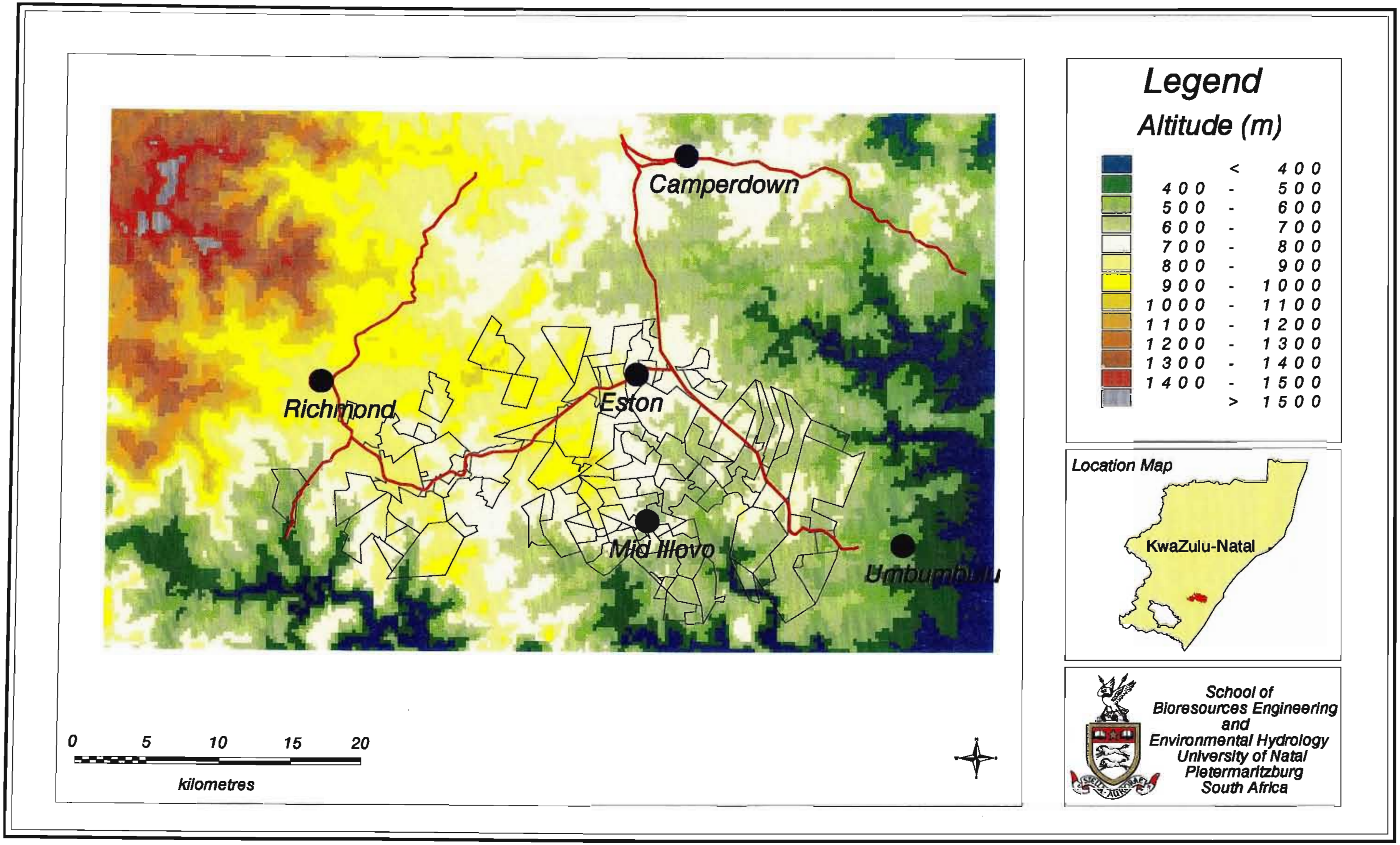


Figure 5 Altitude map of the Eston Mill Supply Area

6 VERIFICATION OF SUGARCANE YIELD MODELS

This chapter serves to verify that the models proposed for crop forecasting are able to accurately estimate the yields of past seasons, given that observed rainfall records are available for the entire season. This is important to establish before introducing rainfall forecasts, as these forecasts add additional uncertainty to the yield predictions. Yield predictions based on observed rainfall should be sufficiently accurate before attempting yield forecasting. The evaluation of model performance further serves to identify which models may be more suited than others to yield forecasting. The yield models that were evaluated were the Simple Rainfall Model, the *ACRU*-Thompson yield model and the CANEGRO-DSSAT growth model.

When operating a model, certain inputs are required in order to predict the crop yield. The inputs required by the proposed models and the preparation of these inputs are addressed in this chapter. A description of the modelling strategies adopted in predicting the MSA yields is also given. The observed yield database against which model simulations are verified, is briefly discussed. Finally, the results of the model output verifications are presented.

6.1 Model Inputs

Rainfall inputs are required by all of the cane yield models. In addition, the *ACRU*-Thompson and CANEGRO-DSSAT models require inputs relating to other climatic variables, soils and growth cycles. Climate inputs to the various models were prepared for the period 1985 to 1995, in order to predict the yields for the 1988 to 1995 harvest seasons. Observed yields were readily available for these seasons, thus allowing for comparative analyses of simulated and observed yields over this period.

6.1.1 Climate

6.1.1.1 Rainfall

The *ACRU*-Thompson and CANEGRO-DSSAT models require daily rainfall input while the Simple Rainfall Model requires monthly rainfall totals. There are two main problems associated with deriving the required rainfall data sets: first, the need to ensure that the available rainfall records are complete and without missing data, and secondly, the extrapolation of the records from point values (rainfall stations) to values representing the rainfall occurring on surrounding farms, i.e. to develop a spatial representation of daily rainfall across the MSA. The first problem was addressed by applying the technique of Inverse Distance Weighting (IDW). This technique makes use of data from surrounding rainfall stations to infill (patch) data that are missing at a station of interest. The data from the surrounding stations are weighted according to their distance and mean annual precipitation, relative to the station whose record is being infilled. Rainfall stations that are closer and more similar in their value of MAP, relative to the station of interest, are given a greater weighting. The IDW equation is as follows :

$$r_s = \frac{\dfrac{R_d r_1}{R_1 d_1^2} + \dfrac{R_d r_2}{R_2 d_2^2} + \dfrac{R_d r_3}{R_3 d_3^2} + \ldots}{\dfrac{1}{d_1^2} + \dfrac{1}{d_2^2} + \dfrac{1}{d_3^2} + \ldots}$$

where r_s = synthesised daily rainfall (mm) for the station whose missing values are being infilled

$r_{1,2,3}$ = actual rainfall (mm) recorded at surrounding stations

$d_{1,2,3}$ = distance (degrees decimal) of surrounding stations relative to the station whose missing values are being infilled

$R_{1,2,3}$ = MAP (mm) of grid cells containing surrounding stations

R_d = MAP (mm) of the grid cell containing the station whose record is being infilled.

The grid cell MAP values are taken from the national MAP grid developed by Dent, Lynch and Schulze (1989). This MAP grid allows for convenient assessment of the MAP of surrounding stations relative to the station being patched. It also has the advantage that the MAP values contained in the grid were calculated from regression equations derived from rainfall data of stations having sufficiently long and uniform record lengths (Dent *et al.,* 1989). The rainfall stations used in this dissertation would not necessarily have had records appropriate for producing representative estimates of MAP. The predictors of the regression equations used to calculate MAP included altitude, distance from sea, aspect, terrain roughness and direction of prevailing rainbearing winds (Dent *et al.*, 1989). The MAP grid has a resolution of one minute by one minute of a degree of latitude and longitude, i.e. approximately 1.6 km by 1.6 km.

In order to address the problem of developing a spatial representation of daily rainfall across the MSA, each farm in the MSA was associated with a representative rainfall station known as a "driver station". Figure 6 shows the location of rainfall stations in and around the Eston MSA, with some of these stations being driver stations and the others being stations used for patching (see above). Figure 6 also contains the gridded image of MAP extracted from the national grid of Dent *et al.* (1989). The data from the driver stations were used to represent the rainfall occurring on the various farms. The point rainfall estimates of the stations were adjusted to be more representative of their associated farms through the application of adjustment factors. An adjustment factor was developed for each farm and its associated driver station. Adjusted daily farm rainfall, r_a, was calculated as follows:

$$r_a = r_s * k$$

where r_a = adjusted daily farm rainfall (mm)

r_s = driver station point estimate of daily rainfall (mm)

k = adjustment factor

where $k = R_a / R_s$

R_a = average MAP of grid cells within the farm (mm)

R_s = MAP of the grid cell containing the driver station (mm) .

Grid cell MAP values were taken from the national MAP grid (Dent *et al.*, 1989) for reasons similar to those discussed above in the procedures for infilling of missing rainfall data.

The development of a spatial representation of rainfall across the MSA (at farm scale) was only required for the *ACRU*-Thompson and CANEGRO-DSSAT models, as the SRM was applied using one representative rainfall station for the entire MSA. The station used in this regard was driver station 5 (Figure 6), located centrally in the MSA. No adjustments were made to the station data for spatial representation of rainfall. This is typical of the manner in which the SRM would be applied in practice.

Table 3 contains general information relating to the driver stations used in this project.

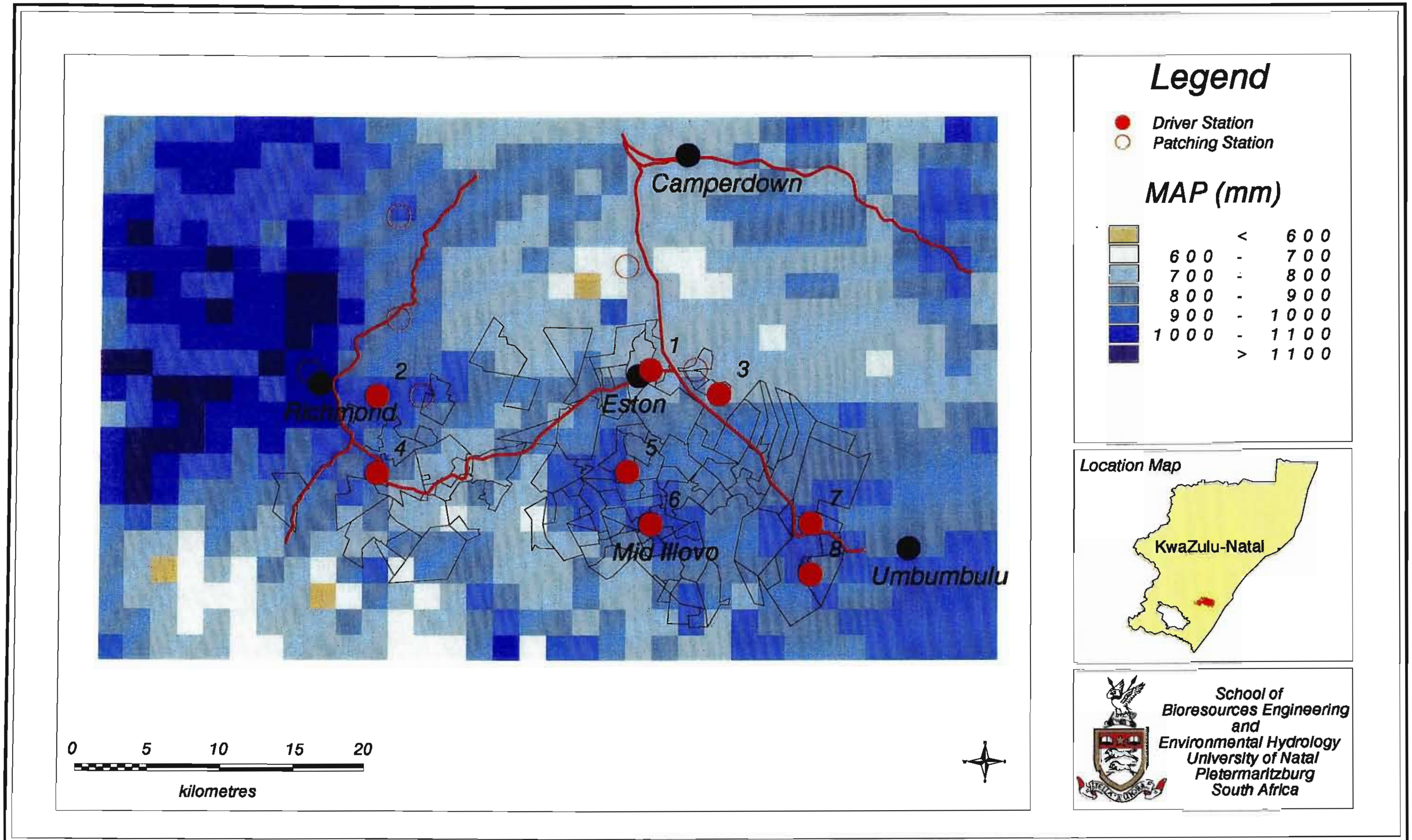


Figure 6 Rainfall stations used in the development of farm rainfall data sets. Gridded rainfall after Dent, Lynch and Schulze (1989).

Table 3 General information relating to rainfall driver stations selected to represent rainfall in the Eston Mill Supply Area

No.	Name	Source	Latitude (° ')	Longitude (° ')	Altitude (m)	MAP (mm)	Extent Infilled (%)
1	Eston	SAWB[1]	29 52	30 31	792	704	1.2
2	Ivanhoe	ISCW[2]	29 53	30 19	860	730	3.8
3	Beaumont	SASEX[3]	29 53	30 32	732	821	N/A
4	Little Harmony	ISCW[2]	29 56	30 19	810	883	0.7
5	Stoke Mid-Illovo	ISCW[2]	29 56	30 30	670	894	0.0
6	Mid-Illovo	SAWB[1]	29 58	30 31	716	932	4.5
7	Powerscourt	SASEX[3]	29 58	30 38	631	1015	N/A
8	Umbumbulu	SASEX[1]	30 00	30 38	637	952	0.0

NOTES: [1] From the database housed at the Computing Centre for Water Research
[2] Obtained directly from the Institute of Soil, Climate and Water (ISCW)
[3] Obtained directly from SASEX

6.1.1.2 Temperature

The *ACRU*-Thompson and CANEGRO-DSSAT models require inputs of daily values of maximum and minimum temperatures. The temperature values for each of the farms in the MSA were determined using a spatial temperature estimation technique (Schulze and Maharaj, 1998). This technique involves consideration of the proximity and relative altitude of the available climate stations measuring temperature in relation to the point of interest (i.e. a farm, in this case). Weightings related to proximity and altitude are assigned to the stations surrounding each farm, with these weightings being used to form a ranking of station suitability for each farm. The observed data from the most suitable station are then extracted to develop the farm temperature data set, with any adjustments for altitudinal differences being made according to regional temperature lapse rates given in Schulze (1995). For any days of missing temperature data in the station record, data from the next most suitable station are extracted and corrected in order to infill the record. If necessary, the process of data extraction is continued successively through all stations, until observed data are found which can represent temperatures on that day. A more detailed description of the spatial temperature estimation technique is given in Appendix A. Figure 7 shows the climate stations whose data were considered for estimation of farm temperatures. General information relating to these stations appears in Table 4.

Table 4 General information relating to climate stations used in the development of temperature inputs at farm level in the Eston Mill Supply Area

No.	Name	Source	Latitude (° ')	Longitude (° ')	Altitude (m)
1	Little Harmony	ISCW[1]	29 56	30 19	810
2	Stoke, Mid-Illovo	ISCW[1]	29 56	30 30	670
3	Aireley Farms	ISCW[1]	29 48	30 35	590
4	Natal Est., Thornville	ISCW[1]	29 45	30 24	860
5	Camperdown - Eston	ISCW[1]	29 52	30 33	697
6	Hillcrest	ISCW[1]	29 48	30 35	457
7	Sevontein	ISCW[1]	29 45	30 08	1375
8	Sapekoe Estate	ISCW[1]	29 56	30 09	1200
9	Baynesfield Estates	ISCW[1]	29 45	30 20	914
10	Double Diamond	ISCW[1]	29 48	30 30	610
11	Powerscourt	SASEX[2]	29 58	30 38	631
12	Beaumont	SASEX[2]	29 53	30 32	732

NOTES: [1] Obtained directly from the ISCW

[2] Obtained directly from SASEX

6.1.1.3 Solar radiation

The CANEGRO-DSSAT model requires daily values of solar radiation in order to predict sugarcane yields. Direct measurements of solar radiation in the Eston MSA were not available for the period of interest. There were, however, measurements of sunshine duration available at several climate stations in and around the MSA. These measurements were converted to values of solar radiation using established equations (Ångström, 1924). Two stations within the MSA were selected as driver stations and their data used to develop solar radiation inputs for each

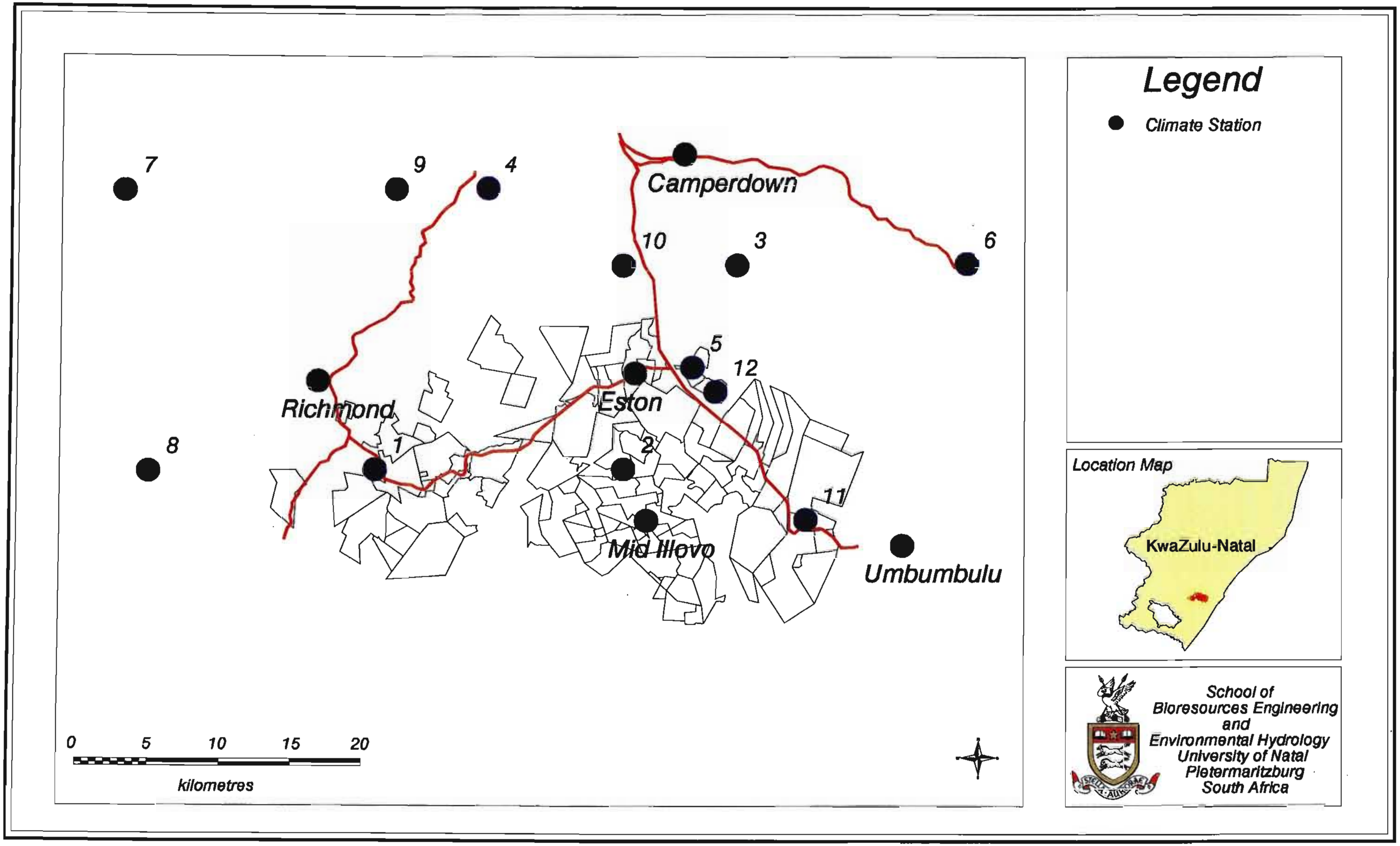


Figure 7 Climate stations used in the estimation of farm temperatures

of the farms in the MSA. The data from other available climate stations were used to infill the missing records of these driver stations, with no adjustments to their data being made. Figure 8 shows the climate stations used in deriving solar radiation inputs, including both driver stations and infilling stations. General information relating to the driver stations appears in Table 5.

Table 5 General information relating to driver stations used in the development of solar radiation inputs at farm level in the Eston Mill Supply Area

No.	Name	Source	Latitude (° ')	Longitude (° ')	Altitude (m)	Extent Infilled (%)
1	Little Harmony	ISCW[1]	29 56	30 19	810	28.3
2	Powerscourt	SASEX[2]	29 58	30 38	631	N/A

NOTES: [1] Obtained directly from the ISCW

[2] Obtained directly from SASEX

6.1.1.4 Reference potential evaporation

The *ACRU*-Thompson and CANEGRO-DSSAT models require inputs of daily potential evaporation to calculate the total evaporation (i.e. transpiration plus evaporation from the soil surface) of a cropped area. Potential evaporation represents the atmospheric demand for evaporation on a day and it is estimated from a reference measurement or equation. The evaporation from a standard Class A evaporation pan is considered the reference for potential evaporation in the *ACRU*-Thompson model, while a short grass surface is used as the reference in the CANEGRO-DSSAT model. In the present study the reference potential evaporation of both models was estimated using temperature (and solar radiation in the case of CANEGRO-DSSAT) based equations. The Linacre (1991) daily temperature based equation for A-pan equivalent potential evaporation was used in the *ACRU*-Thompson model, while the Ritchie modification of the Priestley-Taylor equation was used to derive the reference potential evaporation in the CANEGRO-DSSAT model. The preparation of the inputs required by these equations has been discussed in the preceding sections.

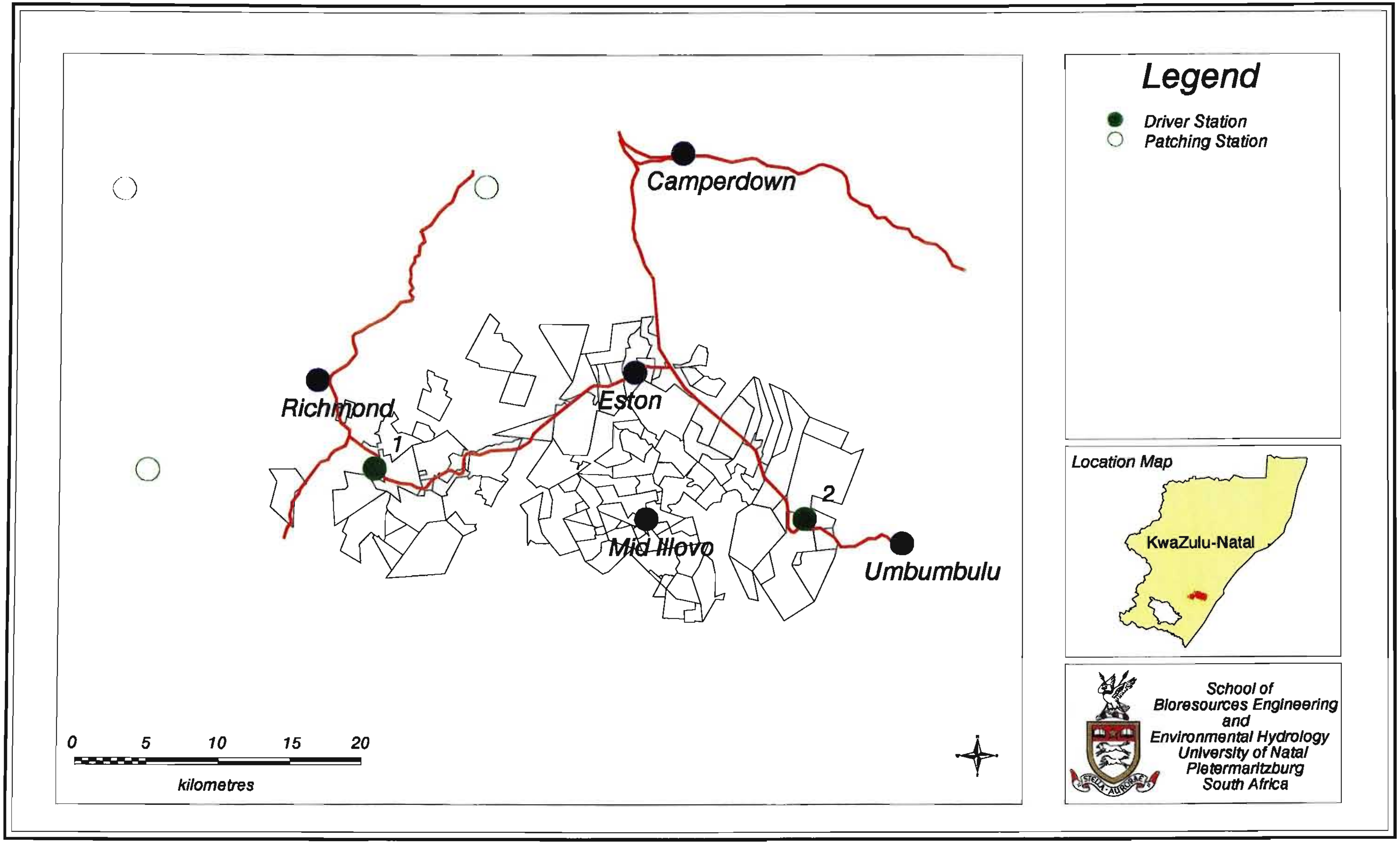


Figure 8 Climate stations used in the development of farm solar radiation inputs

6.1.2 Soils

6.1.2.1 Inputs required

The SRM requires no soils inputs. The key soils inputs required by the *ACRU*-Thompson and CANEGRO-DSSAT models include the thickness of the respective soil horizons and the soil water content of these horizons at critical levels of soil water retention. These levels are at the permanent wilting point, the drained upper limit (formerly known as field capacity) and saturation. The *ACRU*-Thompson model requires that a topsoil and subsoil horizon be specified, while the CANEGRO-DSSAT model requires that multiple soil layers of approximately 100 to 150 mm be specified. For crop yield modelling, the total depth of soil within the active root zone is considered. A further key soil input required by the CANEGRO-DSSAT model is the root growth factor, which is specified for each of the soil layers. This factor, which can range between 0 and 1, controls the amount of root growth in each of the layers.

6.1.2.2 Translation of available soils information into model inputs

Two sources of soils information were available for the development of soils inputs for the farms in the Eston MSA. These sources were the soil parent material (PM) maps which are available for large areas of the sugar industry, as well as the Land Type (LT) maps developed by the Institute of Soil Climate and Water, which are available for many areas across South Africa. Land Types are areas of relatively uniform agricultural potential, where this potential is based partly on the soils occurring within a Land Type. Soil inputs were derived from both sources of soils information, enabling an evaluation to be made of their suitability for crop modelling. The soil PM and LT maps for the Eston MSA are shown in Figures 9 and 10. General characteristics of the soils contained in each Land Type are presented in Table 6. Both sources of soils information required "translation" into the inputs needed by the models. The methodologies employed in this regard are explained below:

<u>Soil parent material information</u>

In the case of the PM maps, soil inputs were derived by first predicting the resulting soil types and depths likely to be associated with each parent material. These predictions were made according

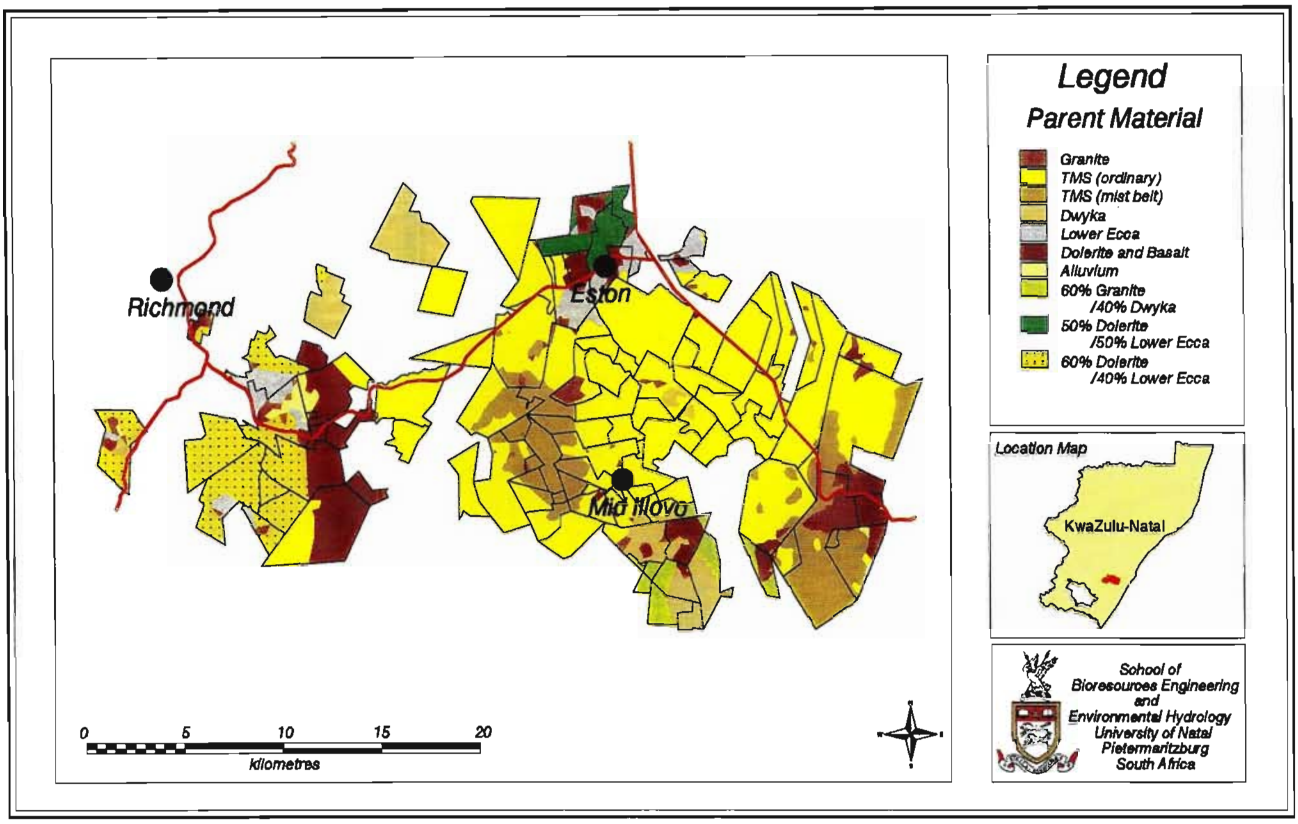


Figure 9 Soil parent material map for the Eston Mill Supply Area

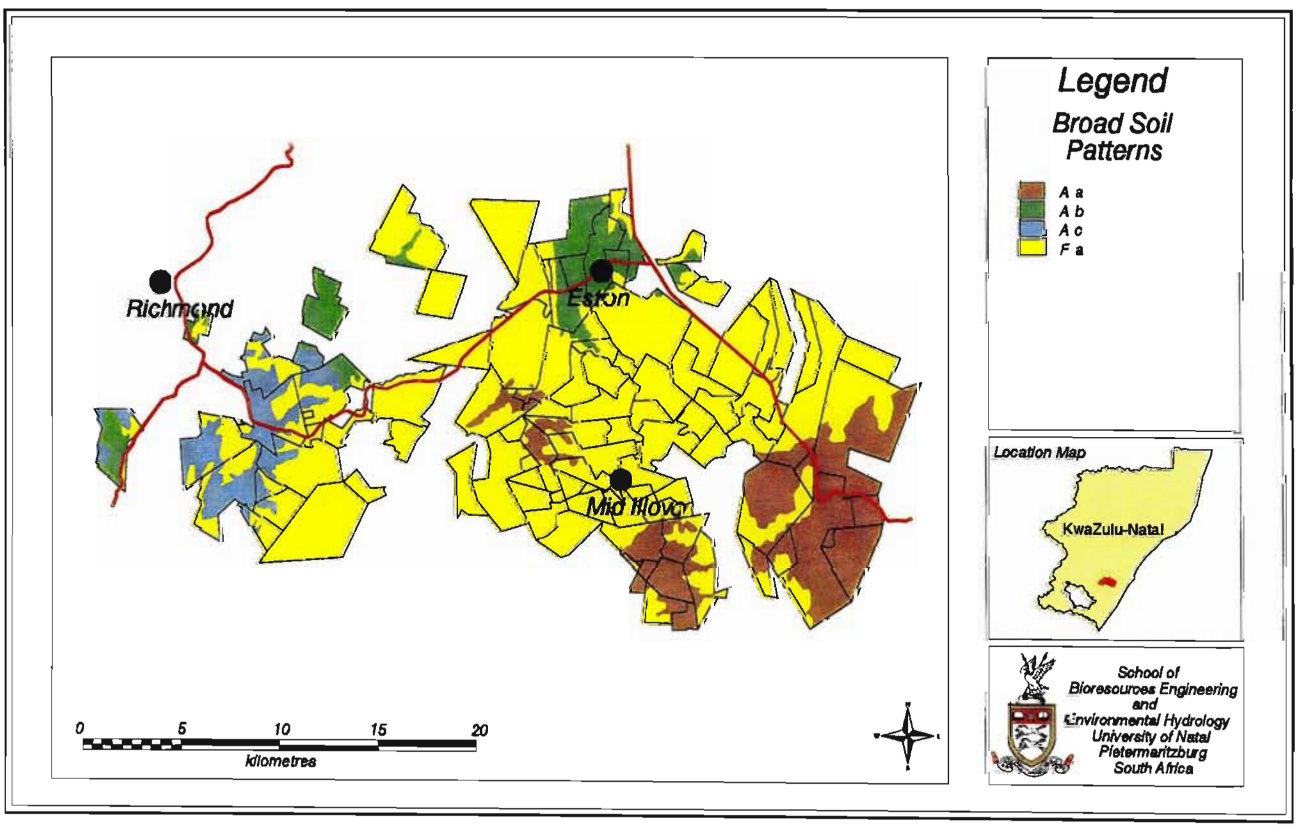


Figure 10 Land Type map for the Eston Mill Supply Area

Table 6 General soil characteristics associated with various Land Types found in the Eston Mill Supply Area

Land Type	General Soil Characteristics
Aa	Freely drained soils with red-yellow apedal and humic horizons.
Ab	Freely drained, dystrophic and/or mesotrophic soils with red apedal horizon.
Ac	Freely drained, dystrophic and/or mesotrophic soils with red-yellow apedal horizon.
Fa	Glenrosa and/or Mispah forms (other soils may occur), lime rare or absent in the entire landscape.

to a set of working rules developed from expert opinion (Mann, Meyer and Hellmann, 1997), where these rules took into account the soil PM, slope position, slope gradient and MAP in the prediction of soil type and depth. In order to implement the rules, grids of each of the above components were created at a 200 m resolution using a GIS. The portion of the MAP grid of Dent *et al.* (1989) extracted for this purpose, was resampled from the original resolution of approximately 1600m. A computer program incorporating the working rules was then written to generate the predicted soil type and depth grid values, with the soil PM, slope position, slope gradient and MAP grid values forming input to this program. An example of the working rules is given in Table 7 for Table Mountain Sandstone (TMS - ordinary) parent material. The working rules for the remaining predominant parent materials are found in Appendix B. As the working rules in Table 7 and Appendix B were not fully inclusive, further assumptions were required regarding the soil types and depths expected on certain parent materials and combinations of parent materials. These assumptions are given in Table 8. The areas of these parent materials and combinations of parent materials made up a relatively small proportion of the total MSA.

All soils contained in the South African soil classification systems (Macvicar *et al.*, 1977 and SCWG, 1991) have been assigned the relevant model inputs through previously developed generic relationships (Schulze, Hutson and Cass, 1985). Thus, the required soil inputs of the models could be developed for the various parent materials occurring in the MSA. An average set of model inputs was determined for each farm by calculating the proportions of areas having unique soil type/depth combinations and performing an area weighted average of the associated model inputs.

Table 7 Example of working rules used in the prediction of soil type and depth in the Eston MSA, based on the consideration of soil parent material (TMS - ordinary), slope position, slope gradient and MAP (Mann *et al.*, 1997)

Parent Material	Slope Position	Slope Gradient (%)	MAP (mm)	Soil Type	Topsoil Depth (m)	Effective Subsoil Depth (m)
TMS (ordinary)	Topslope	5 - 12	< 800	Gs 14	0.3	0.3
			800 - 900	Gs 14	0.3	0.4
			> 900	Gs 17	0.5	0.4
		< 5	< 800	We 21 (40%) / Gs 14 (60%)	0.4 / 0.3	0.2 / 0.3
			800 - 900	We 21 (40%) / Gs 14 (60%)	0.4 / 0.3	0.2 / 0.3
			> 900	Gs 17	0.5	0.4
	Midslope	> 12	< 800	Cf 21 (60%) / Gs 14 (40%)	0.3 / 0.4	0.3 / 0.3
			800 - 900	Cf 21 (60%) / Lo 21 (40%)	0.3 / 0.4	0.4 / 0.4
			> 900	Cf 21 (60%) / Lo 21 (40%)	0.3 / 0.4	0.4 / 0.4
		5 - 12	< 800	Cf 21	0.3	0.3
			800 - 900	Lo 21	0.3	0.3
			> 900	Lo 21	0.3	0.3
		< 5	< 800	Lo 21	0.3	0.3
			800 - 900	Lo 21	0.3	0.3
			> 900	Lo 21	0.3	0.3
	Bottomslope	5 - 12	< 800	Kd 14	0.4	0.2
			800 - 900	Kd 14	0.4	0.2
			> 900	Kd 14	0.4	0.2
		< 5	< 800	Ka 10	0.5	0.0
			800 - 900	Ka 10	0.5	0.0
			> 900	Ka 10	0.5	0.0

Table 8 Assumptions made regarding the soil types and depths expected on certain parent materials and parent material combinations

Parent Material / Parent Material Combination	Assumption
Granite	Treat as TMS (ordinary)
Alluvium	Assume soil type is Du10 (only soil having alluvium layer) topsoil depth = 0.3m, effective subsoil depth = 0.8m
Granite (60%) / Dwyka (40%)	Treat as TMS (ordinary)
Dolerite (50%) / Lower Ecca (50%)	Treat as Lower Ecca
Dolerite (60%) / Lower Ecca (40%)	Treat as Lower Ecca

Land Type information

For translation of the LT information, all the soil series contained in each Land Type have had model input values assigned to them through use of the above mentioned generic relationships (Schulze *et al.*, 1985). To determine an average set of Land Type derived model inputs for each farm, use was made of the AUTOSOIL computer program (Pike and Schulze, 1995). This program requires that the proportions of areas having unique soil type/depth combinations be specified for each Land Type. This information is then used by the program to calculate an area weighted average of the model inputs associated with those soil type/depth combinations. The proportional areas of Land Types occurring on the various farms must further be specified to allow for an area weighted average set of model inputs to be determined for each farm.

For the CANEGRO-DSSAT model, soils inputs (derived from both PM and LT information) were duplicated in the layers contained in the topsoil and subsoil horizons, as inputs were not available for each of these layers (only one generic relationship was available for a top - or subsoil horizon for the prediction of model inputs). Values of the root growth factor were estimated for each of the horizon layers, by considering the location of that layer within the soil profile and its rooting characteristics. Deeper layers, and layers known to hinder root growth, were given lower values of the root growth factor.

6.1.2.3 Evaluation of model inputs derived from different sources of soils information

The model inputs derived from the two available sources of soils information were evaluated to assess their differences. Assessing these differences was considered necessary for selecting the more appropriate source of soils information for application in cane yield modelling. Also considered necessary, was a comparison of the resulting yield simulations derived from the two sets of soils inputs. This comparison is carried out in the results section of this chapter (Chapter 6.4).

The two sets of model inputs were compared by calculating and mapping the respective soil total available moisture (TAM) values for the farms in the MSA. Figures 11 and 12 show TAM based on PM and LT derived soils inputs, respectively. The maps indicate that the two sources of soils information mostly give rise to similar TAM values, except in parts of the central and eastern areas of the MSA, where PM derived soils inputs result in higher TAM. The range in TAM values extended from less than 20mm for both LT and PM soils, to over 120mm in the case of PM soils. The boundaries of the PM and LT maps and the values of the subsequently derived soil TAM are similar, and appear to indicate a strong influence of the soil parent material on the formation and properties of the soils in the area. Not all farms indicated in the maps are under cane (including the single farm having a TAM less than 20mm). These non-cane farms are indicated later in Figure 15 under the discussion of the observed cane yield database used in this study.

6.1.3 Growth Cycles

It is known that growth cycles affect the yields of sugarcane crops (Hellmann, 1993). Bearing this in mind it was considered necessary to investigate the growth cycles practised in the Eston MSA, with a view to incorporating cycles in the modelling framework. The investigation of growth cycles is divided into sections relating to classification of cycles, the effects of cycles on yields, the proportions of cycles cut in the Eston MSA and the modelling of cycles.

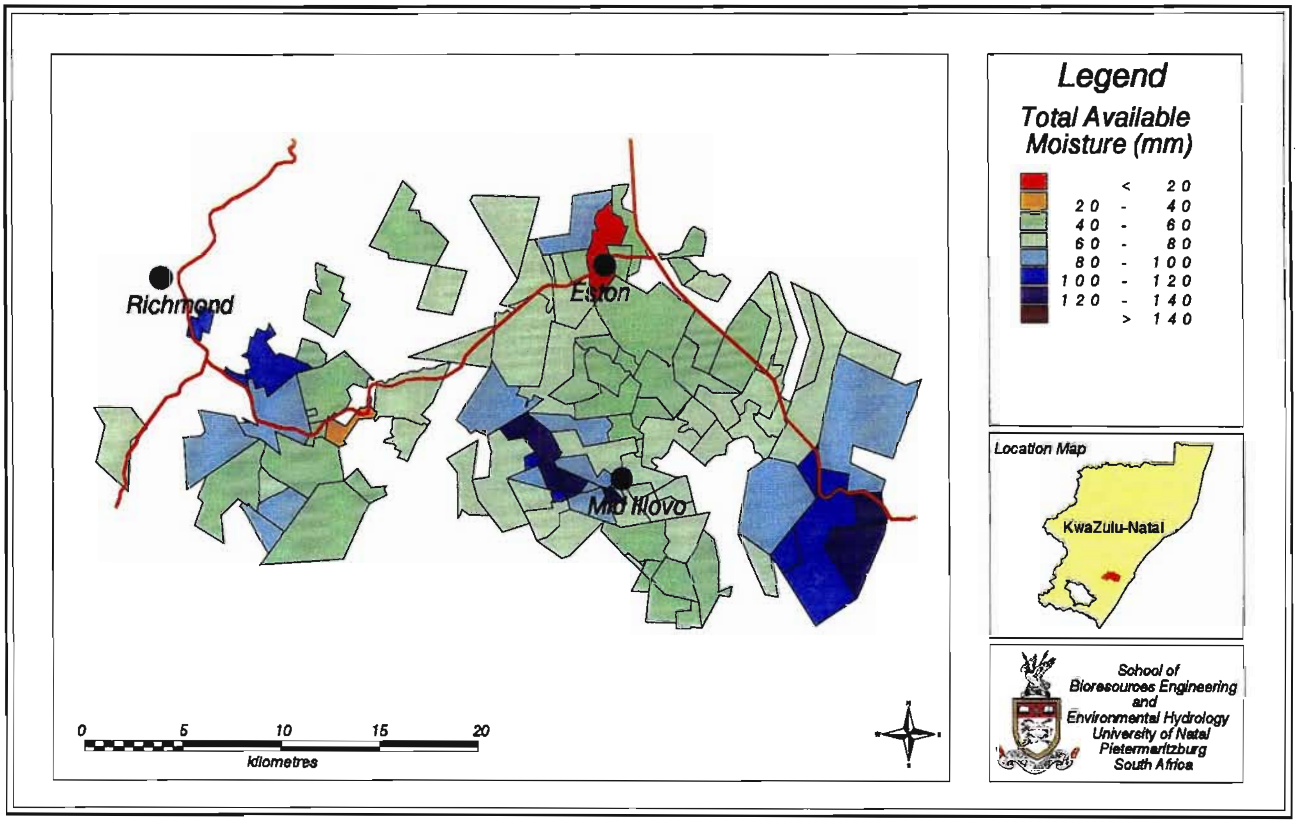


Figure 11 Total plant available soil moisture (TAM) derived from soil parent material information

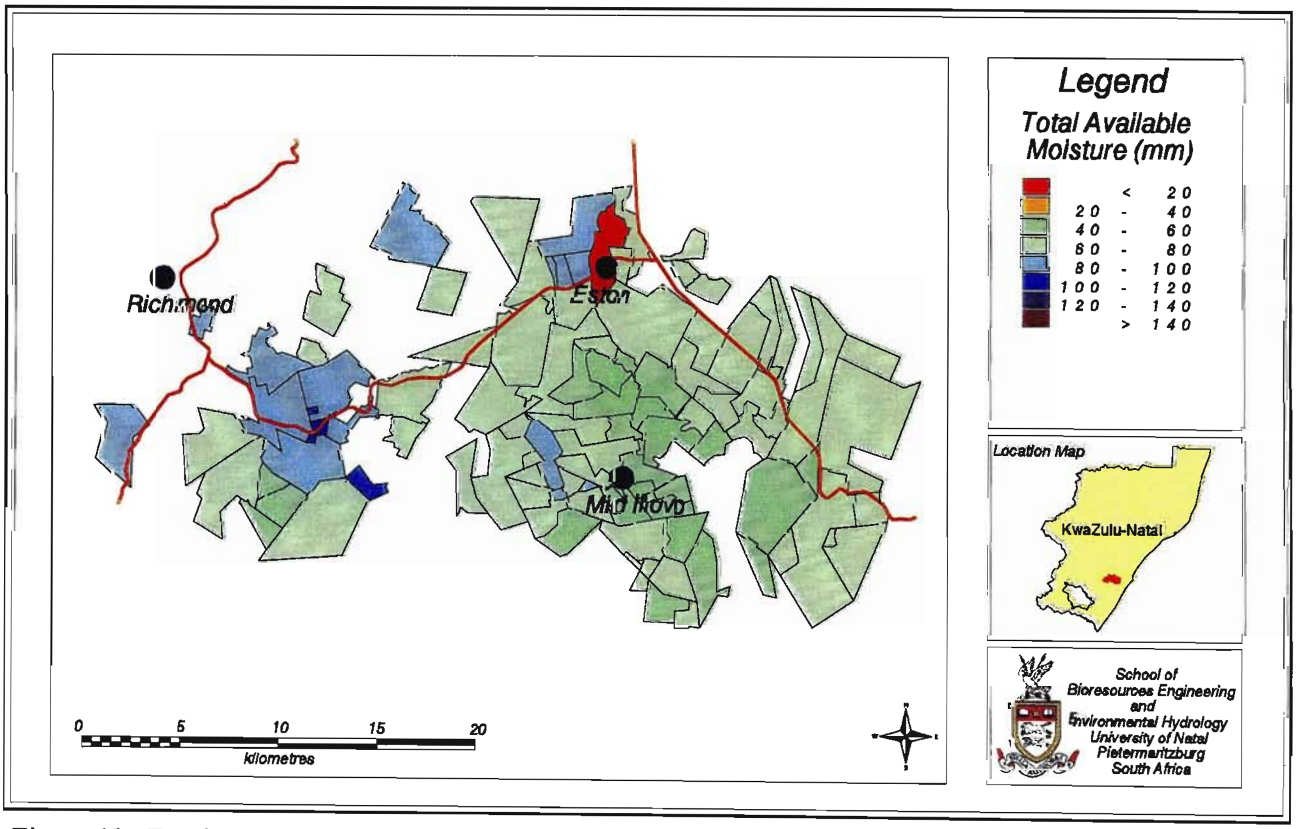


Figure 12 Total plant available soil moisture (TAM) derived from Land Type information

6.1.3.1 Classification of growth cycles

As cane in the Eston MSA is typically harvested throughout the period from April to December, it was necessary to categorize cane crops into discrete growth cycles that could be represented in the modelling framework. Growth cycles were classified according to a system developed in a previous study in the Eston MSA (Hellmann, 1993). The seasonal timing of growth commencement and harvest, as well as the length of time a crop grows for, including the number of summer seasons experienced, were identified as being important in classifying the growth cycles practised. In the study by Hellmann (1993), 11 growth cycles were classified according to the above criteria, with codes being assigned to each of them. The timing of the seasons was defined as follows:

- Autumn (A) : 1 February to 30 April
- Winter (W) : 1 April to 31 July
- Spring / Summer (S) : 1 August to 31 December .

The assigning of codes to the various cycles is illustrated by way of the following examples:

- PS-W2S : This indicates a plant crop (first letter P), planted in spring / summer (second letter S), harvested in winter (third letter W) and growing through two summers (2S).
- RW-S1S : This indicates a ratoon crop (R), commencing growth in winter (W), harvested in spring / summer (S) and growing through one summer (1S).

The 11 cycles and the period of time over which they span are shown in Figure 13. The time span of the cycles is indicated by way of a time line. The months relating to the start and end of seasons are shown on the horizontal axis, and the cycles are represented by lines which span from the beginning of the start season to the end of the harvest season. The crops would not necessarily grow for the entire time represented by the lines, but would begin and end within the relevant season intervals which are indicated below the horizontal axis. The longest cycle in the Eston MSA is the PA-S2S cycle, with a maximum possible length of 35 months, while the shortest cycles are the RS-S1S and RW-W1S cycles, having maximum possible lengths of 17 months.

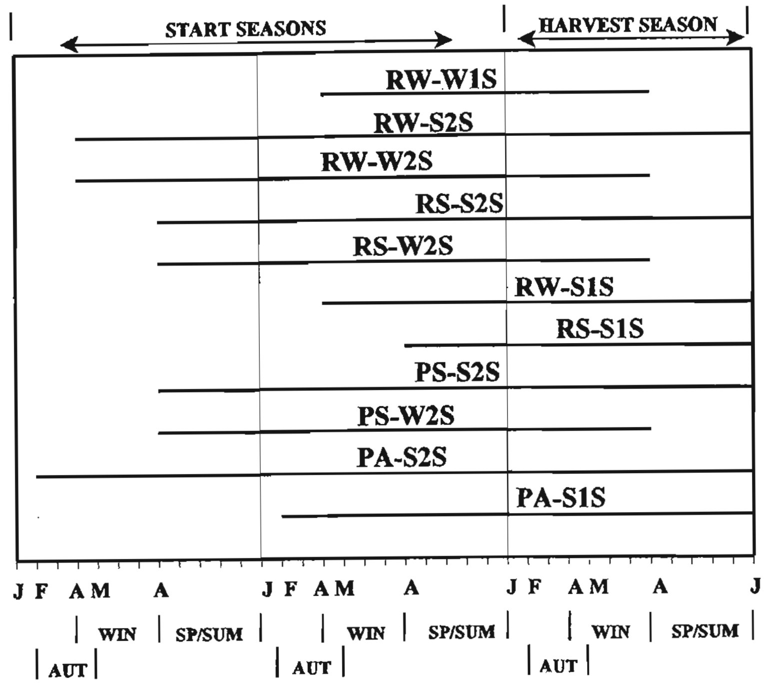


Figure 13 Diagram indicating the 11 growth cycles practised in the Eston Mill Supply Area and the time periods over which they span

6.1.3.2 Effect of growth cycles on observed yields

When incorporating growth cycles into a modelling framework, it is important to have an understanding of how growth cycles affect crop yields. To this end, an investigation was carried out on observed yields extracted from the Field Record System (FRS) for a representative farm in the MSA. Each crop harvested was classified into one of the 11 growth cycles identified in the Eston area, with data being available from the selected farm for 8 of these cycles. The yields of crops assigned to each growth cycle were averaged over the period considered, this being from 1986 to 1993. Yields were also averaged for the individual years of 1987, 1991 and 1992, these corresponding to average, high and low yielding harvest seasons. This was done to assess whether the relative yields of the various cycles were consistent regardless of the season. Figure 14 shows the average yields calculated for each growth cycle. The graph indicates that cycles do affect yield. These effects are related to the influence of planting versus ratooning, the seasonal timing of ratooning and whether a crop grows through one or two summers. Plant crops generally produce yields which are higher than those of ratoon crops (eg. PS-S2S vs RS-S2S), summer ratoons generally yield higher than corresponding winter ratoons (eg. RS-W2S vs RW-W2S) and crops growing over two summers generally produce higher yields than those growing over a single

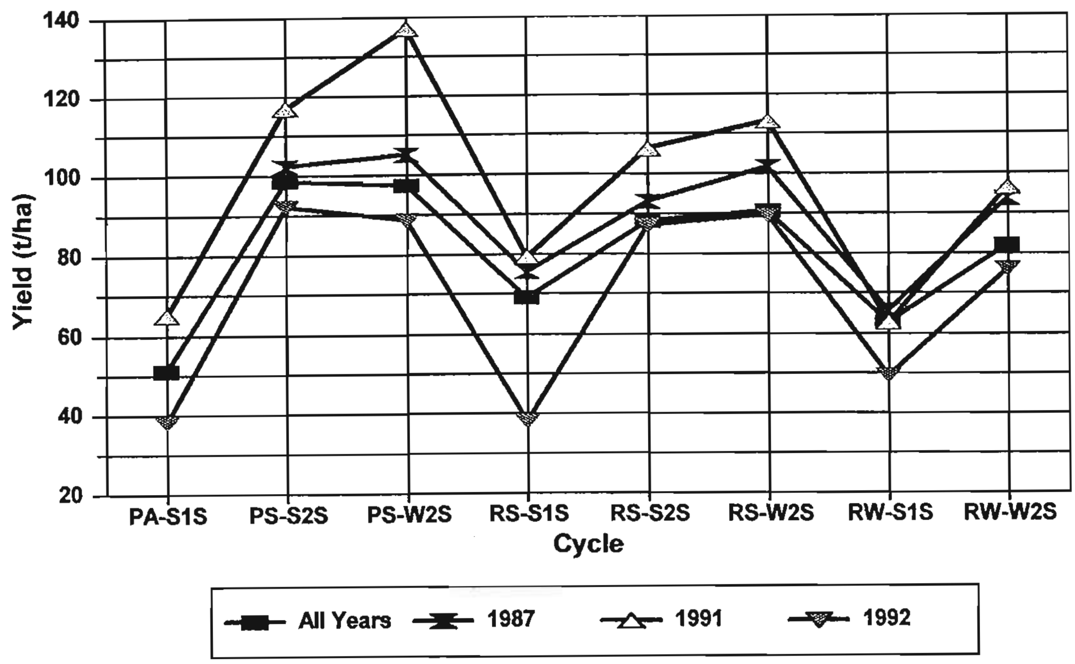


Figure 14 Plot of mean observed (field scale) yield versus growth cycle for a selected farm in the Eston Mill Supply Area, for all and selected years in the period 1986 to 1993

summer (eg. RS-S2S vs RS-S1S). The curves for the individual years generally follow the trend of the curve for all years, although their actual values differ substantially in some cases. The consistency in yield-cycle trends over different seasons simplifies the modelling of cycles.

6.1.3.3 Average proportions of growth cycles in the study area

When incorporating growth cycles into a modelling strategy, the proportions in which the various cycles are cut should be assessed for the study area. The average proportions of cycles cut in the Eston MSA were assessed from the FRS data of 11 representative farms in the area (for the period 1986-1993). The proportions were calculated as percentages based on the area cut under each cycle, and are tabulated in Table 9.

There are three cycles of approximately equal proportions which constitute the majority of the area harvested (total of 60.5 %). These cycles are the RS-W2S, RW-W2S and RS-S2S cycles. The next most common cycle is the PS-S2S cycle which constitutes 11.3% of the area. All other cycles constitute less than 10% of the area cut.

Table 9 Average proportions of growth cycles cut (by area) for selected farms in the Eston Mill Supply Area over the period 1986 to 1993

Cycle	Proportion (%)
PA-S1S	1.7
PA-S2S	0.1
PS-W2S	5.3
PS-S2S	11.3
RS-S1S	6.7
RW-S1S	8.2
RS-W2S	22.8
RS-S2S	17.7
RW-W2S	20.0
RW-S2S	5.4
RW-W1S	0.8

6.1.3.4 Modelling of growth cycles

Growth cycles differ in their length and in the seasonal climatic environment experienced by the crop. Rainfall, temperature, solar radiation and evaporation all form part of the climatic environment and are reflected in the yield predictions of the *ACRU*-Thompson (no solar radiation) and CANEGRO-DSSAT models. Only monthly rainfall is reflected in the yield predictions of the SRM. Growth cycles are input in the yield models through varying the start and harvest dates of the crops. The *ACRU*-Thompson model was modified to further improve its ability to represent growth cycles through the introduction of dynamic equations relating crop water use to daily temperature (see Chapter 4.2.3). Many of the growth relationships contained in the CANEGRO-DSSAT model are related to temperature, thus rendering this model well suited to the representation of growth cycles.

6.1.4 Other inputs

Inputs required by the yield models that have been discussed thus far relate to climate, soils and growth cycles. A number of other inputs are required by the yield models. These inputs are related to processes not directly linked to crop yields, or alternatively, are inputs that are treated as being relatively invariant (in a study such as this), and thus do not receive a great deal of attention during input preparation. Default, or literature derived values, are often used to develop the latter inputs. In the case of the *ACRU*-Thompson model, inputs include those relating to the generation of runoff and to crop characteristics (eg. canopy interception loss, root distribution, point of stress commencement). Other inputs to the CANEGRO-DSSAT model include inputs relating to the cultivar used and to planting (eg. row spacing, plant population). Both models also require that appropriate methods be selected for the simulation of various processes represented within the models.

6.2 Modelling Strategy

A modelling strategy was required to account for the growth cycles practised in the Eston MSA. In the case of the *ACRU*-Thompson and CANEGRO-DSSAT models, it was assumed that crops falling into each of the 11 growth cycle categories would be harvested in all of the 8 years of simulation. Practical constraints precluded the simulation of continuous crops where one crop followed another, and where all 11 cycles are harvested in each year. It was assumed instead, that new simulations would be performed for each year of harvest, and that the start dates of the various cycles would be determined by subtracting the ages of the cycles (in months) from their harvest dates. All ratoon crops were assumed to be first ratoons. The start and harvest dates assumed for the various cycles and their corresponding ages, are contained in Table 10. These dates represent the middle of the start/harvest seasons indicated in Figure 13. Yields were simulated for 11 crops (cycles) for each of the farms in the MSA. The mean yield of each farm was determined by averaging the weighted yields of the 11 cycles, where weightings were assigned to cycles according to their proportion of the total area cut (see Table 9).

The modelling strategy for the SRM model involved first determining the average area weighted cycle length, which was found to be 22 months. For this model to produce the best results, the MAP (of rainfall driver station 5, Figure 6) and average observed MSA yield were determined

Table 10 Start and harvest dates and corresponding ages of growth cycles simulated in the Eston Mill Supply Area

Number	Cycle	Start Date	Harvest Date	Age (months)
1	PA-S1S	1 Feb	16 Oct	20.5
2	PA-S2S	1 Feb	16 Oct	32.5
3	PS-W2S	16 Oct	1 Jun	19.5
4	PS-S2S	16 Oct	16 Oct	24.0
5	RS-S1S	16 Oct	16 Oct	12.0
6	RW-S1S	1 Jun	16 Oct	16.5
7	RS-W2S	16 Oct	1 Jun	19.5
8	RS-S2S	16 Oct	16 Oct	24.0
9	RW-W2S	1 Jun	1 Jun	24.0
10	RW-S2S	1 Jun	16 Oct	28.5
11	RW-W1S	1 Jun	1 Jun	12.0

over the eight year period of yield prediction, this being1988 to 1995. These statistics (values of 794 mm and 83 t/ha respectively) were then used along with the average cycle length to determine the mean rate of yield accumulation. This rate was found to be 5.70 t/ha/100mm. Yields were then predicted by the SRM according to the methodology described in Chapter 4.1.

6.3 Observed Yield Database Used in Verifications

A database of observed farm yields was required for comparison with model simulated yields. The database used in comparisons was based on the South African Sugar Association Annual Survey of Area Under Cane and Cane Production. Data were readily available in electronic form for the period 1988 to 1995. The greater availability of data over this period determined the period of yield simulation, as it was not considered feasible to expend time and effort in obtaining a longer record of observed yields in an (electronic) format suitable for comparisons. The tonnages of cane contained in this database were those measured at the mill. Yields of cane (tonnage per hectare) were calculated for each year of harvest and included in the database. Where yields could not be calculated owing to a lack of data, the area under cane (rather than cane tonnage) was usually found to be missing. The calculated yields were checked to ensure that they fell within a range normally encountered in the Eston MSA for average farm yields, this range extending from 30 to 175 t/ha. A small number of data were found to fall outside of this range (below 30t/ha). These suspect data were then omitted from further analyses. Figure 15 shows the number of years of missing observed farm yield data. Most farms were found to have one year or less of missing data, while a minority of farms were found to have as many as six years of missing data out of the eight years considered. For farm-by-farm comparisons of simulated and observed yields, where it is desirable to have as many years of data available as possible, it was decided that farms having more than four years of record missing would not be considered. It should be noted that all data obtained were associated with quota numbers and that these quota numbers were assigned to individuals, some of whom own more than one farm. In these cases, separate data were not available for each farm, resulting in the available data being associated with all farms in the ownership of those particular individuals. There were a total of 68 quota numbers associated with the 85 farms considered in comparative (simulated vs observed) analyses.

6.4 Results

Cane yields were simulated on a farm-by farm basis for the 1988 to 1995 harvest season using the *ACRU*-Thompson, CANEGRO-DSSAT and SRM models. In the first part of this section, simulation results are analysed in terms of their accuracy relative to observed yields. The simulations based on the two sets of soils inputs are then compared to each other in the second part. In the third part, comments are made regarding the suitability of the models for yield forecasting. The simulation of growth cycles is then discussed in the final part of this section.

6.4.1 Model simulations relative to observed yields

<u>Farm scale simulations</u>

In order to assess the *ACRU*-Thompson and CANEGRO-DSSAT yield simulations at farm scale, the mean simulated and observed yields over the period of yield simulation were calculated for each of the considered farms in the MSA. The percentage differences between the mean simulated and observed yields were then calculated and mapped. These maps may be found in Figures 16

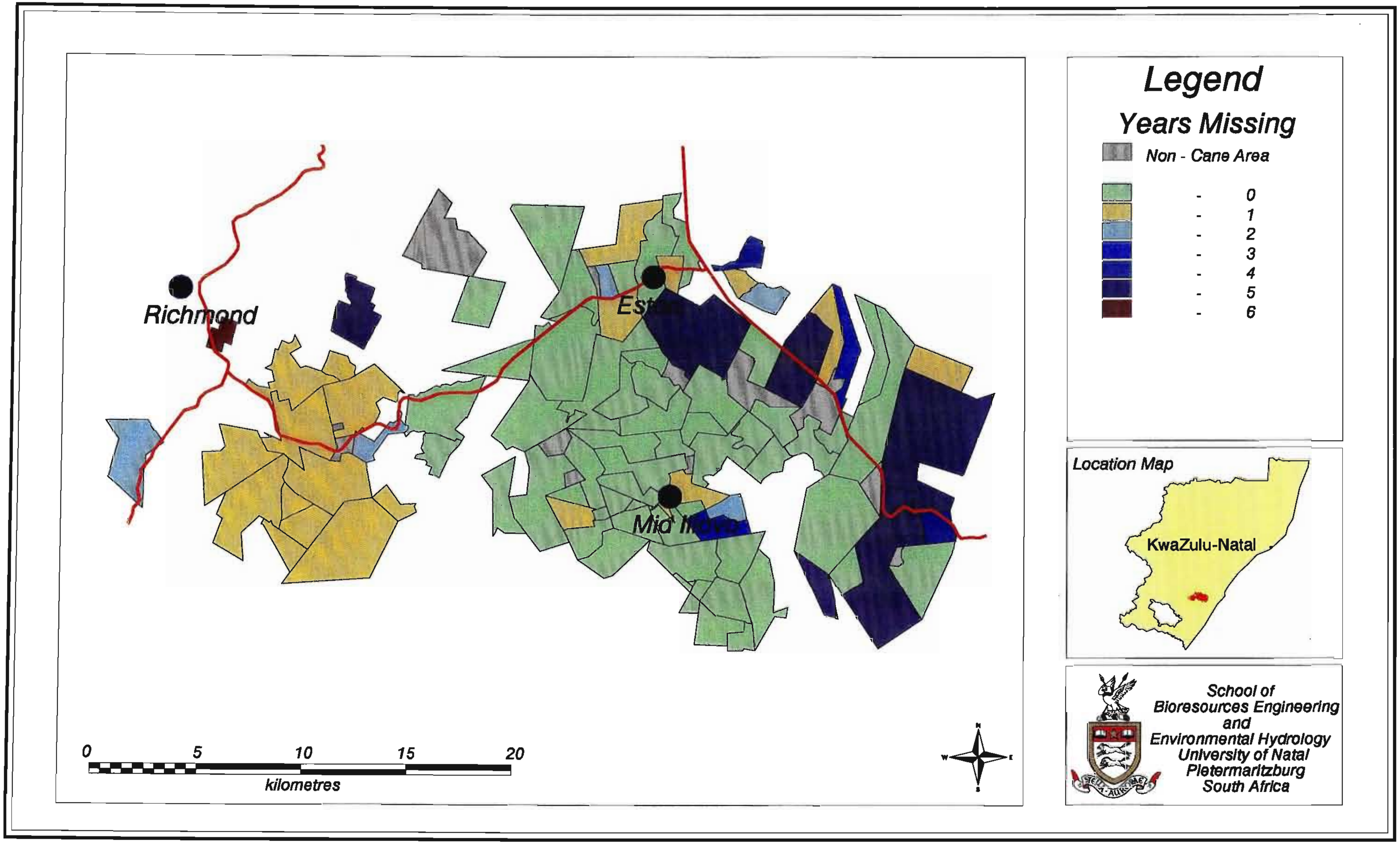


Figure 15 Number of years of missing observed farm yield data over the period 1988 to 1995

and 17 for the *ACRU*-Thompson and CANEGRO-DSSAT simulations respectively, where soils inputs have been derived from Land Type information. Farms on these maps having positive percentage differences indicate over-simulation, while negative values indicate under-simulation. Both maps have large areas shaded in green indicating that simulations are within 20% of the observed yield. The CANEGRO-DSSAT model tends to have more areas that are under-simulated by more than 20%, than does the *ACRU*-Thompson model. The *ACRU*-Thompson model tends to over-simulate yield, with the yields of 67% of farms being over-simulated, while 44% are over-simulated by CANEGRO-DSSAT.

An analysis of the variability of simulated yields in relation to the variability of observed yields was also conducted. Variation in yields was expressed through the coefficient of variation (CV). This statistic was calculated for yields simulated by the *ACRU*-Thompson and CANEGRO-DSSAT models over the period of simulation (soils inputs derived from Land Type information), as well as for the corresponding observed yields. For both sets of model simulations, the percentage difference between the CVs of simulated and observed yields was calculated on a farm-by-farm basis. These percentage differences were then mapped for the *ACRU*-Thompson (Figure 18) and CANEGRO-DSSAT (Figure 19) simulations. Positive percentage differences indicate that the simulated yields were too variable, while negative percentage differences indicate insufficient variation.

Reference to Figures 18 and 19 indicates that the *ACRU*-Thompson model tends to under-simulate variability in yields while the CANEGRO-DSSAT model over-simulates variability. The degree to which the former does not accurately represent yield variability is generally less than that of the latter, with most *ACRU*-Thompson simulated farm yields having a variability within 40% of the corresponding observed yields, in contrast to CANEGRO-DSSAT simulated farm yields where most simulations differ in variability by more than 40%. The variability in yields experienced in the MSA is thus better captured by the *ACRU*-Thompson model.

Aggregated MSA results

For yield forecasting, the average yield of the MSA is important as this affects activities such as mill operations planning. To verify the performance of the models at this scale, the simulated and observed farm yields were aggregated to obtain average MSA yields for each year of simulation.

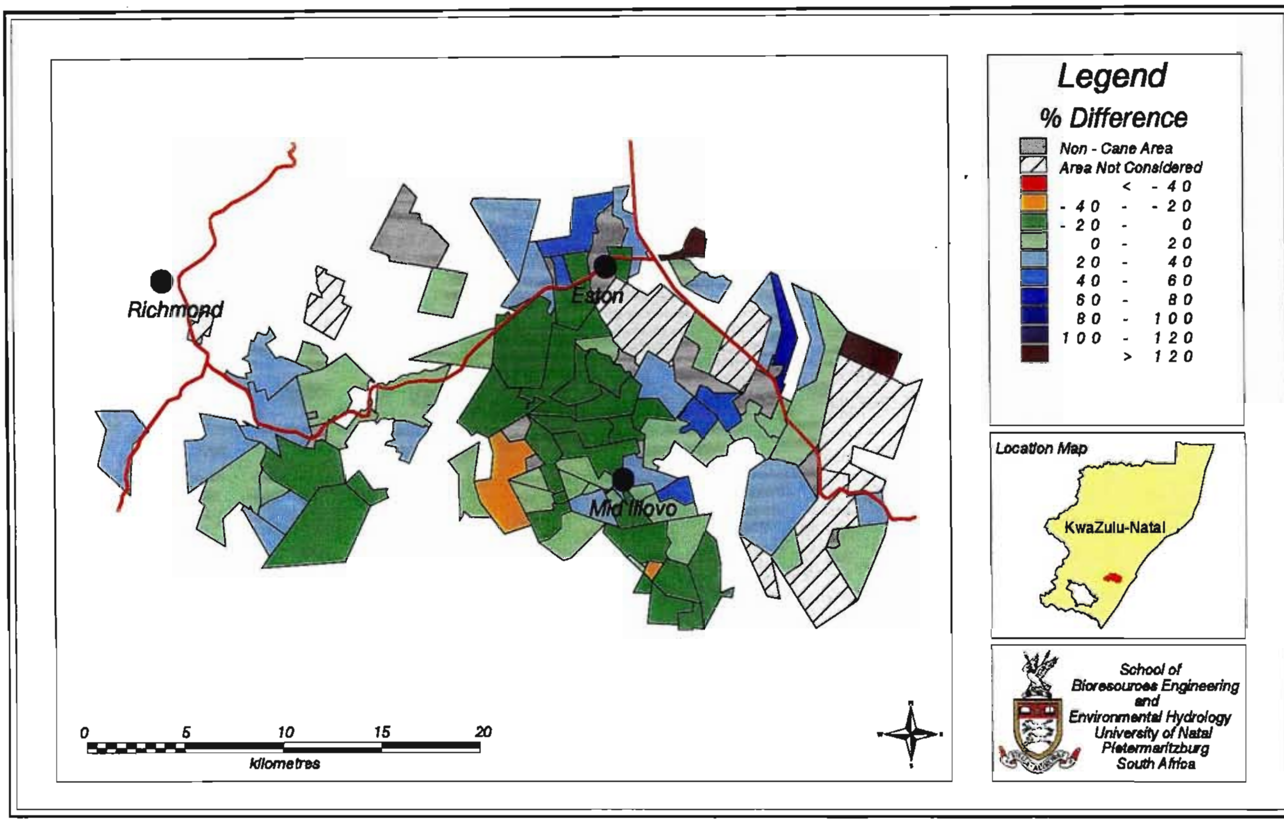


Figure 16 Percentage differences between means of simulated and observed yields: ACRU –Thompson model using Land Type derived soils

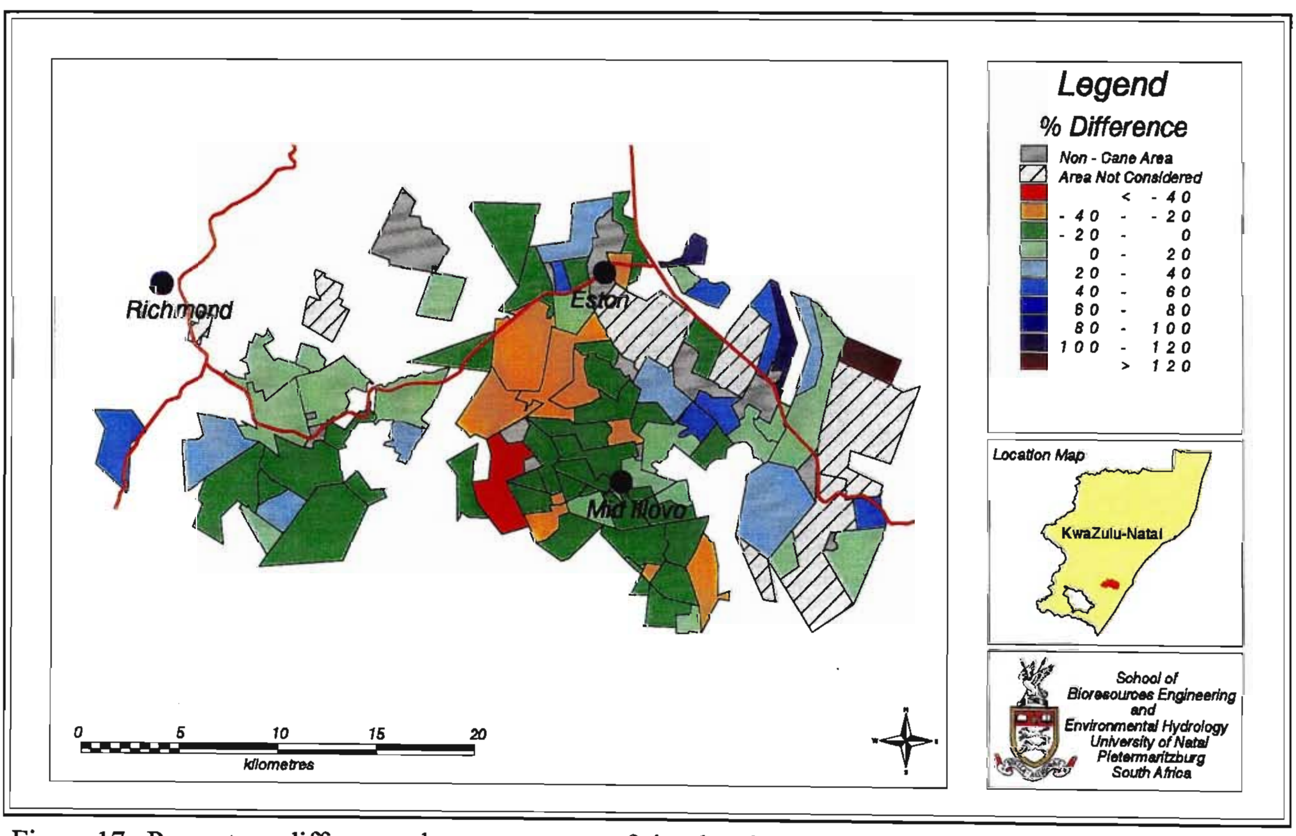


Figure 17 Percentage differences between means of simulated and observed yields: CANEGRO –DSSAT model using Land Type derived soils

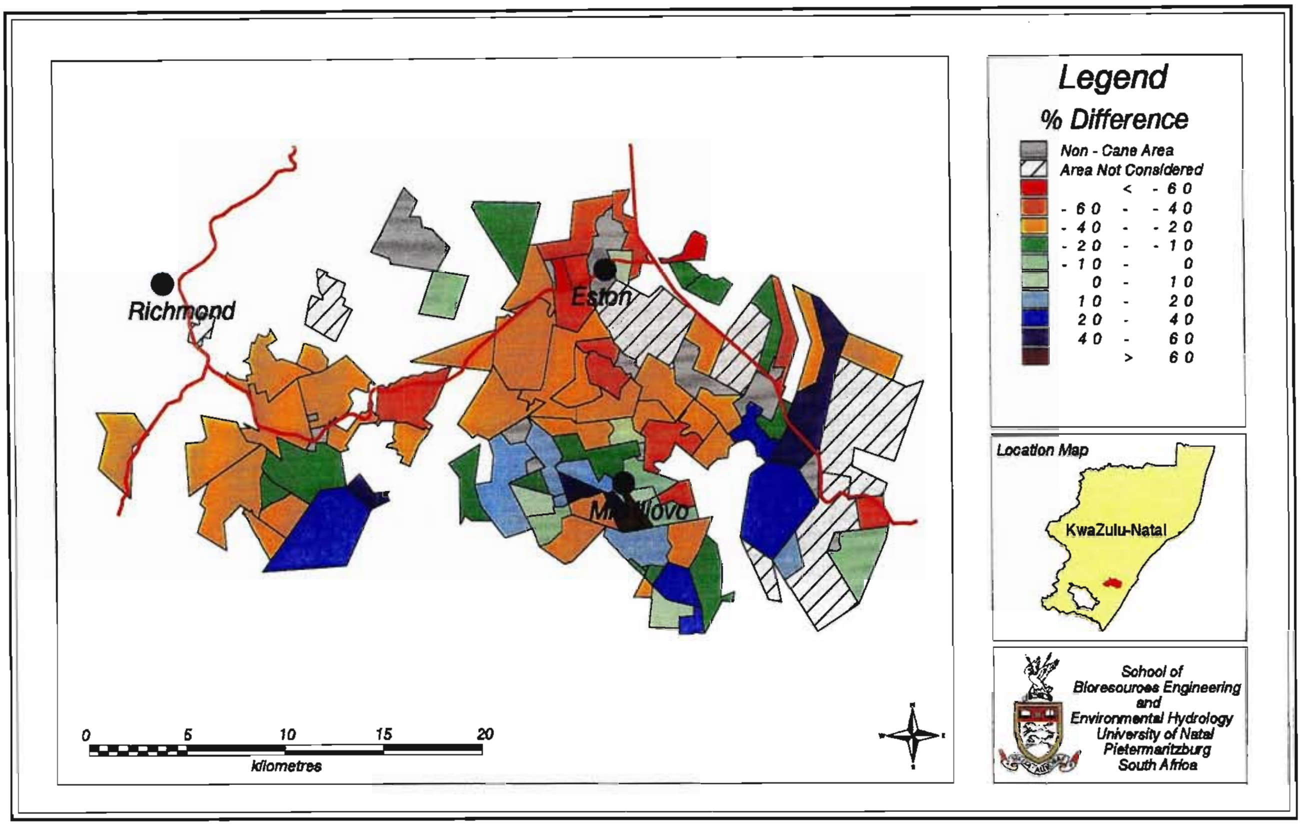


Figure 18 Percentage differences between the coefficients of variation of simulated and observed yields: ACRU –Thompson model using Land Type derived soils

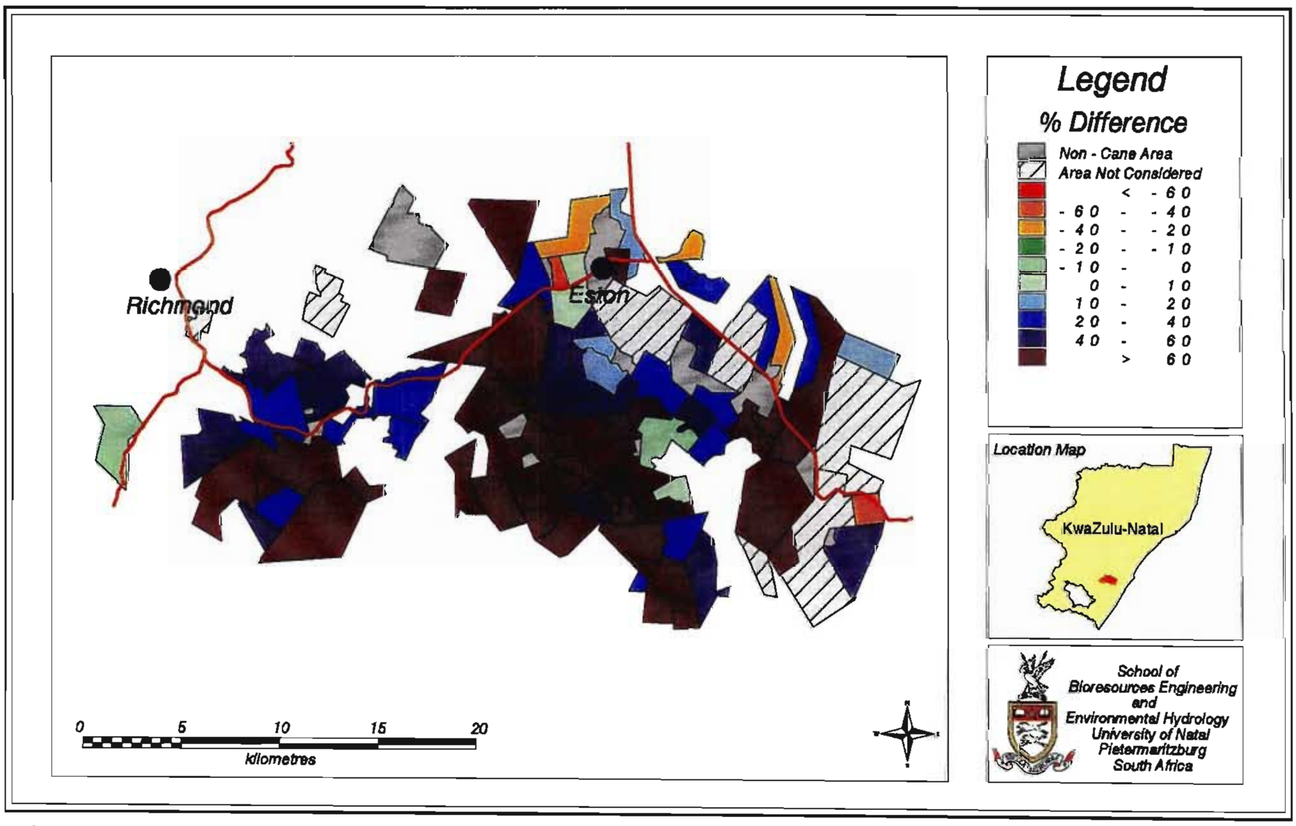


Figure 19 Percentage differences between the coefficients of variation of simulated and observed yields: CANEGRO –DSSAT model using Land Type derived soils

The average MSA yields for *ACRU*-Thompson, CANEGRO-DSSAT (both using LT derived soils) and the SRM are plotted against time in Figure 20, along with the observed yields. These plots verify that the *ACRU*-Thompson model has generally over-simulated yields, but has captured the trend in the year to year variation of yield. The CANEGRO-DSSAT model has over- and under-simulated yields and has not captured the trend in year to year yield variation as well. The yield for the low-yielding year of 1993 was markedly under-simulated. The SRM simulated yields well, except in the years of 1988, 1989 and 1995. This model is calibrated against observed rainfall, and would thus be expected to give good simulations in those years where rainfall has exerted a strong influence on seasonal yields.

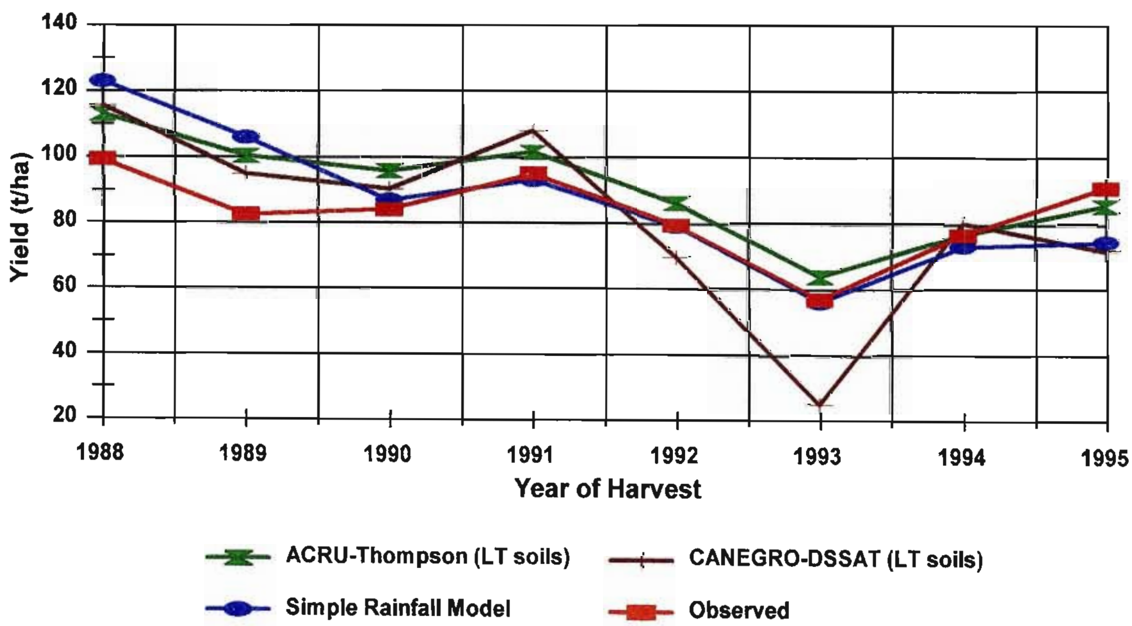


Figure 20 Average simulated (*ACRU*-Thompson, CANEGRO-DSSAT, Simple Rainfall Model) and observed mill supply area yields versus time

6.4.2 Model simulations based on parent material versus Land Type derived soils inputs

<u>Farm scale simulations</u>

In order to assess whether the two sources of soil information give rise to different yield simulations, the mean simulated yields using the two sets of soils inputs were calculated at farm scale for the *ACRU*-Thompson and CANEGRO-DSSAT models. For each model the percentage difference between the two sets of simulations was calculated and mapped, with positive percentage differences indicating that yields derived from PM soils are higher than those derived from LT soils (and vice versa). These maps are shown as Figures 21 and 22 for the *ACRU*-Thompson and CANEGRO-DSSAT models respectively. The indication from these maps is that there is little difference between the two sets of yield simulations for *ACRU*-Thompson (mostly within 10% of each other), while for CANEGRO-DSSAT there are notable differences between the two simulation sets (many areas have differences greater than 30%).

Aggregated MSA results

The farm scale simulations using different soils inputs were averaged to obtain aggregated MSA yields. The average MSA yields for the *ACRU*-Thompson model using both sets of soils inputs are plotted against time in Figure 23, while the corresponding plot for the CANEGRO-DSSAT model is given in Figure 24. The observed yields are also plotted in both figures. The findings at farm scale are again applicable at MSA scale, with the *ACRU*-Thompson model indicating almost no difference between yield simulations using different soils inputs, and the CANEGRO-DSSAT model indicating notable differences. In both plots there is no clear indication as to whether one set of soils inputs give rise to better yield simulations than the other.

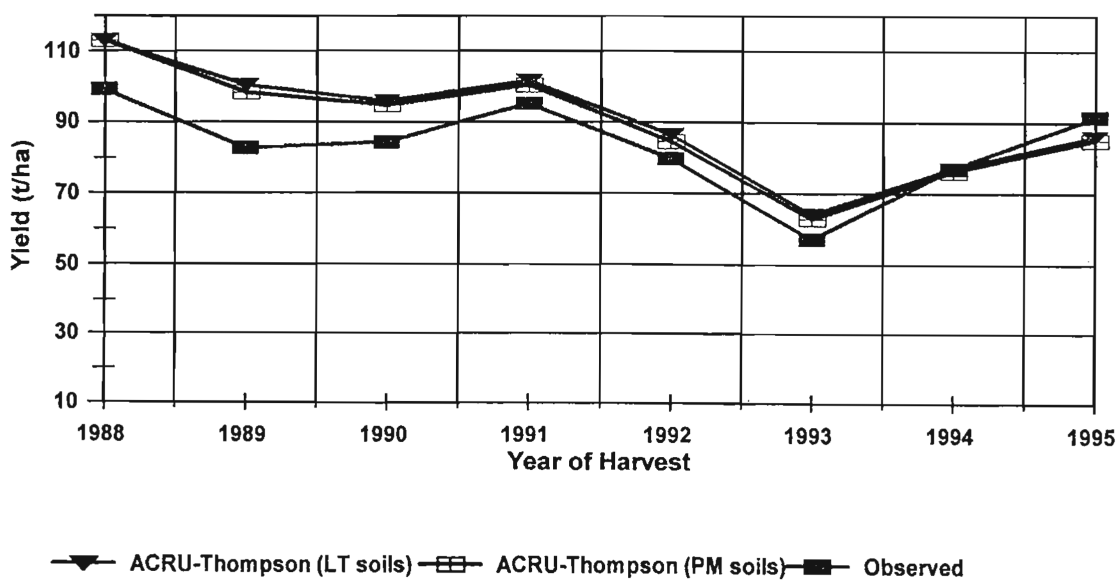


Figure 23 Average mill supply area yields simulated by the *ACRU*-Thompson model versus time, using Land Type and soil parent material derived soils inputs

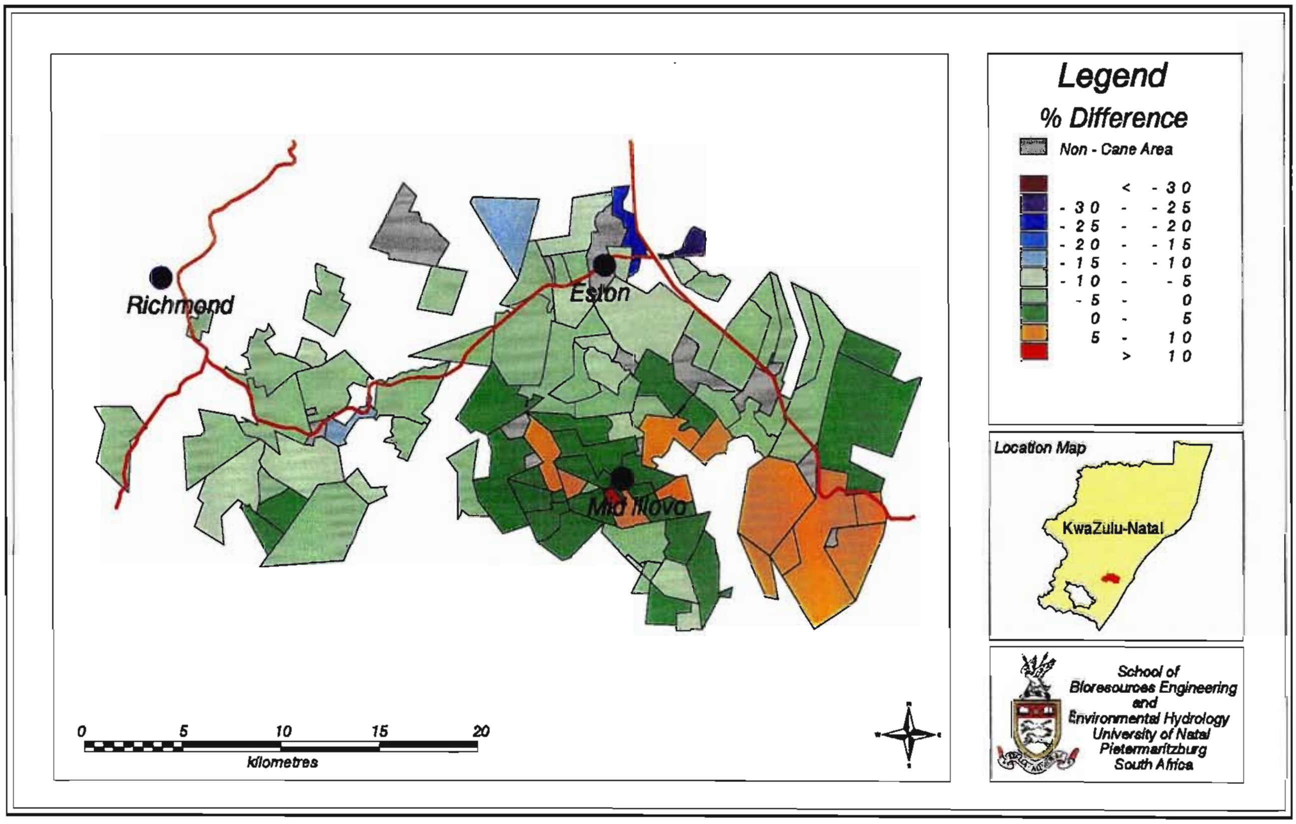


Figure 21 Percentage differences between the means of yields simulated using parent material and Land Type derived soils : ACRU–Thompson model

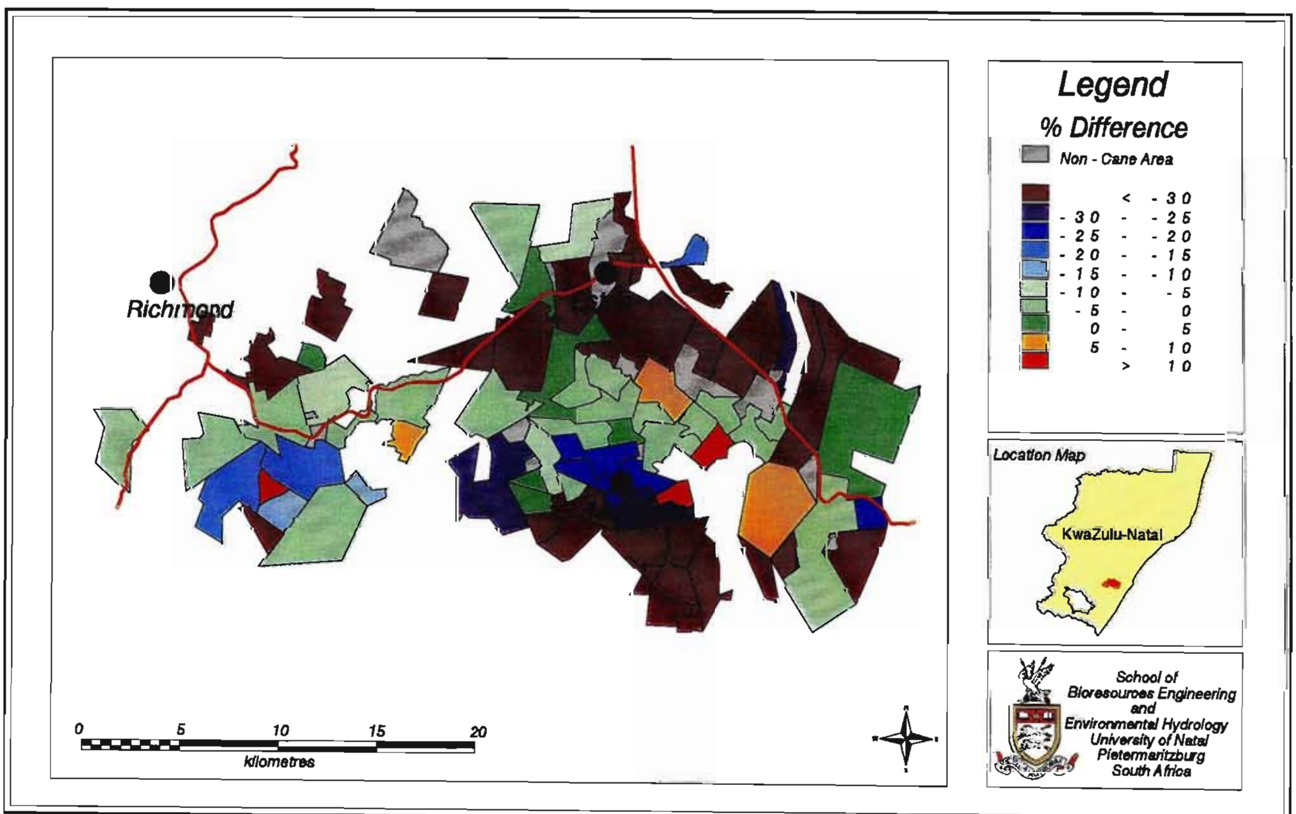


Figure 22 Percentage differences between the means of yields simulated using parent material and Land Type derived soils : CANEGRO–DSSAT model

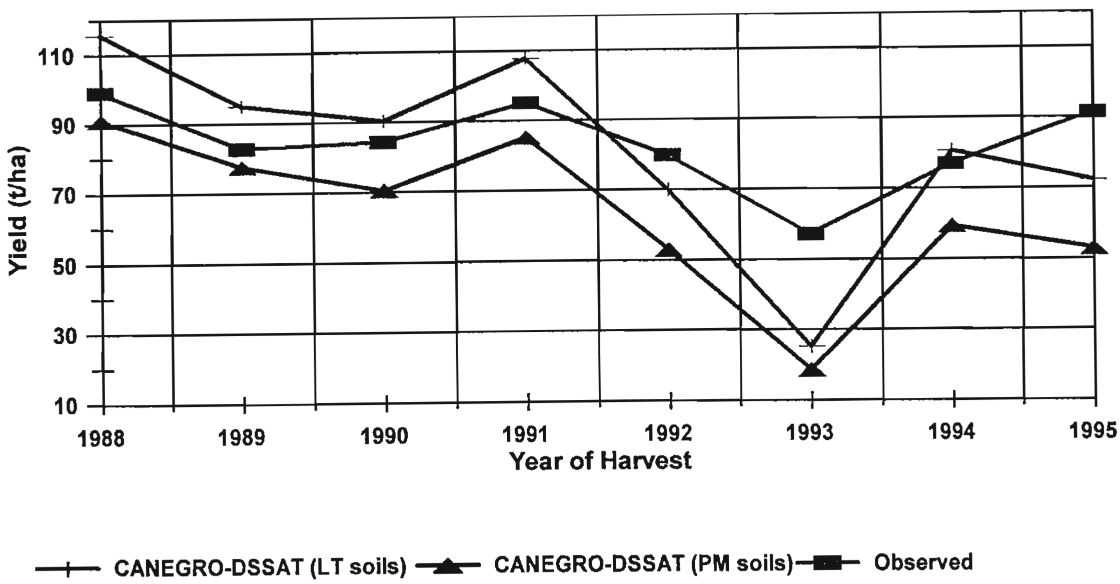


Figure 24 Average mill supply area yields simulated by the CANEGRO-DSSAT model versus time, using Land Type and soil parent material derived soils inputs

6.4.3 Model suitability for yield forecasting

The *ACRU*-Thompson and SRM models were selected for application in yield forecasting. From the results presented, it was noted that the *ACRU*-Thompson model over-simulated yield, but captured the trends in year-to-year variations of yield. The over-simulation of yields becomes less important if yields for the various years are expressed as fractions of the yield attained in the previous year. Expression of yields in this way allows for direct comparisons of yield sets having distinctly different means. This practice is common in industry when comparing simulated and observed yields, as yield models are often based on assumptions of perfect crop management, thus giving rise to the simulation of potential (non-management limited) yields. Such yields are seldom attained in practice, hence the need to convert yields to a common base. Figure 25 is a plot of *ACRU*-Thompson (LT soils), CANEGRO-DSSAT (LT soils), SRM and observed MSA yields versus time, where yields are expressed as yield ratios (ie as fractions of the previous year's yield). The 1994 harvest season yield ratio (value of 3.2) for the CANEGRO-DSSAT model is not shown in the plot, in order that the trends in yield ratios of the other models be maintained in perspective. The value for 1994 is high for the CANEGRO-DSSAT model because of the marked under-simulation of the 1993 harvest season yield. The removal of the trend in over-simulation is evident in this plot for the *ACRU*-Thompson model. The CANEGRO-DSSAT model (using LT soils),

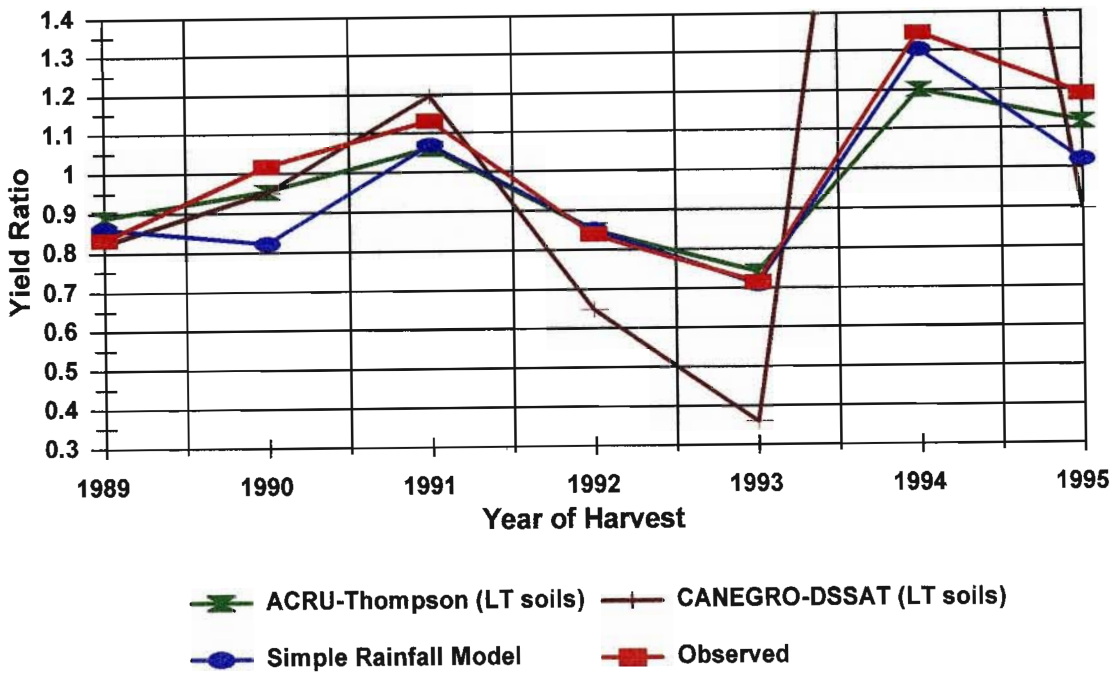


Figure 25 Simulated (*ACRU*-Thompson, CANEGRO-DSSAT, Simple Rainfall Model) and observed mill supply area yield ratios versus time

which over- and under-simulated observed yields, does not capture well the trend in year-to-year variation of yields.

It is evident from Figures 20, 24 and 25 that there is no additional accuracy in simulations derived from use of the CANEGRO-DSSAT model versus the *ACRU*-Thompson model. Given also the greater sensitivity of the CANEGRO-DSSAT model to soils inputs, and thus the need for a higher level of accuracy in these inputs, it was decided that the *ACRU*-Thompson model would be selected for yield forecasting. Yield forecasting requires a model producing robust simulations of yield at the required scale, especially under the conditions of greater uncertainty that prevail in the soil-plant-atmosphere environment.

As yield simulations using the *ACRU*-Thompson model indicated no significant benefit in using the Land Type versus soil parent material derived soil inputs, it was decided that the LT derived inputs would be used. These inputs were simpler to prepare as the LT information is a standard source of soils information, and does not require the soil types to be derived. Hence LT information would also be used in preference if similar work were carried out elsewhere. The SRM was applied in yield forecasting for reasons discussed in the introduction of this dissertation.

6.4.4 Representation of growth cycles in simulations

An evaluation was made between using all 11 growth cycles versus using only the 4 most dominant cycles, to assess what accuracy was lost in simulations when using fewer cycles in the *ACRU*-Thompson model. The use of fewer cycles would considerably reduce the number of simulations required during yield forecasting. Figure 26 compares simulated MSA yield ratios using 4 versus 11 growth cycles. A 1:1 line is drawn to represent the line along which the ratios would fall, were there a perfect correlation between the two sets of simulations. The ratios fall close to this line, indicating that the simulations using four growth cycles have adequately represented yields. The use of only four growth cycles was thus considered adequate for yield forecasting.

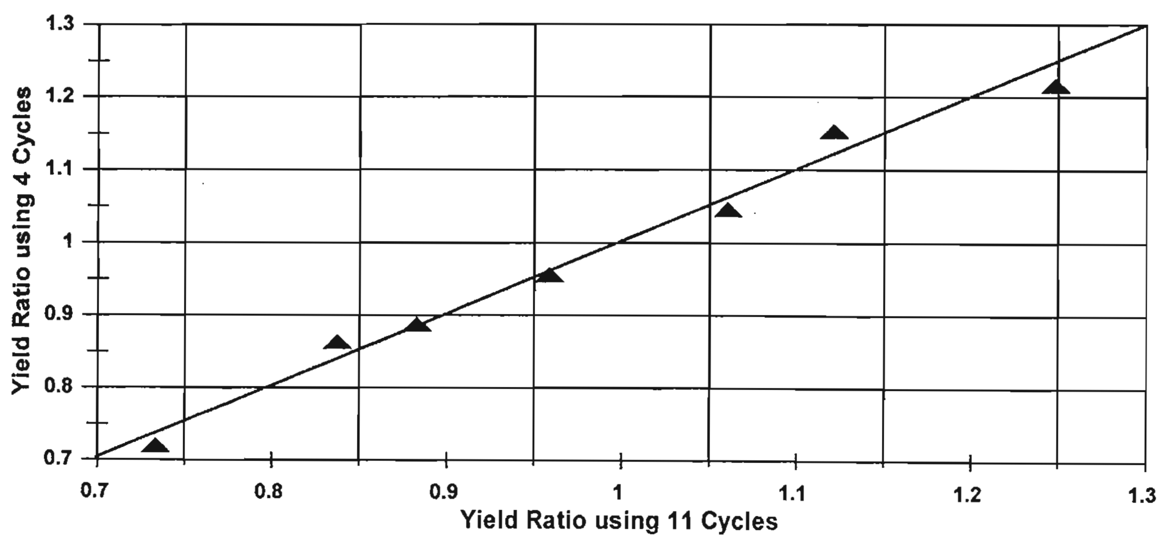


Figure 26 Mill supply area yield ratios simulated by the *ACRU*-Thompson model, using 4 versus 11 growth cycles

In this chapter it has been verified that the *ACRU*-Thompson and SRM models are able to accurately predict historical cane yields given an observed climate record. The preparation of inputs to these models has been described. Decisions regarding the choice of soils inputs and the number of cycles to be considered in yield forecasting, have also been made. The next chapter of this dissertation relates to the application of the models in yield forecasting.

7 CROP FORECASTING USING SUGARCANE YIELD MODELS

In this chapter, seasonal rainfall forecasts are applied together with the selected yield simulation model (*ACRU*-Thompson), to produce forecasts of sugarcane yield. These yield forecasts are then compared against observed yields and forecasts produced using traditional industry methods. The chapter begins with an evaluation of the seasonal rainfall forecasts applied in the model. This evaluation is then followed by consideration of the modelling strategy employed for yield forecasting. The yield forecasts are then presented and evaluated at the end of the chapter.

7.1 Evaluation of Seasonal Rainfall Forecasts

7.1.1 Description of rainfall forecasts

Categorical seasonal rainfall forecasts were obtained from SAWB (Landman and Bartman, 1998) for 11 locations in and around the Eston MSA. These rainfall forecasts were derived from statistical models employing canonical correlation analysis, where rainfall over South Africa is related to sea surface temperatures of various oceans around the globe. The relationships upon which the models are based are linked to the ENSO phenomenon, which has been shown to influence rainfall over southern Africa, even though the origins of the phenomenon are distant from the region (Lindesay, Harrison and Haffner, 1986; Van Heerden, Terblanche and Schulze, 1988).

The SAWB rainfall forecasts were considered to be as appropriate for application as any other seasonal rainfall forecasts available for South Africa. They are categorical in nature with rainfall being forecast as either above normal, near normal or below normal. On each date of forecast, rainfall is predicted for two successive three-month periods, allowing for a total lead time of six months. Forecasts are updated at monthly intervals. The rainfall forecasts are usually issued by SAWB for broad areas within the provinces of South Africa, as this scale gives the greatest skill in forecasting. However, for this study, forecasts were obtained for specific locations (rainfall stations) in and around the Eston MSA, in order that the forecasts be consistent with the proposed scale of modelling. Forecasts were obtained for the period 1988 to 1998. Categories of rainfall derived retrospectively from the observed rainfall records of those seasons were also obtained.

These categories of rainfall correspond to a "perfect" forecast while those generated by the SAWB statistical rainfall model are the "actual" rainfall forecast.

7.1.2 Selection of rainfall forecast locations for skill assessment

The 11 locations (rainfall stations) for which rainfall forecasts were obtained are shown in Figure 27. Given that rainfall forecasts give rise to greatest forecast skill when applicable over broad areas, an investigation of various groupings of locations was carried out, in order to assess whether the collective forecast from a group of locations would yield greater accuracy in the rainfall forecasts than a single location would. Before assessing the accuracy of a collective forecast, however, it was necessary to ensure that groupings of forecasts were relatively consistent in predictions for a particular forecast period. All 11 locations combined were considered as a grouping, as were various sub-groupings thereof (minimum of five station locations). Analysis of the forecasts from these groupings, indicated that the cluster of five locations falling within and immediately around the Eston MSA were most consistent in their forecasts. This cluster of locations was identified for assessment of accuracy of their collective rainfall forecasts.

7.1.3 Assessment of rainfall forecast skills

The collective forecasts from the cluster of five locations identified above were evaluated in terms of their skill. The forecasts from the single location falling within the MSA (Eston Station) were also evaluated. Above normal rainfall forecasts were assigned a categorical value of 3, while near normal and below normal forecasts were assigned values of 2 and 1 respectively. Forecast skill was assessed by subtracting the values of the forecasts had they been correct (i.e. perfect) from those of the actual forecasts. This assessment is presented in Figure 28 for the Eston Station location as well as for the modal (collective) forecast of the cluster of five locations.

In both cases the differences between actual and perfect forecasts are presented in the form of a colour-coded table, where rows relate to the month of forecast generation, and columns to the year of the forecast period. Each block in the table corresponds to a 3 month forecast period (horizontal direction). If actual and perfect forecasts are identical, then the relevant block in the table is coded green, to indicate an identical forecast. Forecasts not identical are coded yellow for a difference of 1 category, and red for a difference of 2 categories. The percentage occurrences

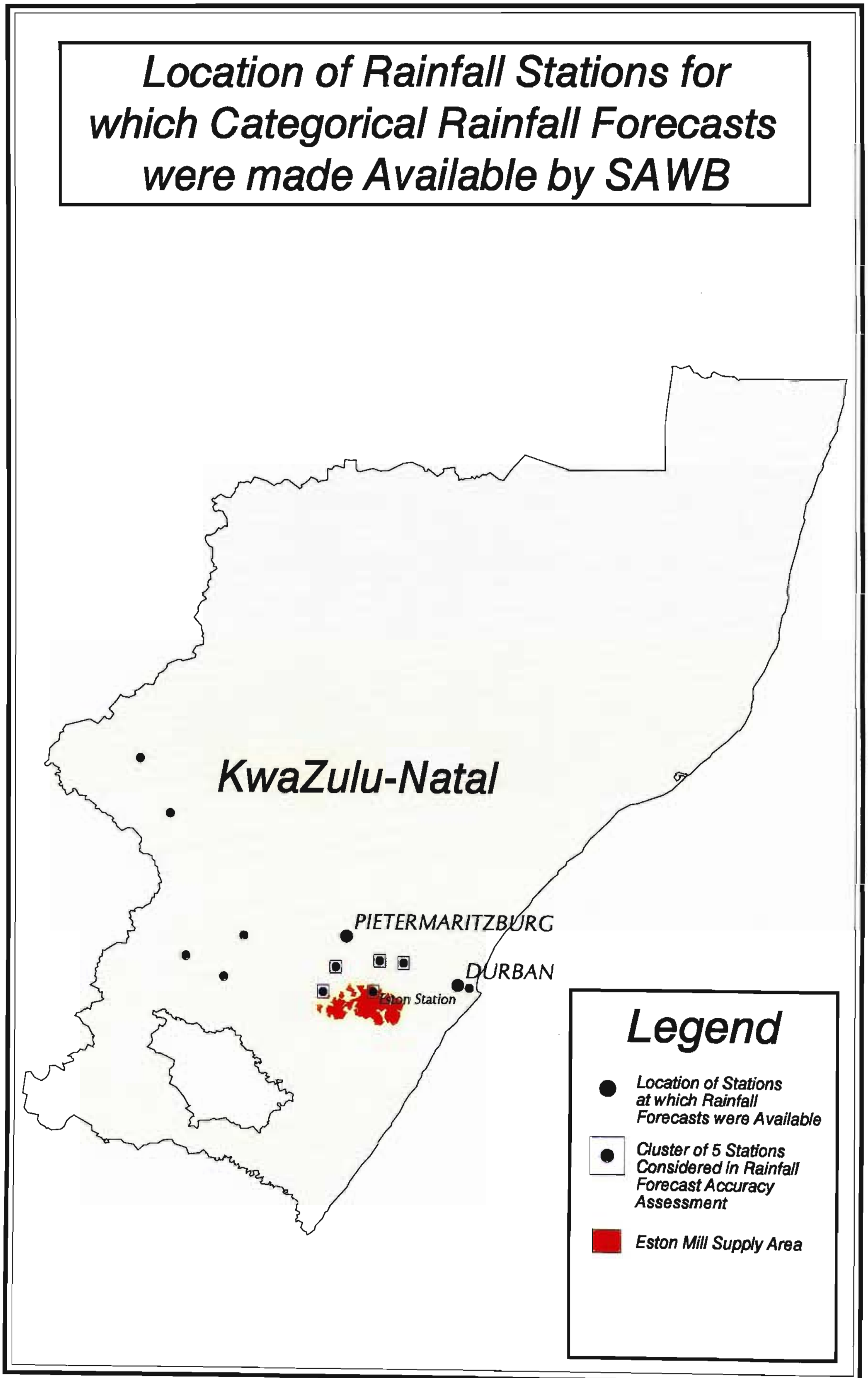


Figure 27 Location of rainfall stations for which categorical rainfall forecasts were made available by SAWB

Evaluation of Rainfall Forecast Skills

Eston Station Forecasts:

Forecast Difference = Actual Forecast - Perfect Forecast

Month of Forecast	Lead Time (months)	87	88	88	89	89	90	90	91	91	92	92	93	93	94	94	95	95	96	96	97	97	98	98
SEP	12.5 / 0.5	1	-2	0	0	-1	0	1	-1	2	0	2	1	-1	0	0	0	0	-2	0	0	0	-1	
DEC	9.5		-2	-2	0	1	0	-1	-1	1	1	1	1	2	1	2	0	-1	-2	2	0	-2		
JAN	8.5		-2	0	0	0	-2	0	1	0	1	0	1	0	1	0	0	0	-1	0	1	0	0	0
MAR	6.5		-2	2	-1	2	-1	1	1	1	0	2	2	1	1	0	-2	2	2	-2	-2	2	0	1
MAY	4.5		1	2	1	0	0	1	2	2	2	2	1	2	-1	1	0	1	0	1	0	0	1	1
JUN	3.5		2	-1	2	-2	1	0	1	2	2	2	1	1	0	0	2	0	-2	0	2	0		

KEY		% Occurence
Green	Forecasts Identical	34
Yellow	Forecasts Differ by 1 Category	35
Red	Forecasts Differ by 2 Categories	31

Modal Forecasts of 5 Locations:

Forecast Difference = Actual Forecast - Perfect Forecast

Month of Forecast	Lead Time (months)	87	88	88	89	89	90	90	91	91	92	92	93	93	94	94	95	95	96	96	97	97	98	98
SEP	12.5 / 0.5	0	-2	0	-1	-1	-1	0	-1	1	0	2	1	-1	0	2	0	0	-2	0	0	0	-2	
DEC	9.5		-2	-2	-1	0	-1	-1	-1	1	1	1	1	2	1	1	0	-1	-2	1	0	-2		
JAN	8.5		-2	0	-2	0	-2	0	1	0	1	0	1	0	1	0	0	0	-1	0	1	0	-1	0
MAR	6.5		-2	2	0	2	-1	1	1	1	0	2	2	1	1	1	-2	2	1	-1	-2	0	0	1
MAY	4.5		0	1	1	0	0	1	2	0	2	2	2	1	-1	2	-1	0	0	1	0	0	1	1
JUN	3.5		2	-1	2	-2	1	-1	1	1	2	2	1	1	1	2	2	0	-2	0	0	0		

KEY		% Occurence
Green	Forecasts Identical	32
Yellow	Forecasts Differ by 1 Category	41
Red	Forecasts Differ by 2 Categories	27

Figure 28 Evaluation of rainfall forecast skills for Eston Station forecasts and the modal forecasts of a cluster of 5 locations

of the various block colours are indicated for both tables. These percentages indicate that the forecasts exhibit poor skill for both Eston Station and the group of 5 locations, and that there is no perceivable benefit in the grouping of stations. On this basis it was decided that the forecasts of Eston Station would be used in yield forecasting. The poor skill of the forecasts indicate that they are not, as yet, ideally suited to a scale of application the size of a MSA. The location of the MSA in KwaZulu-Natal may also be a contributing factor, as this region is known to have relatively poor forecast skills when compared with other regions of the country (Schulze, Hallowes, Lynch, Perks and Horan, 1998).

7.2 Modelling Strategy for Yield Forecasting

In developing a modelling strategy for yield forecasting, two issues were considered important, namely, the scale at which yield forecasting should be carried out, and a methodology for translating the categorical seasonal rainfall forecasts into a form suitable for application in the *ACRU*-Thompson and SRM models. These two issues are addressed in this section.

7.2.1 Scale of modelling for the *ACRU*-Thompson model

For the generation of yield forecasts using the *ACRU*-Thompson model, it was decided that a coarser scale of modelling than that used in model verification (Chapter 6) would be adopted, given the large number of simulations anticipated in this phase of the research. The MSA was divided into three sub-areas for yield forecasting, these areas being delineated according to mean annual precipitation. The Land Type derived soils inputs developed at farm scale (Chapter 6.1.2), were then averaged for each of the sub-areas in order to form the inputs required at the new scale of modelling. The three sub-areas delineated and the soil TAM derived from the averaged soils inputs, are indicated in Figure 29. The rainfall driver stations selected for the three sub-areas are Stations 4, 8 and 1 in Figure 6, for Sub-areas 1, 2 and 3 respectively. Apart from being representative of the sub-areas, the stations were also selected based on their long length of observed rainfall record, an important characteristic when considering the translation of rainfall forecasts into daily rainfall data sets (Chapter 7.2.2 below). For yield forecasting, the rainfall data from the driver stations were adjusted to better represent the rainfall in the sub-areas. These adjustments were made on a monthly basis in a manner similar to that outlined in Chapter 6.1.1.1 (relative gridded median monthly rainfalls were considered as opposed to relative gridded MAP).

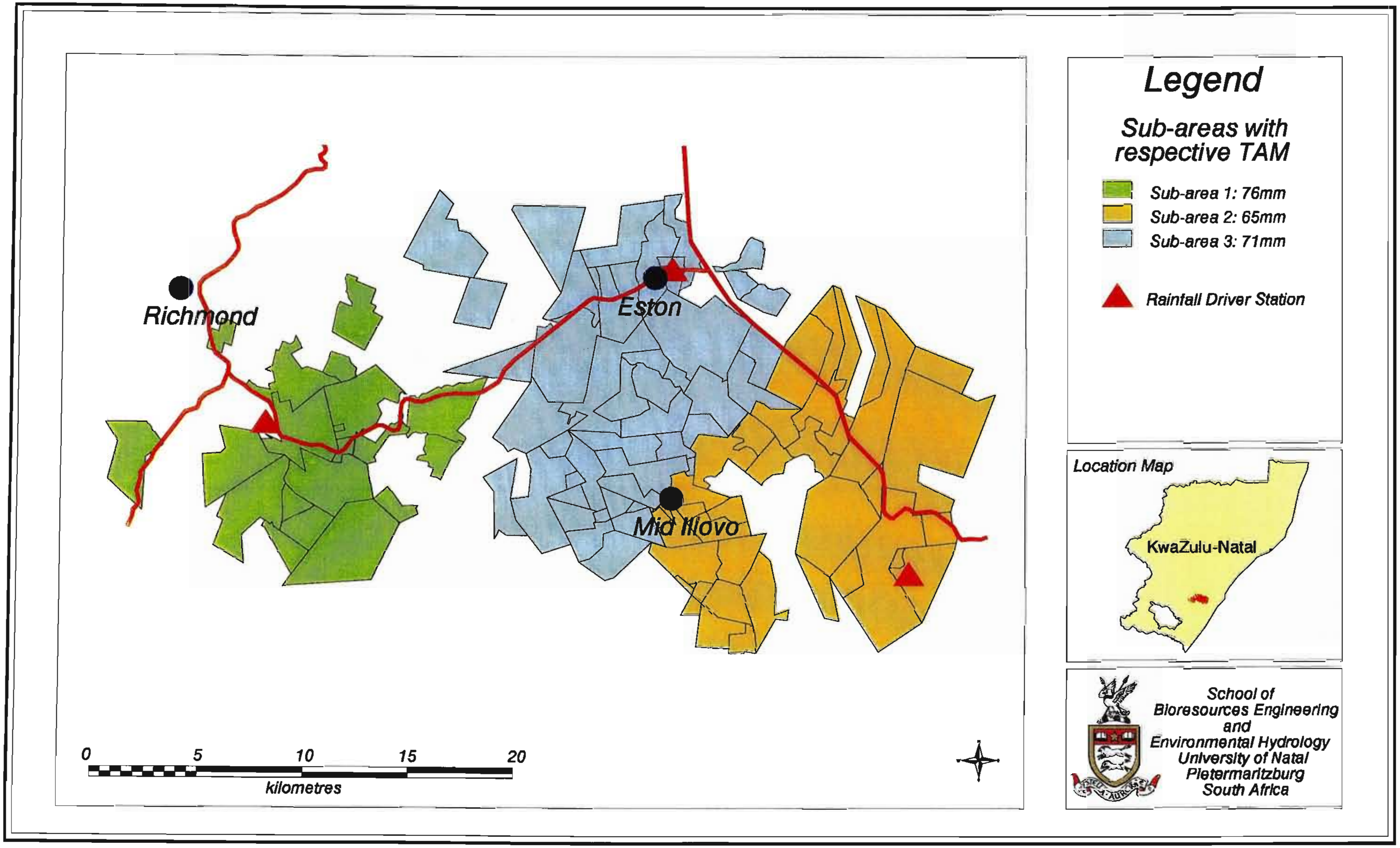


Figure 29 Sub-areas of the Eston Mill Supply Area used in yield forecasting: Average total available soil moisture (TAM)

Further changes in model inputs were introduced to simplify the yield forecasting simulations. These included the simulation of only 4 of the 11 growth cycles, as was recommended in Chapter 6.4.4. The option in the *ACRU*-Thompson model to estimate reference potential evaporation through a daily temperature-based equation (Linacre ,1991) was no longer used, with the option allowing for the use of monthly averages of A-pan equivalent evaporation being invoked in its place. The method of estimating reference potential evaporation was changed, as climate forecasts were only obtained for rainfall and not for temperature, resulting in simulation options requiring daily temperatures being no longer feasible. As the *ACRU*-Thompson model does not offer the facility to change the method of estimating reference potential evaporation during the course of a simulation, the option of using monthly averages of A-pan equivalent evaporation was used throughout the simulations. In the absence of daily temperatures, the *ACRU*-Thompson model requires monthly means of daily temperatures to be input into the model. The required monthly averages of potential evaporation and temperature were obtained for the three sub-areas from the national climate database described in Schulze (1997).

In order to assess whether the above changes would affect the accuracy of the *ACRU*-Thompson model, yields were simulated for the same period as in Chapter 6, allowing for the two sets of simulations to be compared. When determining the average MSA yield based on the coarser scale of modelling, the simulations of the three sub-areas were weighted according to the proportion of the MSA that each sub-area constituted. The average MSA yield ratios (ratio of yield to previous year's yield) for the simulations based on farm and sub-area scales of modelling, are plotted against time in Figure 30, along with the corresponding observed yield ratios. The plots indicate that the sub-area scale of modelling still gives rise to yield simulations that closely mimic the observed yields. Therefore this scale of modelling was considered suitable for yield forecasting.

7.2.2 Translation of categorical seasonal rainfall forecasts for application in yield models

The seasonal rainfall forecasts obtained were categorical in nature, and required translation into daily and monthly rainfall inputs suitable for application in the *ACRU*-Thompson and SRM models. The translation of the forecasts was achieved through the use of an analogue year concept developed by Lecler (Lumsden *et al.*, 1999), which is described in the following paragraphs.

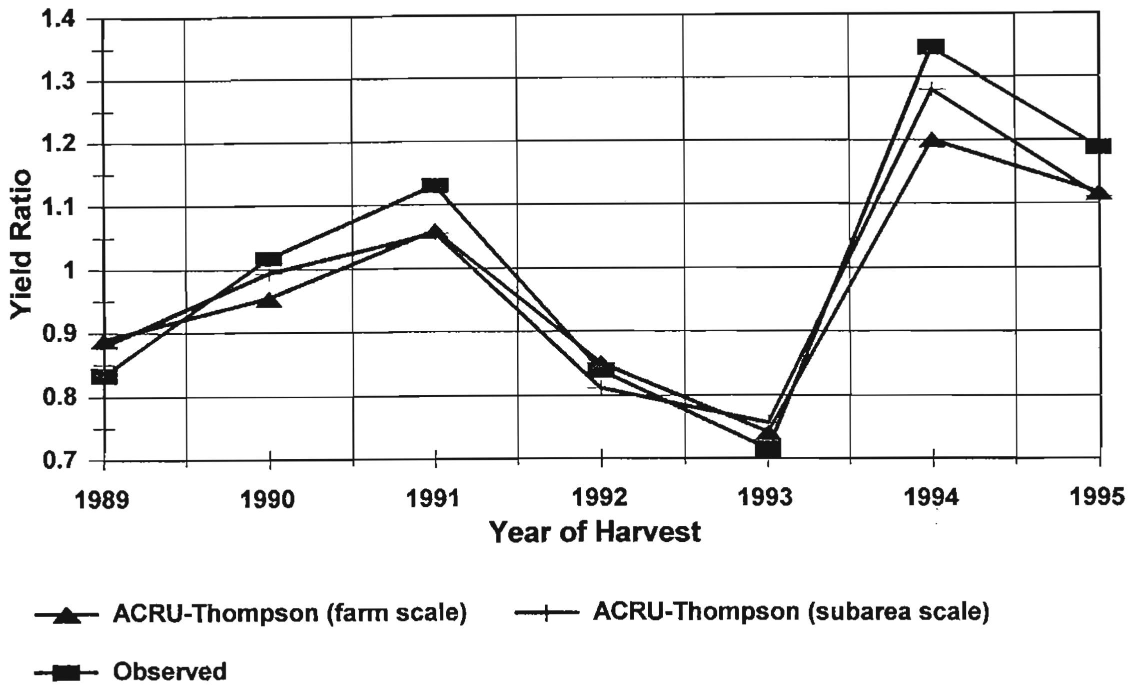


Figure 30 Average simulated mill supply area yield ratios for simulations based on farm and sub-area scales of modelling

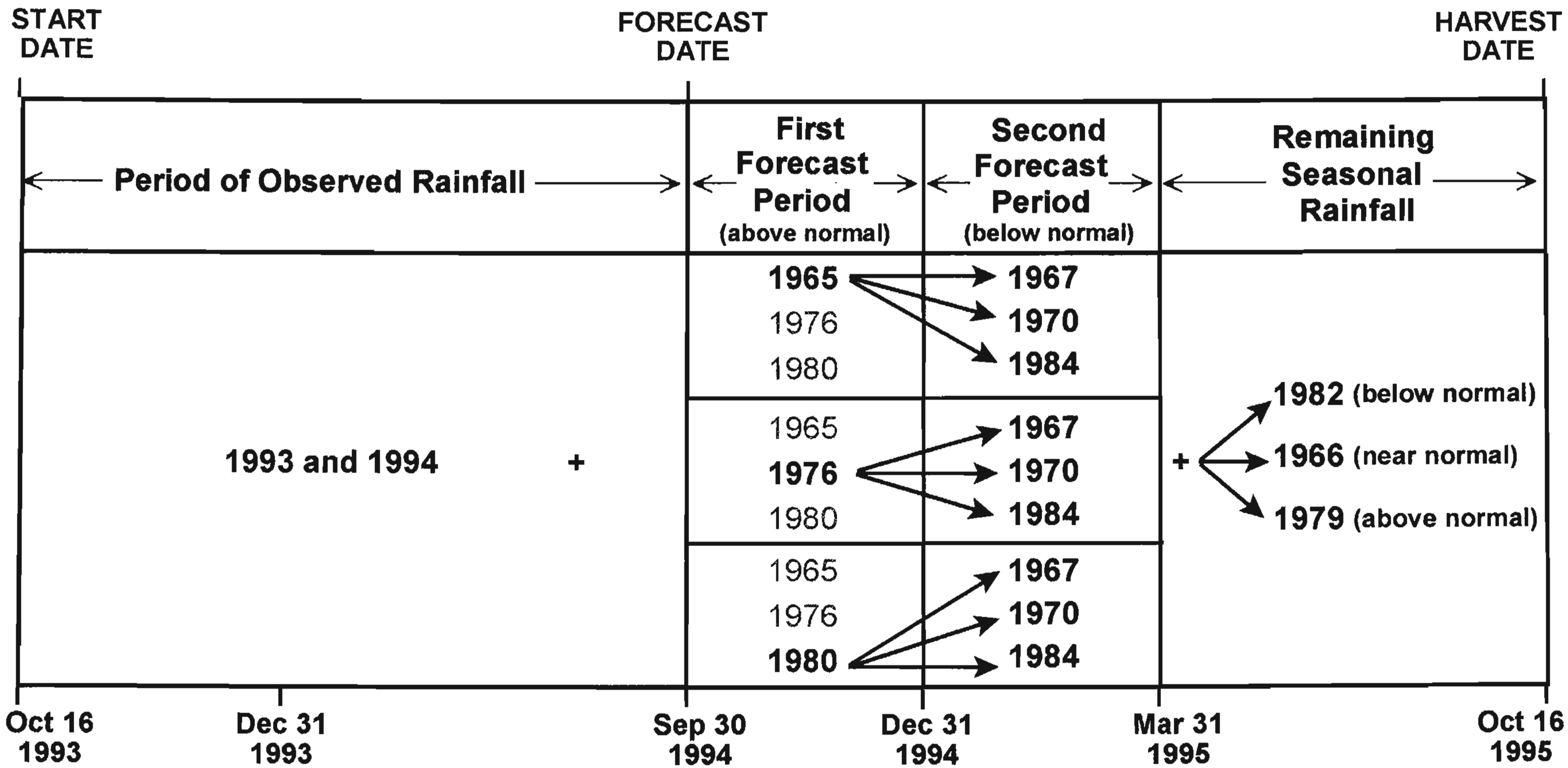
For decision-making purposes, certain dates within a season were recommended by industry representatives as being appropriate for the generation of yield forecasts. These dates of forecast and the corresponding lead times are shown in Table 11.

Table 11 Recommended dates of forecast and corresponding lead times for the generation of yield forecasts

Forecast	Recommended Date of Forecast	Lead Time (months)
1	End of September	12.5
2	End of January	8.5
3	End of March	6.5
4	End of May	4.5
5	End of September	0.5

The rainfall forecasts obtained were used to infer rainfall for the months following the various forecast dates. The analogue year concept used in the translation of rainfall forecasts involved the identification of years in the historical rainfall record resembling a particular forecast. The data from those years were extracted and used to create forecast rainfall files. In the case of the *ACRU*-Thompson model, all years resembling a particular forecast were used successively to create a number of daily rainfall files, each of which was then used, in turn, to simulate a yield for the season. This resulted in a number of yield outcomes being generated for a particular season. A forecasting methodology giving rise to a range in possible yields for a season is desirable, as it allows for an appreciation of the risk associated with any decisions based on the yield forecasts.

Figure 31 illustrates the development of forecast rainfall files for the *ACRU*-Thompson model for a hypothetical crop harvested in October 1995. The forecast represented in Figure 31 is the first for the 1995 season (as at 30 September 1994) for a crop starting on 16 October 1993 and harvested on 16 October 1995. The rainfall file is filled with observed daily rainfall for the period leading up to the forecast date. During the first and second three-month rainfall forecast periods, all combinations of years (indicated by use of arrows) corresponding to the given categorical forecast are identified, and their daily data used to fill the rainfall file for these periods. Figure 31 indicates that the categorical forecasts associated with the first and second three-month forecast periods, were above normal and below normal respectively, and that three associated years were identified in each case. In practice there were generally many more (in the order of 10 to 15) years in the 40 year historical record that were identified as having rainfall resembling that of a particular categorical forecast. For the months following the second forecast period, there is an equal chance of above normal, near normal or below normal rainfall occurring. In order to represent this equal probability all simulations were performed in triplicate, with the period of remaining seasonal rainfall being filled with above, near and below normal rainfall for that period. This increased the number of yield outcomes for a season, and reflected a wider range in yields possible. As a season progresses and the forecast date becomes later, the period of observed rainfall increases, while the period of remaining seasonal rainfall decreases, thus resulting in greater certainty in the representation of rainfall for that season. The categorical rainfall forecasts used in yield forecasting were those for Eston Station, which was also the rainfall driver station for Sub-area 3. These forecasts were assumed to apply to the other two driver stations. The historical records of the other two stations were used in the development of forecast rainfall files for their respective sub-areas. Both "actual" and "perfect" seasonal rainfall forecasts were used to create two distinct



Previous "above normal" years (Sep 30 - Dec 31): 1965, 1976, 1980

Previous "below normal" years (Dec 31 - Mar 31): 1967, 1970, 1984

Figure 31 Use of combinations of years of rainfall data in the development of a forecast rainfall file for the *ACRU*-Thompson model for a hypothetical crop harvested in October 1995 sets of rainfall files, both of which were used to forecast yields. These two sets of yield forecasts were then later compared during the analysis of results.

The forecasting approach adopted for the *ACRU*-Thompson model was not appropriate for the SRM, as this model is spreadsheet-based and was not suited to the complex representation of multiple rainfall scenarios. In addition, such an approach would not have followed the philosophy which was adopted in regard to the application of the SRM in this research. The approach adopted was rather to identify the year having a rainfall total for the first or second forecast period that was closest to the 25th, 50th or 75th percentile value, depending on the nature of the categorical forecast. The data from that year were then input into the model. For the period of remaining seasonal rainfall, the year having rainfall closest to the median (50th percentile) rainfall for that period was identified, and the data for that period input into the model. This resulted in only one yield outcome being predicted by the SRM for a particular forecast date and season. The rainfall driver station used in Chapter 6 (driver station 5, Figure 6) was also used for yield forecasting. The values of average growth cycle length and mean rate of yield accumulation were maintained at the values used in Chapter 6 (22 months and 5.70 t/ha/100mm respectively). Yield forecasts were generated using both "actual" and "perfect" rainfall forecasts.

7.3 Results

Yield forecasts were generated using the *ACRU*-Thompson and SRM models for the 1988 to 1998 harvest seasons. Forecasts were generated at the five recommended forecast dates within each season. In the case of the *ACRU*-Thompson model, results were averaged for the three sub-areas to give aggregated results for the MSA. These results were then compared to the observed yields for those seasons, and the yields estimated by the SRM and Mill Group Board. For reasons discussed in Chapter 6, all yield results were presented in the form of yield ratios (ratio of current year's yield to previous year's yield). The median (50th percentile) *ACRU*-Thompson yields were selected for comparative analyses from the range of yields forecasted for each season.

The results of yield forecasting are presented in sections relating to accuracy of forecasting, the application of actual versus perfect rainfall forecasts in yield forecasting and ranges in yields forecasted by the *ACRU*-Thompson model. Benefit analyses of the application of the different forecasting methods are also presented.

7.3.1 Accuracy of yield forecasts

The *ACRU*-Thompson, Simple Rainfall Model and Mill Group Board forecasts were plotted against time in Figure 32 for the various lead times. The observed yields for those seasons where data were available, were also plotted. The *ACRU*-Thompson and SRM forecasts were those derived from actual rainfall forecasts. The MGB forecasts only became available at a 6.5 month lead time, and are thus not indicated in all of the plots. MGB forecasts from the Illovo Mill were used in analyses for the period prior to the opening of the Eston Mill. The forecasts for the 0.5 month lead time are not shown for any of the forecasting methods, as there were negligible differences between the 0.5 month lead time forecasts and those of the preceding 4.5 month lead time. The graphs indicate that forecasts become more accurate as the lead time shortens. The *ACRU*-Thompson and SRM forecasts were in many cases very similar, except for some seasons (eg. 1990, 1995) where the SRM undersimulated yield ratios. The MGB forecasts, when they became available, gave a good representation of observed yield ratios. To give a better representation of the relative performances of the forecasting methods, the mean absolute difference (over a number of years) between each of the forecasts and the corresponding observed yields was calculated at each of the lead times and plotted in Figure 33. This plot indicates that, on average, the *ACRU*-Thompson yield forecasts were closer to the observed yields than the other forecasts, and that this trend was consistent across all lead times. The differences between the various forecasts were, however, noted to be relatively small. All forecasts were closer to the observed yields than to the median of the observed yields.

Reference to Figures 32 and 33 indicates that yield forecasts improve as a season progresses, with the biggest improvement occurring between the 12.5 and 8.5 month lead times (September and January forecasts). There is little improvement after the 6.5 month lead time, except for the MGB forecasts. The MGB forecasts allow for improvements later in the season, as they are adjusted at each lead time based on available crop production figures for earlier periods in the season. The other methods of forecasting were only adjusted by more indirect means, through the updating of model inputs (observed rainfall records), which then influence the yields forecasted.

The *ACRU*-Thompson model gives a good indication of seasonal yield at the 8.5 month lead time forecast (January forecast). This indication improves even further at the subsequent 6.5 month lead time forecast. The SRM behaves in a similar manner to the *ACRU*-Thompson model, except

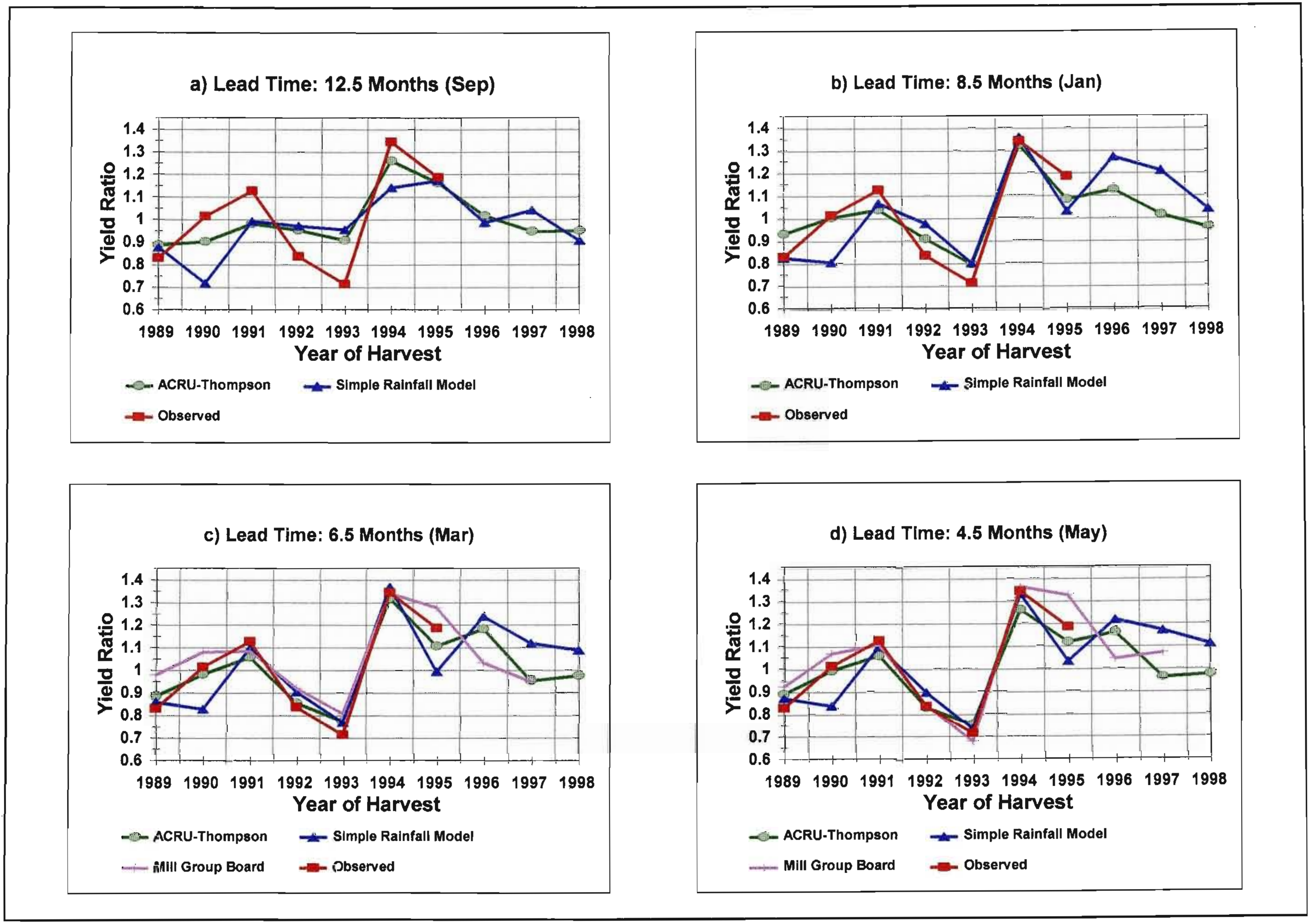


Figure 32 Eston Mill Supply Area yield forecasts (*ACRU*-Thompson, Simple Rainfall Model, Mill Group Board) at various lead times

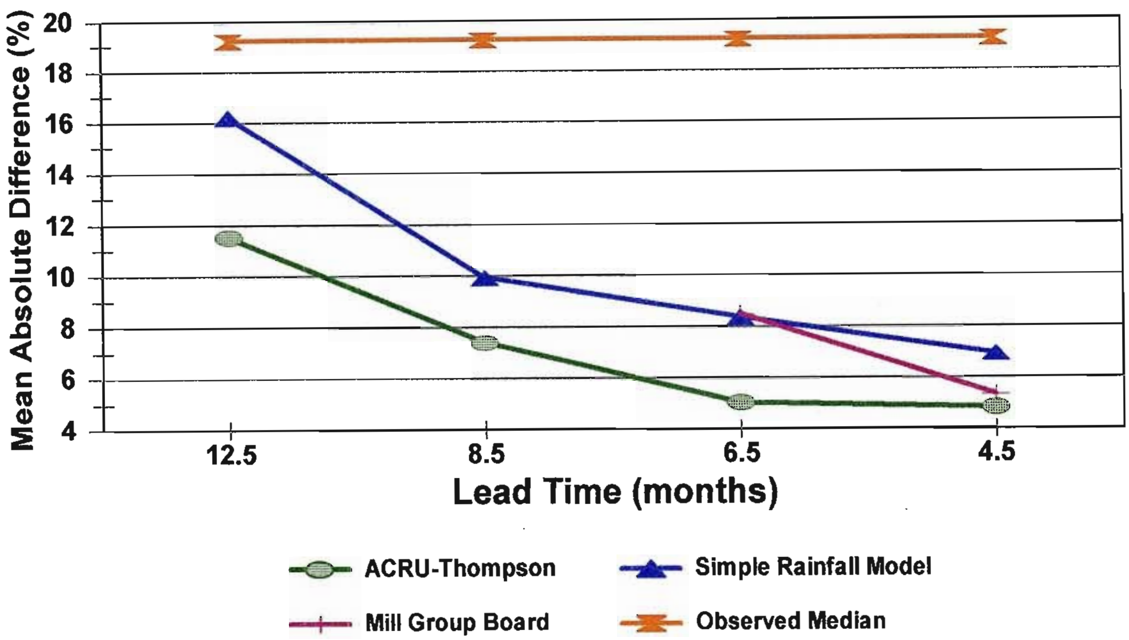


Figure 33 Mean absolute difference (over a number of seasons) between forecasted and observed yields at various lead times

that some seasons are poorly forecasted throughout the various forecast dates (eg. 1990, 1995). On average, at a lead time of 8.5 months, both the *ACRU*-Thompson and SRM models give rise to forecasts that are within 10% of the observed yield. The MGB forecasts, when they become available at the 6.5 month lead time, are also within 10% of the observed yields, on average.

To assess whether certain seasons were forecasted better than others, plots of yield ratios versus lead time were created for all of the harvest seasons where observed yield data were available. These plots are shown in Figure 34 for the 1989 to 1995 harvest seasons. The indication from these plots is that the *ACRU*-Thompson model performs better, relative to the SRM, in certain seasons. For example, the *ACRU*-Thompson forecasts for the 1990 and 1995 harvest seasons were noticeably more accurate than those of the SRM, while for other seasons the yield forecasts derived from the two methods tended to be relatively similar. The better performance of the *ACRU*-Thompson model in the 1990 and 1995 harvest seasons could be attributable to the weak, or variable, ENSO activity observed in the growing cycles of those seasons crops. This would have resulted in observed yields being influenced less by rainfall. As the *ACRU*-Thompson model accounts for a variety of yield influencing factors in addition to rainfall, it was better able to represent the observed yields of those seasons. The performance of the *ACRU*-Thompson model

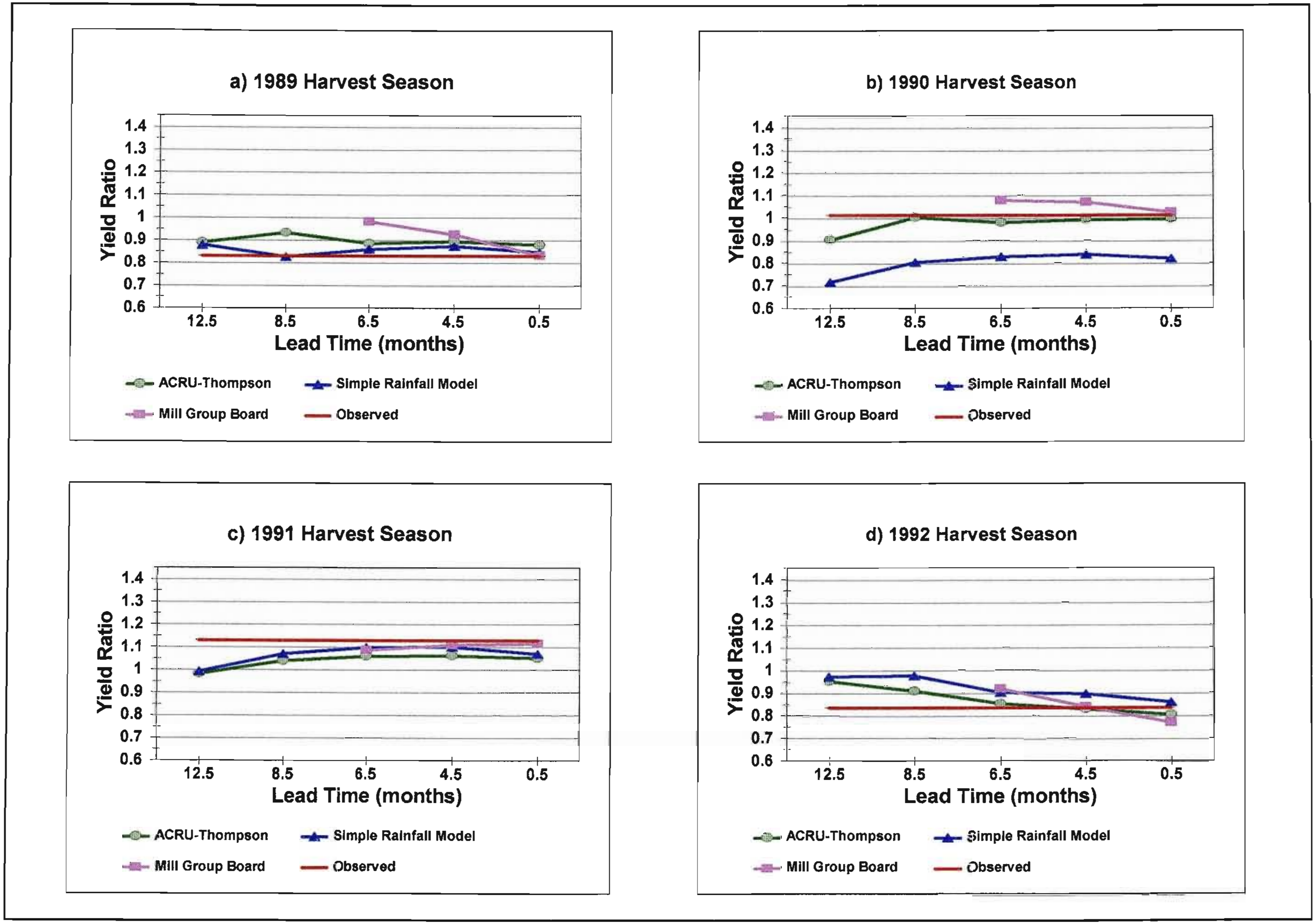


Figure 34 Eston Mill Supply Area yield forecasts (*ACRU*-Thompson, Simple Rainfall Model, Mill Group Board) for the 1989 to 1995 harvest seasons

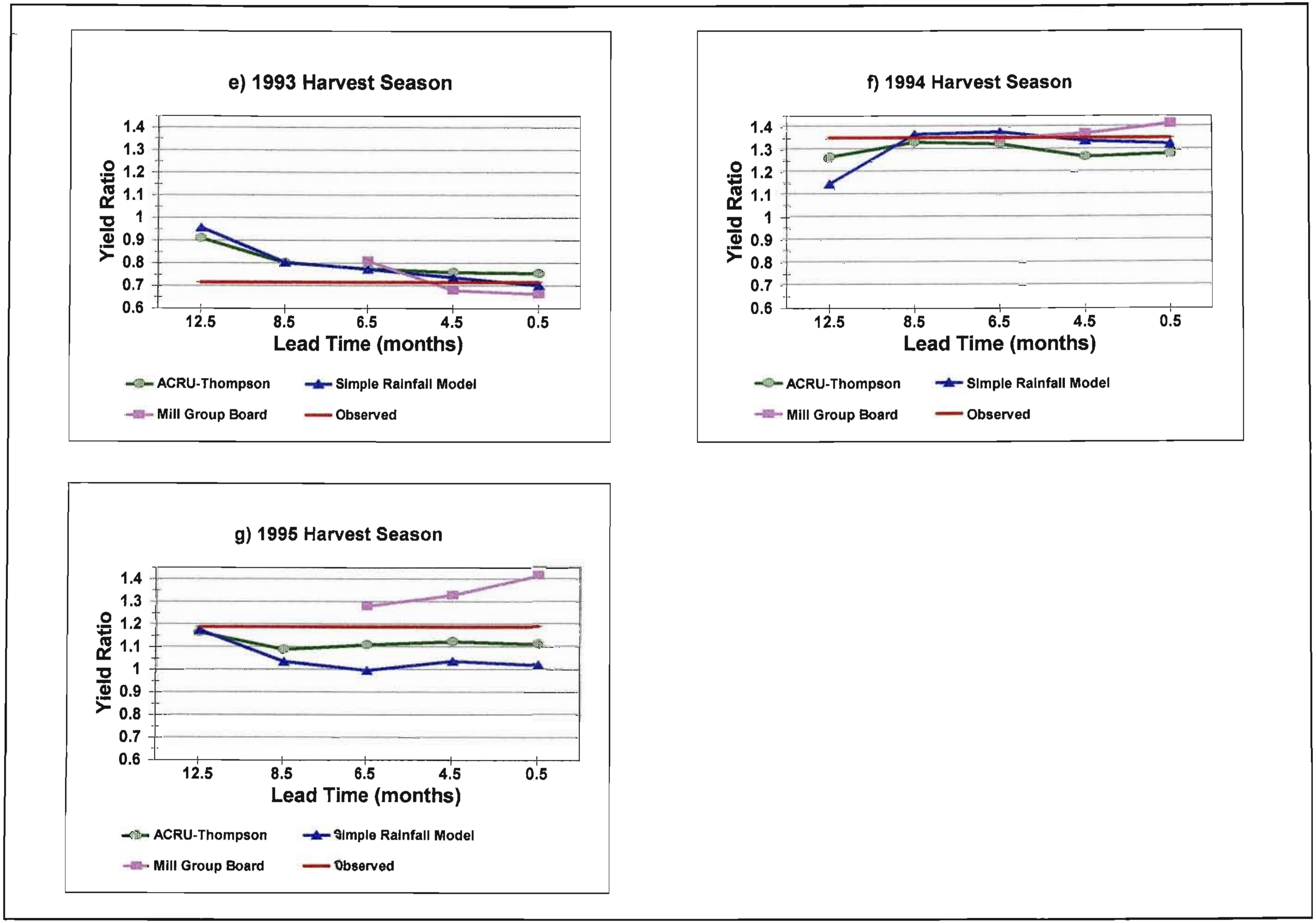


Figure 34 continued Eston Mill Supply Area yield forecasts (*ACRU*-Thompson, Simple Rainfall Model, Mill Group Board) for the 1989 to 1995 harvest seasons thus tends to be more consistent than the SRM. The performance of the MGB forecasts was generally consistent across the seasons. The accuracies of the forecasts for the 1995 harvest season were poor relative to the other seasons.

7.3.2 Yield forecasts derived from actual versus perfect rainfall forecasts

In order to assess whether yield forecasts would improve if perfect rainfall forecasts were available, yield forecasts generated using perfect and actual rainfall forecasts were compared for both the *ACRU*-Thompson and SRM models. These comparisons are shown in Figures 35 and 36 for the *ACRU*-Thompson and SRM models respectively, where forecasts are plotted against time for the various lead times. The plots for the *ACRU*-Thompson model indicate that there is little difference between the two sets of yield forecasts, except at the 8.5 month lead time where differences are slightly larger. This is so despite the actual and perfect rainfall forecasts being generally very different (Figure 28). By the 6.5 month lead time there is almost no difference between the sets of yield forecasts. This could be ascribed to the onset of winter, where growth is reduced as a result of lower temperatures and rainfall. The accuracy of rainfall forecasts would be less critical during such a period. The crops are also nearing the end of their cycles. For the SRM model, there are larger differences between yield forecasts derived from actual and perfect rainfall forecasts, most notably at the longer lead times. These differences can be ascribed to the greater sensitivity of the SRM to rainfall, as compared to the *ACRU*-Thompson model, where other factors influencing growth are considered. For the SRM model, yield forecasts were generally more accurate when derived from perfect rainfall forecasts.

7.3.3 Ranges in yields forecasted by the *ACRU*-Thompson model

Thus far in the analyses of results, only the median of the range in yields forecasted at each lead time has been considered for the *ACRU*-Thompson model. Any yield falling within this range, has a probability of occurring in the season being considered. When analysing yield forecasting results, it is important to gain an appreciation of the range in yields likely for a season. To assess the range in yields forecasted for the various seasons, the 25th percentile (representative of low probability of occurrence) and 75th percentile (representative of high probability of occurrence) yields were extracted for each season, and plotted against time for the various lead times (Figure 37). The median (50th percentile) yields presented in earlier plots are also shown.

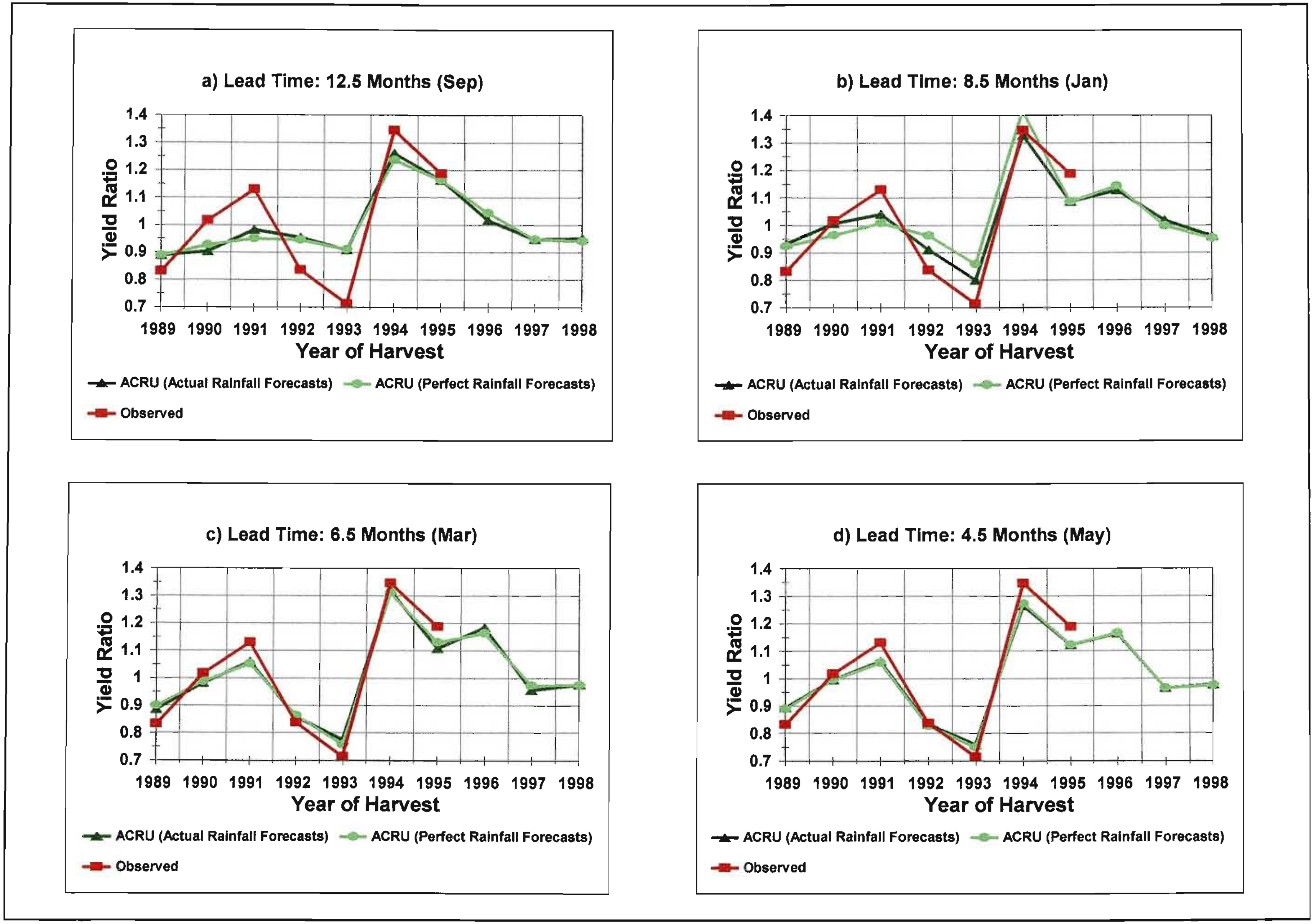


Figure 35 *ACRU*-Thompson yield forecasts at various lead times for the Eston Mill Supply Area, using actual and perfect rainfall forecasts

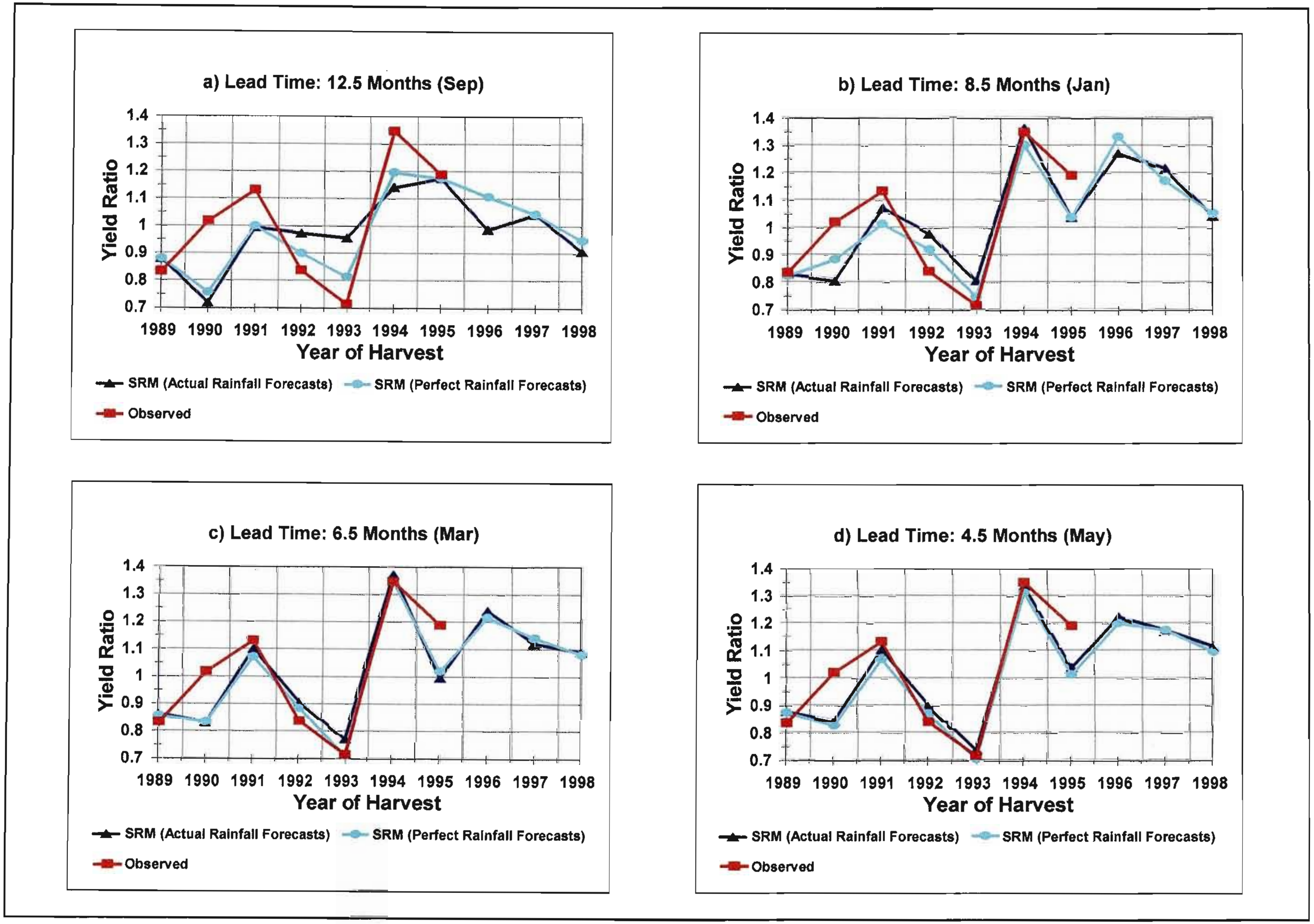


Figure 36 Simple Rainfall Model yield forecasts at various lead times for the Eston Mill Supply Area, using actual and perfect rainfall forecasts

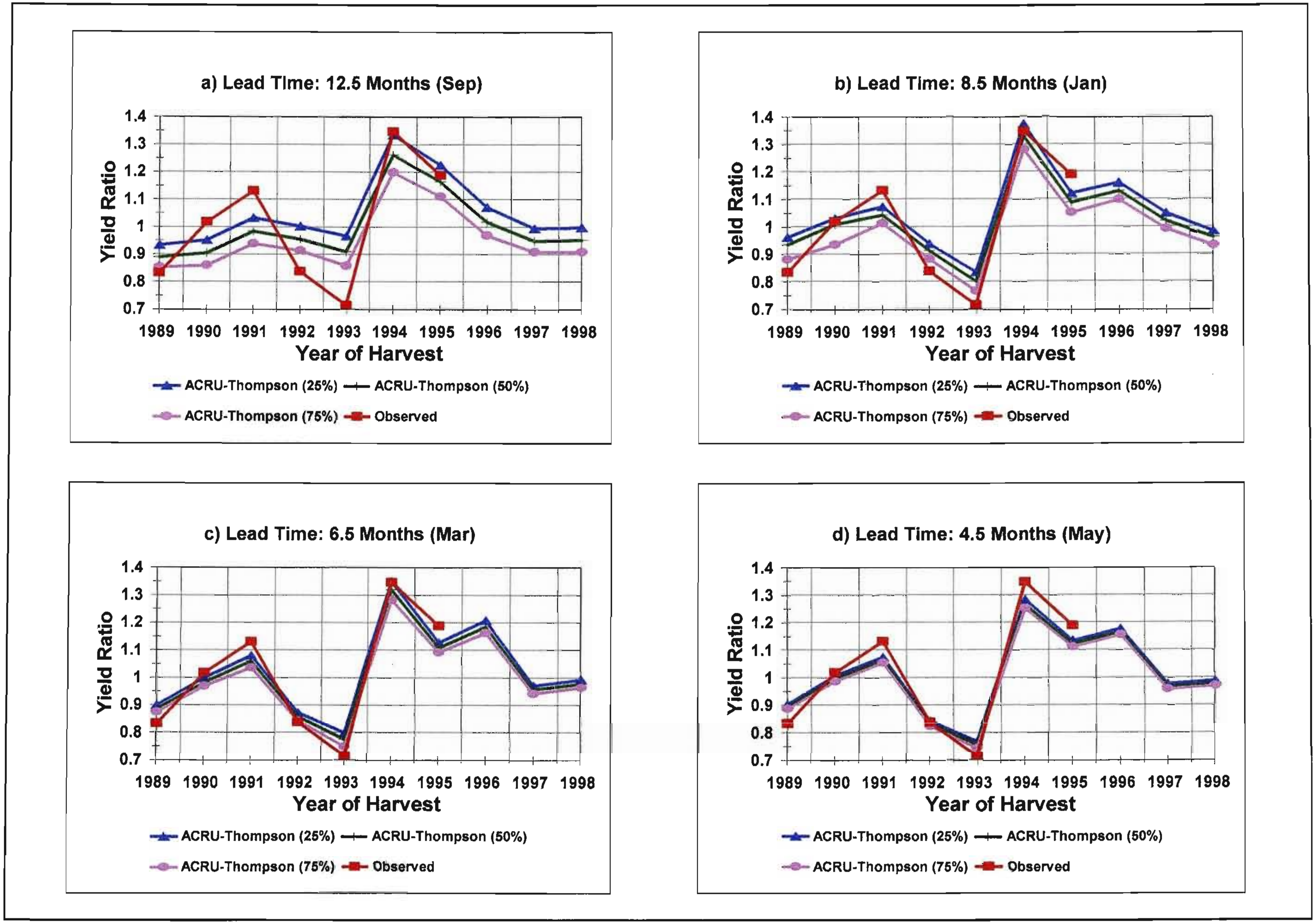


Figure 37 *ACRU*-Thompson yield forecasts for the Eston Mill Supply Area at various lead times and probabilities of exceedence

The plots indicate that the range in yields possible for a season becomes progressively smaller as lead time decreases. This is as a result of the reduced uncertainty associated with the remaining seasonal rainfall. It was calculated, on average, that the 25th percentile yields at the 12.5 month lead time varied by 5.3% from the median, while at the 6.5 month lead time they varied by only 1.9%. This variation in yields is relatively small, considering the large number of analogue rainfall years used in yield forecasting.

7.3.4 Benefit analyses of the application of different yield forecasting methods

Simple analyses were performed to assess whether benefits could be derived from the use of *ACRU*-Thompson yield forecasts, compared with yield forecasts derived from traditional forecasting methods. These analyses first involved assessing the relative accuracy of the methods, with "benefits" and "losses" being associated with the *ACRU*-Thompson yield forecasts, depending on whether they were more accurate, or less so, than the other methods. This was followed by a consideration of the cost-benefits associated with applying the different methods.

7.3.4.1 Relative accuracy of methods

For the various harvest seasons and lead times, comparisons were made between the *ACRU*-Thompson yield forecasts and the corresponding forecasts of the other methods. The methods were compared in terms of their accuracy relative to observed yields. Comparisons of *ACRU*-Thompson yield forecasts with the observed median yield, MGB forecasts and SRM forecasts are shown in Figure 38. The medians of the ranges in yields forecasted by the *ACRU*-Thompson model were again considered. Yield forecasts derived from both actual and perfect rainfall forecasts were considered. If an *ACRU*-Thompson forecast was better than the other method's forecast, then the block representing that harvest season and lead time was shaded green, in order to indicate a "benefit" from the use of the *ACRU*-Thompson forecast. If it was worse than the other forecast the block was shaded red, indicating a "loss", and if the two forecasts were within 2.5% of each other, the block was shaded grey, indicating a "neutral" situation.

In the case of the comparison with the median yield, the *ACRU*-Thompson forecasts gave rise to benefits throughout all lead times, except for the 1990 and 1991 harvest seasons. This applied to

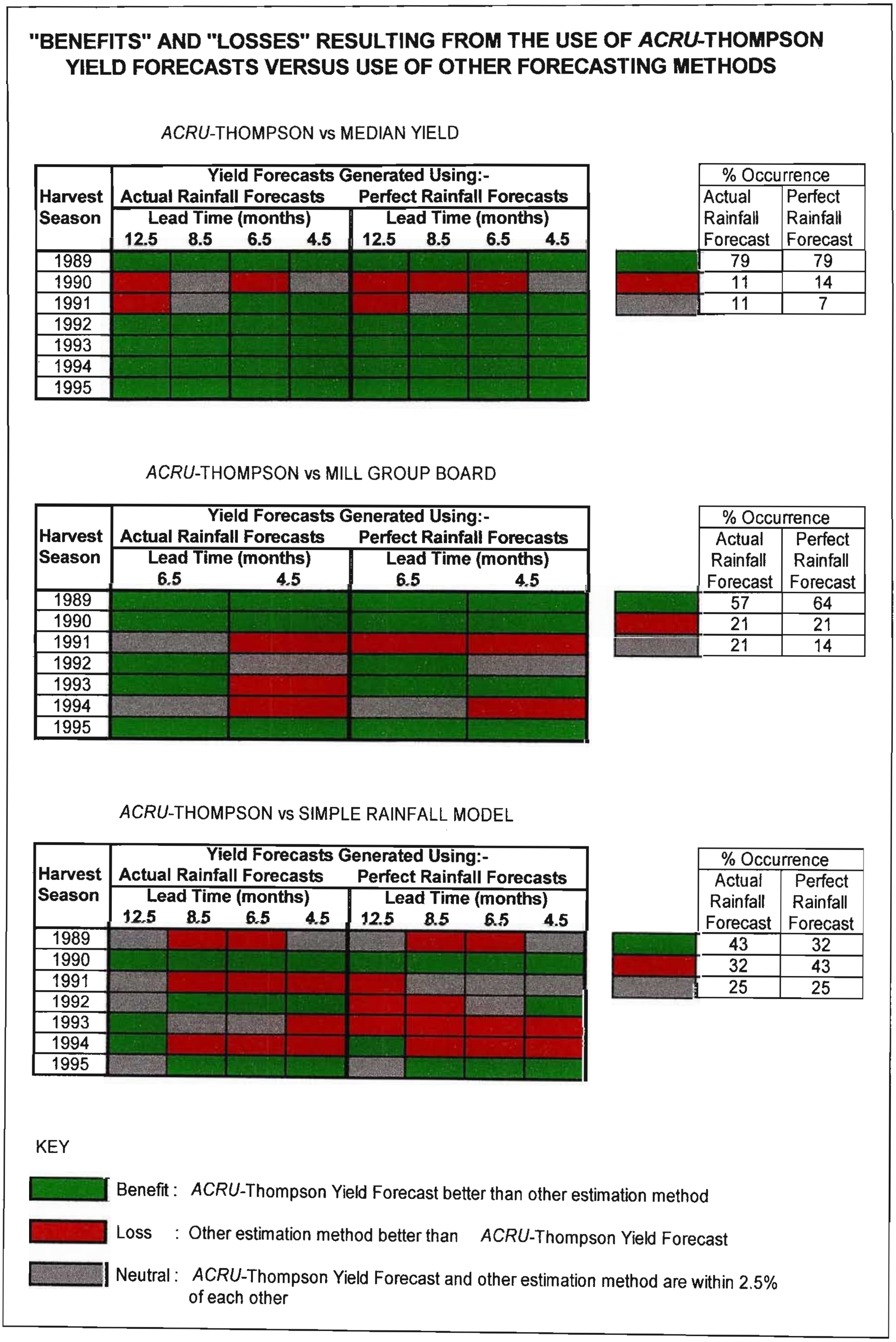


Figure 38 "Benefits" and "losses" resulting from the use of *ACRU*-Thompson yield forecasts versus use of the observed median yield, MGB forecasts and SRM forecasts

ACRU-Thompson yield forecasts derived from both actual and perfect rainfall forecasts. This is likely to be as a result of the observed yields in those years being close to the median yield. This might be explained by the weak, or variable, ENSO activity experienced during the growth cycles of those seasons' crops.

In the comparison with MGB forecasts, the *ACRU*-Thompson yield forecasts gave rise to more benefits at the 6.5 month lead time than at the 4.5 month lead time. This applied to yield forecasts derived from both actual and perfect rainfall forecasts. The greater proportion of losses to benefits at the 4.5 month lead time may be explained by the nature of the MGB, where adjustments for greater accuracy are applied, based on consideration of earlier season production figures and other factors.

In the comparison with SRM forecasts, *ACRU*-Thompson forecasts derived from actual rainfall forecasts gave rise to slightly more benefits than losses. When perfect rainfall forecasts were applied, however, there were slightly more losses than benefits. As discussed earlier, the SRM is entirely dependent on rainfall, and would thus be more sensitive to a change in rainfall input. The *ACRU*-Thompson model gave rise almost entirely to benefits during the 1990 and 1995 harvest seasons (using both actual and perfect rainfall forecasts) in which the weak, or variable, ENSO activity experienced during the growth cycles of those seasons' crops would have resulted in observed yields being less influenced by rainfall. As the *ACRU*-Thompson model accounts for a variety of yield influencing factors in addition to rainfall, it was better able to represent the observed yields of those seasons.

The comparisons in Figure 38 do not give an indication of the degree to which one method is more representative of observed yields, than another. This has been considered in earlier sections of this chapter, and should be borne in mind when assessing the benefits and losses of applying the various methods.

7.3.4.2 Cost-benefits of methods

The above analysis relating to the relative accuracies of the methods, was taken further through consideration of the economic consequences of their application in yield forecasting. To this end, a simple (first-level) cost-benefit analysis was conducted, where the economic benefits of applying the methods were assessed against their costs of implementation. The medians of the ranges in yields forecasted by the *ACRU*-Thompson model were again considered. Only yield forecasts derived from actual rainfall forecasts were assessed. The approach adopted in performing the cost-benefit analysis was based on the ideas of Schmidt, as reported in Lumsden *et al.* (1999).

Economic benefits of yield forecasting

In order to assess the economic benefits of forecasting, consideration was given to the costs associated with inaccurate yield forecasts derived from the different methods, including use of the observed median yield. The costs associated with the different methods were compared to those of the observed median, which formed the base for comparisons. The improved accuracy of the various methods relative to the observed median, and the associated reductions in cost, were considered to be the economic benefits (savings) of forecasting.

The economic costs associated with inaccurate forecasts were derived from those predicted by the LOMS model (Chapter 2) for the Noodsberg Mill Supply Area. Costs accounted for in the LOMS model relate to both the miller and the growers in the area. LOMS model results were not available for the Eston Mill Supply Area, but they were for the Noodsberg Mill, situated approximately 60 km away in a similar climate and soil region (Schulze, 1997). Although the cost structures at the Noodsberg Mill are different to those at Eston, it is believed that some value can be derived from use of the Noodsberg costs, as they provide an idea of the relative costs associated with application of the different methods. The absolute costs of applying the different methods would not necessarily be representative for Eston.

The costs of over- or underestimating a 1.5 million ton crop for a milling season, as discussed in Chapter 2, are plotted in Figure 39. The costs of underestimating a crop are greater than overestimating, as an underestimation implies that a mill must operate later into the wet season in order to complete the crushing of a crop. This period is less favourable for milling.

The costs associated with inaccurate yield forecasts generated by the various yield forecasting methods were estimated for the 1989 to 1995 harvest seasons. These costs were estimated for the March forecasts (6.5 month lead time), as it was believed that these projections would have the most influence on mill operating decisions. The costs were determined by first considering the

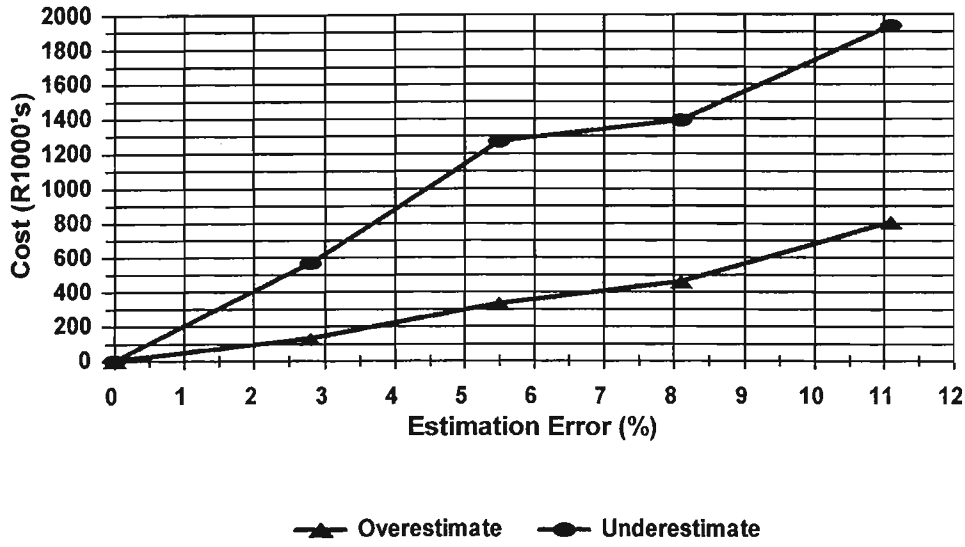


Figure 39 Costs to the Noodsberg miller and growers of poor estimation of a 1.5 million ton cane crop according to the LOMS model (Hildebrand, 1998a)

estimation errors of the various forecasting methods, i.e. differences in yield were calculated between each model's forecast and the corresponding observed yield. These errors had been calculated previously in the development of the "benefit"/ "loss" tables in Figure 38. The errors that were calculated are tabulated in Table 12. The estimation errors were then used to predict the

Table 12 Estimation errors of the various yield forecasting methods for the 1989 to 1995 harvest seasons

Harvest Season	Estimation Error (%)			
	Median Yield	Mill Group Board	Simple Rainfall Model	*ACRU*-Thompson
1989	22.1	18.1	3.6	6.5
1990	0.0	6.3	-18.2	-3.3
1991	-10.0	-3.8	-2.7	-6.2
1992	21.3	10.1	8.1	2.2
1993	42.2	13.2	8.0	8.4
1994	-24.5	-0.3	1.8	-2.1
1995	-14.3	7.7	-16.1	-6.6

costs of forecast inaccuracy through reference to the cost curves in Figure 39. Where estimation errors were greater than the maximum errors shown in Figure 39, a linear extrapolation of the cost curves was assumed. The average seasonal costs over the period considered were then calculated for each forecasting method. These costs are shown in Figure 40 for the Eston Mill Supply Area for the different forecasting methods applied (based on Noodsberg LOMS model cost figures). The use of the observed median yield as a predictor of seasonal yield is also represented in Figure 40. The average seasonal cost of inaccurate yield forecasts is lowest for the *ACRU*-Thompson forecasts, followed by the MGB and SRM forecasts. The MGB forecasts are associated with lower costs than the SRM because the estimation errors of the former are more consistent, whereas the SRM forecasts result in both large and small estimation errors, depending on the season. The SRM gave rise to a large underestimation in the 1995 harvest, thus further contributing to the lower average costs of the MGB forecasts relative to the SRM forecasts.

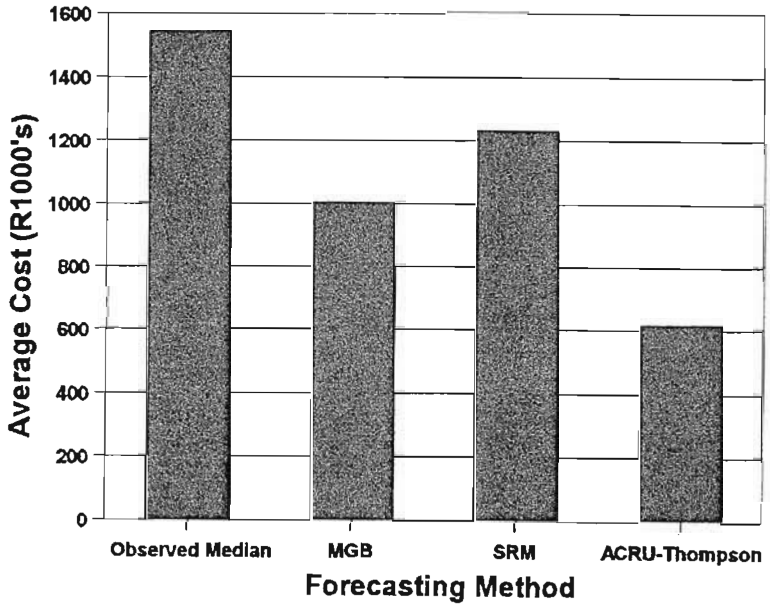


Figure 40 Average seasonal cost for the Eston MSA of inaccurate crop forecasting in March when applying different forecasting methods

The average seasonal economic benefits (reductions in costs) of the different forecasting methods relative to the observed median are presented in Table 13. These savings are determined from the costs reflected in Figure 40.

Table 13 Average seasonal economic benefits of applying various yield forecasting methods in the Eston MSA relative to use of the observed median yield

Method of Forecasting	Economic Benefit (R1000's)
Mill Group Board	543.3
Simple Rainfall Model	314.7
ACRU-Thompson	929.9

Costs of implementation of different yield forecasting methods

The costs associated with implementing the *ACRU*-Thompson and SRM forecasting methods on a monthly interval were estimated based on the time required to set up and maintain each method. A labour cost of R150 per hour was assumed. No costs were assigned to computer hardware, software or data. The estimated time required to complete each step in the setup and maintenance of the methods is contained in Appendix C, along with the associated costs. Maintenance costs were separated into monthly and annual costs. The implementation costs of the two methods are summarized in Table 14, with a total of costs being given after one year of implementation.

Table 14 Summary of implementation costs of *ACRU*-Thompson and Simple Rainfall Model based yield forecasting systems for the Eston MSA

Task	*ACRU*-Thompson		Simple Rainfall Model	
	Required time (h)	Cost (R)	Required time (h)	Cost (R)
Initial set-up	494	74100	105	15750
Monthly maintenance	12 * 41	73800	12 * 8	10800
Annual maintenance	1 * 36	5400	1 * 5	750
Total after one year	1022	153300	182	27300

The estimated cost after one year of implementing the SRM system was approximately 18% of that estimated for the *ACRU*-Thompson system over the same period.

No set-up costs were attributed to the MGB forecasts as this system of forecasting is already operational. Maintenance costs relating to this system were not accounted for as they were not readily available.

Cost-benefits of the yield forecasting methods

The economic benefits (savings) of the different yield forecasting methods can be assessed against their costs of implementation in order to estimate the net economic benefit of the methods. The annual net economic benefits that could be expected for the forecasting methods were calculated and plotted in Figure 41. The corresponding gross benefits and costs of implementation, as calculated in Tables 13 and 14, were also plotted.

The net economic benefit of the *ACRU*-Thompson method was found to be highest, followed by the MGB and SRM methods. These net economic benefits account for initial set-up costs, implying that in a year subsequent to implementation of the system, net benefits would increase for the *ACRU*-Thompson and SRM methods. The net economic benefit associated with the MGB forecasts may be inflated relative to the other methods, as the costs of maintaining this system were not accounted for.

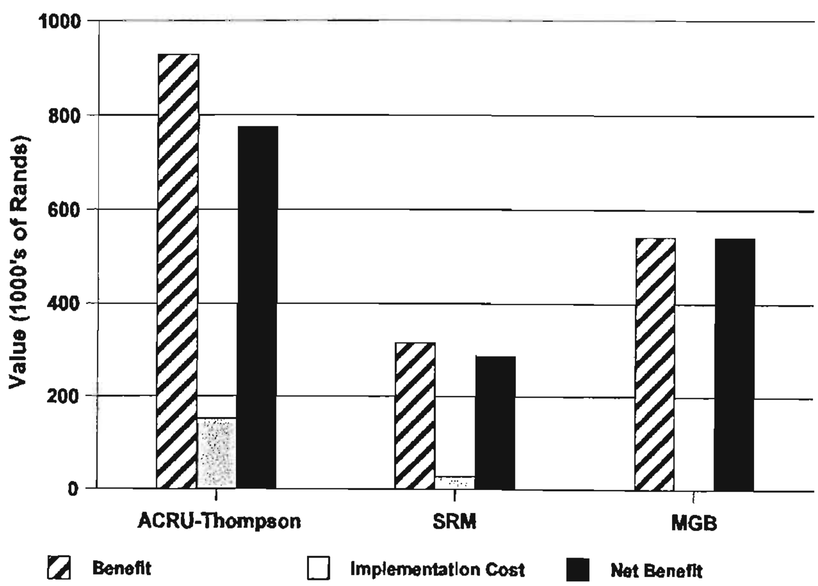


Figure 41 Annual benefits, costs of implementation (including set-up costs) and resulting net benefits of various yield forecasting methods at Eston

8 DISCUSSION

The main objective of this project involved the development and evaluation of a yield forecasting system for the Eston Mill Supply Area using yield simulation models and seasonal rainfall forecasts. Two daily time step yield simulation models, namely the *ACRU*-Thompson and CANEGRO-DSSAT models, were initially evaluated to verify their ability to accurately simulate historical yields given an observed rainfall record. The model found to be most appropriate for yield forecasting at Eston, the *ACRU*-Thompson model, was then used to generate yield forecasts for a number of seasons, through the application of seasonal rainfall forecasts in the model. These rainfall forecasts had previously been translated into daily rainfall values for input into the model. The sugarcane yield forecasts were then evaluated against observed yields as well as against forecasts generated by more traditional methods, these methods being represented by a Simple Rainfall Model and Mill Group Board estimates.

In the model verification phase of the research, the preparation of climate inputs required that missing records of climate be infilled and adjusted to represent the climate at each of the farms associated with those climate stations. For the period of simulation, there were numerous climate stations in and around the MSA that facilitated the preparation of climate inputs. Eight driver stations were available, with an additional six stations available for infilling missing values.

In the preparation of soils inputs, the Land Type and soil parent material information used to develop the required input values, was found to give very similar values over much of the MSA when expressed in terms of total available soil moisture. Neither source of soils information appeared to result in simulations that were more accurate than the other, although a greater sensitivity of the CANEGRO model to soils inputs became evident. If the similarities in TAM derived from LT and PM soils information is found in other regions of the industry, then the use of LT information in modelling would be recommended, given the widespread availability of this information and the ease with which model inputs can be developed.

Based on a study of observed field scale yields in the Eston MSA, it was verified that growth cycles do affect the yields of sugarcane. An investigation of the proportions of the various growth cycles practised in the MSA indicated that certain cycles were more dominant. It was later shown that the yields in the MSA could be simulated adequately by using only the four most dominant cycles of the 11 that are used in practice.

When verifying the performance of the different models, not all possible factors affecting production were accounted for in the modelling approach. For example, factors such as frost and low plant population following drought were not accounted for. These factors may have a significant effect on yields in certain years. If these influences are believed to be significant for a large proportion of the MSA, then they should be accounted for in some way in the modelling approach.

The *ACRU*-Thompson model was selected for yield forecasting based on its consistent, robust performance at the required scale. The information available for preparation of soils inputs was believed to be of a level of detail more suited to this model. The ability of the *ACRU*-Thompson model to predict MSA yields at sub-area scale was also shown. This coarser scale of modelling, which was adopted for yield forecasting, significantly reduced the number of simulations required to implement the yield forecasting methodology. The preparation of model inputs for sub-area scale modelling was found to be greatly simplified, as a result of inputs having previously been developed at farm scale.

In the yield forecasting phase of the research, the skill of the statistically generated categorical seasonal rainfall forecasts was found to be disappointingly poor, although the accuracy of these forecasts was later found to exert little influence on the resulting yield forecasts, especially in the case of the *ACRU*-Thompson forecasts. The very localized scale of the rainfall forecasts could possibly account for their inaccuracy. The location of the MSA within KwaZulu-Natal may also be a contributing factor, as seasonal rainfall tends to be less predictable in this region. Climatologists may in future make use of atmospheric general circulation models (GCMs) to improve seasonal rainfall forecasting. These models account for various rainfall influencing factors, and are not restricted to accounting primarily for ENSO related factors. In addition, GCMs are dynamic in that they attempt to simulate the physical processes governing rainfall for a particular period, as opposed to statistical models which predict rainfall based on static relationships developed from historical records. At present, the lead time of GCMs is restricted to a monthly time scale, thus limiting their application in seasonal rainfall forecasting.

The *ACRU*-Thompson yield forecasts were found to be, on average, the most accurate of the various forecasting methods assessed. When compared to the SRM, it was found that the *ACRU*-Thompson model was more consistent in the accuracy of its predictions over the various seasons considered. This is a result of its accounting for a variety of yield influencing factors such as daily rainfall, a daily soil water budget and temperature as a growth driver, compared to the SRM which only takes account of monthly rainfall. This consistency was particularly noticeable in seasons when rainfall did not exert a strong influence on yields. In seasons where rainfall did influence yields strongly, the SRM tended to perform well, particularly when perfect rainfall forecasts were used. When *ACRU*-Thompson yield forecasts were compared to MGB forecasts, the accuracy of the MGB forecasts was noted to improve relatively at the 4.5 month lead time. This lead time is relatively short and would not allow a great deal of time for planning decisions. The *ACRU*-Thompson yield forecasts would thus offer a better alternative for longer lead time planning. When comparing *ACRU*-Thompson yield forecasts to the median yield of the Eston MSA, the *ACRU*-Thompson model gave a better representation of the seasonal yield for those seasons which were more strongly influenced by ENSO events.

The period of rainfall between the date of forecast and date of harvest was shown to exert little influence on the yields predicted by the *ACRU*-Thompson model. This was reflected in the small ranges in yield predicted by the model (corresponding to the multiple analogue rainfall years considered in each forecast), as well as in the small differences found between yield forecasts derived from actual and perfect rainfall forecasts. At longer lead times, the SRM gave rise to larger differences between yields derived from actual and perfect rainfall forecasts, as a result of its greater sensitivity to rainfall inputs. Since the rainfall forecasts were found to be inaccurate, any accuracy attained in yield forecasting must have been derived from the observed rainfall available at the time of forecasting. The crops must therefore have been well established by this time. The influence of winter towards the end of the season would also be a contributing factor, as lower rainfall and temperatures would result in growth being less significant during this period.

The above findings, however, may not necessarily be applicable in areas having shorter growth cycles, where crops are more dependent on the rainfall of a single summer season, such as in rainfed coastal areas. In such areas, it is likely that there would be a greater reliance on the accuracy of the rainfall forecasts, in order for accurate yield forecasts to be obtained.

The methodology developed by Lecler, as reported in Lumsden *et al.* (1999), to translate the categorical seasonal rainfall forecasts into daily rainfall values required by the *ACRU*-Thompson model, represented a complex first approach. A more simplified methodology was used to translate the rainfall forecasts into monthly rainfall required by the SRM, as this model was spreadsheet-based and was not suited to the complex methodology employed for the *ACRU*-Thompson model. For application of the *ACRU*-Thompson model in areas similar to Eston, it is likely that the large number of simulations could be reduced, by selecting only certain analogue years to represent rainfall for a season. For example, years of rainfall representative of the current categorical forecast that correspond to the 15th, 50th and 85th percentile level, could be selected. In rainfed coastal areas, however, it may be appropriate to adopt a more thorough approach to representing the rainfall which occurs between forecast and harvest dates, given the greater reliance of the crops on this rainfall.

In the simple cost-benefit analysis conducted in the study, the *ACRU*-Thompson model gave rise to the highest net economic benefits, on average, when compared to benefits derived from the traditional forecasting methods of the SRM and MGB. Although the benefits associated with applying the forecasting methods were derived from economic figures for the Noodsberg area, the calculation of the benefits for each method relative to the median yield ensured that the analysis was more appropriate for the Eston MSA. The relative net economic benefits of the methods were of primary interest in the analysis. This cost-benefit analysis centred around improvements in forecast accuracy (and thus improvements in mill operating decisions) and the cost of implementing a yield forecasting system. There are potentially many other benefits and costs associated with yield forecasting, particularly at scales other than the Mill Supply Area, eg. at farm and national scale.

The observed yield records against which the various forecasting techniques were evaluated was relatively short. A longer record would allow for further verification of the conclusions made in this study. For example, marked differences between the various forecasts were evident for the 1996 to 1998 harvest seasons, a period when observed yields were not readily available for forecast verification. The relative performances of the different forecasting methods during this time may give a further indication of their respective strengths and weaknesses.

Of the methods assessed in this study, the *ACRU*-Thompson model gave rise, on average, to the greatest accuracy in yield forecasting and the highest net economic benefits. This lends support to the adoption of a simulation-based yield forecasting system in the Eston MSA. It is stressed that the above trends in accuracy and net economic benefit were observed on average, and that these trends would not necessarily be evident in every year. If the *ACRU*-Thompson model were applied in a rainfed coastal area, where the accuracy in yield forecasting is uncertain, the benefits associated with applying the method may not outweigh the costs of implementation, as observed previously for the Eston MSA. Thus, if the *ACRU*-Thompson model were to be considered for application in such an area, it is recommended that the relative accuracies of the methods, and the resulting net economic benefits, be re-assessed.

The crop forecasting methodologies assessed in this study, with the exception of the MGB, give rise to forecasts of cane yield (t/ha). For many purposes, a forecast of the actual production for a season is required, i.e. the tonnage of cane. This implies that estimates of the area under cane are also required. The use of remote sensing may allow for timely and accurate estimation of area under cane. In forecasts of seasonal production, practical management issues such as transfers of cane between mills, also need to be taken into account.

9 CONCLUSIONS

The main conclusions drawn in this study may be summarized as follows:

a) Values of soil TAM derived from Land Type and soil parent material information were found to be very similar.

b) The CANEGRO-DSSAT model was found to be more sensitive to soils inputs than the *ACRU*-Thompson model. The level of soils information available for development of model inputs at the selected scale of modelling, led to the conclusion that the latter model would be more suited to application in yield forecasting.

c) Growth cycles of sugarcane affect yields. The crop at Eston can be represented with the *ACRU*-Thompson model by considering the four most dominant cycles (of the 11 practiced).

d) The categorical seasonal rainfall forecasts applied in this study exhibited poor forecast skill. This could be as a result of their very localized scale and the location of the Eston MSA within the country.

e) All yield forecasting techniques generally gave rise to improved yield forecasts as lead time decreased.

f) The *ACRU*-Thompson yield forecasts were, on average, more accurate than those of the other yield forecasting methods. The SRM yield forecasts were relatively more accurate in seasons where rainfall exerted a strong influence on yields. Both methods of forecasting generally produced forecasts within 10% of observed yields at a lead time of 8.5 months (January forecast). MGB forecasts, which are available later in the season, showed good improvements in accuracy as the seasons came to a close. Forecasts were generally within 10% of observed yields at a 6.5 month lead time (March forecast).

g) The period of rainfall between dates of forecast and dates of harvest was shown to exert little influence on the yields forecasted by the *ACRU*-Thompson model. This was reflected:
   - in the methodology for translating the categorical seasonal rainfall forecasts into daily rainfall through use of multiple analogue years of rainfall, which gave rise to small ranges in forecasted yields; and
   - in the small differences found between yield forecasts derived from actual and perfect rainfall forecasts.

h) Point g) above indicates that for the Eston MSA:

- crops are well established when yield forecasting commences and the observed rainfall available is sufficient for the *ACRU*-Thompson model to give a good representation of the seasonal yield; and
- the influence of winter towards the end of the season results in growth being less significant, owing to lower rainfall and temperatures occurring during this period.

i) The rainfall occurring between forecast and harvest dates was shown to exert more influence on the yields of the SRM at longer lead times, owing to its greater sensitivity to rainfall inputs.

j) The findings in g) and h) above may not necessarily be applicable in areas having shorter growth cycles, where crops are more dependent on the rainfall of a single summer season, such as in rainfed coastal areas.

k) A simple cost-benefit analysis indicated that for the Eston MSA the *ACRU*-Thompson system could potentially give rise to greater net economic benefits when compared to those using traditional yield forecasting methods. This cost-benefit analysis, which accounted for the relative accuracies of the methods, lends support for the adoption of a simulation-based yield forecasting system. The above trends, however, were observed when several years' simulations were averaged, and would not necessarily be true for every individual year.

10 RECOMMENDATIONS

The research project upon which this dissertation is based, required that recommendations be made for the practical implementation of a yield forecasting system in the South African sugar industry. Areas of future possible research were also to be identified. The recommendations made are discussed in this chapter.

10.1 Practical Application

It is believed that a yield forecasting system involving the use of a yield simulation model and seasonal rainfall forecasts may be appropriate for practical application in the sugar industry. To be more certain of the economic viability of such a system for a particular area, a detailed cost-benefit analysis should be conducted in conjunction with economists. This analysis would need to take into account the likely accuracy of forecasts in that area. If a simulation-based yield forecasting system were to be implemented, the following recommendations would be made:

a) Crop yield simulation models based on a daily soil water budget, such as the *ACRU*-Thompson model, are well suited to sugarcane crop yield forecasting. A model based on a simple rainfall-yield relationship is likely to perform well in seasons where rainfall exerts a strong influence on yields, but would be less successful in other seasons. Complex crop growth models such as CANEGRO, are more appropriate for assessing the impact of management practices on production, particularly at field scale, where detailed model inputs are available.

b) The mill supply area is a practical scale at which to implement a cane yield forecasting system. MSAs can be divided into several sub-areas each displaying a reasonably homogenous climate.

c) Each sub-area should be assigned a driver climate station with good quality records which can be used to form climate inputs to the model. The data from these stations should be maintained up to date so that the most recent climate data are reflected in the model forecasts. Consideration should be given to the methods employed in this dissertation for:
   - infilling of missing records using data from surrounding stations; and
   - adjusting rainfall driver station records to be representative of their respective sub-

areas through consideration of relative mean annual/median monthly rainfalls.

d) Appropriate average soil inputs to the selected model should be determined for each sub-area. As yield simulations in this project were shown to be relatively insensitive to the source of soils information used, and as the range in values of soil inputs for the various sub-areas was relatively small, it is likely that estimates of soils inputs made by local soil specialists would be sufficient for operational yield forecasting. This point is particularly relevant when yields are expressed as fractions of the previous year's yield. If yield forecasting were to be implemented in an area known to have a wide range of soil characteristics, or if a finer scale of modelling was required, such as farm scale modelling, then it is recommended that soils inputs be derived from Land Type information. Soils inputs can be prepared from this source of information with relative ease. If Land Type information is not readily available, then soil inputs can be derived from soil parent material information and a knowledge of the soils occurring on these parent materials. Slope position, slope gradient and MAP could be considered in the prediction of the soil types and depths occurring on the various parent materials.

e) The dominant growth cycles in the mill supply area should be identified and represented in the modelling strategy. Representation of growth cycles can be achieved by assigning weightings to forecasted yields, according to the proportions of area that the relevant growth cycles constitute.

f) In yield forecasting, seasonal rainfall forecasts can be used to develop suitable daily rainfall data sets to fill the climate record of a season between the dates of forecast and harvest. Rainfall forecasts specific to a localized point or area are not recommended, as skill in rainfall forecasting was found to be poor at this scale. Standard monthly rainfall forecasts, issued by the South African Weather Bureau for broad areas within South Africa, give rise to higher forecasting skills than those found in this study. The correlation between these broad scale forecasts and observed rainfall at local scale is, however, uncertain. In order to translate categorical seasonal rainfall forecasts into daily rainfall values required by a yield simulation model, analogue years in the historical time series resembling the relevant category of rainfall could be identified for the period of interest. The data from these years could then be used to fill the seasonal record for this period. It is likely that only selected years of rainfall, corresponding to certain levels of probability, would be necessary to form an adequate range of rainfall scenarios representative of the range in rainfall likely for a season. For example, years of rainfall corresponding to the 15th, 50th and 85th percentile level within the relevant category of rainfall, could be selected. A single year of data corresponding to the 50th percentile condition for the period of remaining seasonal rainfall not represented by a rainfall forecast, is also likely to be sufficient for the completion of the seasonal record (if necessary). In areas having short growth cycles, such as rainfed coastal areas, it may be appropriate to adopt the more detailed approach used in this dissertation to represent the rainfall occurring between forecast and harvest dates, given the greater reliance of the sugarcane crops on this rainfall. When updating the observed climate record of a season, all available climate data should be used. If necessary, the yield simulation model should be modified so as to incorporate daily observations of climate (eg. temperature, evaporation) up to the time of forecast, and monthly averages thereafter for the remainder of the season. If a strong relationship exists between rainfall and temperature, daily values of temperature recorded in the selected analogue rainfall years can be used to fill the temperature (and hence evaporation) records for the remainder of the season (if data are available). If forecasts of temperature were also available, then the selection of analogue years could be restricted to those years which represent the general climate (rainfall, temperature and evaporation) of the season being forecast.

g) It is recommended that operational yield forecasts be generated on a monthly basis to ensure close monitoring of the expected seasonal yield and the availability of up to date information for decision-making throughout the season.

h) The area under cane that is expected to be harvested should be determined accurately for each sub-area if cane production (tonnes) is to be forecasted. Use could be made of technology such as global positioning systems or remote sensing, to facilitate the accurate determination of area to be harvested.

i) Historical yield data from a representative sample of farms supplying cane in each sub-area should be collated to verify the accuracy of model predictions of year to year yield variation. This should preferably be carried out prior to yield forecasting, but should also be performed after each season to ensure that errors in yield forecasting are minimized as much as possible to inaccuracies in the rainfall forecasts, and to limitations in the translation of these forecasts to representative daily rainfall values.

j) Crop yield forecasts can be presented in a number of ways. Graphs or tables indicating expected yield as a fraction of the preceding year's yield remove any systematic error relating to the influence of crop management. In order to aid decision-making, crop forecasts should be presented in a manner that gives an indication of the range in yields likely for a season. The risks associated with making decisions based on the forecasts can then be assessed. Results in the form of graphs or tables can be presented for sub-areas of the MSA and/or the entire MSA.

k) The potential exists for the linking of remote sensing to an operational yield forecasting system. Remote sensing can be used as an aid in developing model inputs that are spatially representative of the area. Remote sensing can also be used to assess the vigour of crops across the mill supply area. This assessment can then be used to ascertain if the model simulations of yield are likely to represent the spatial patterns in yields of crops harvested during the season.

10.2 Future Research

The following possible areas of research are recommended for the advancement of simulation-based yield forecasting systems in the sugar industry:

a) The evaluation of the yield forecasting system in rainfed, coastal areas is recommended, as shorter growth cycles imply greater reliance of the crops on single summer rainfall seasons. The need for accurate rainfall forecasts is likely to become more apparent, as the period of observed rainfall record will be short when forecasting commences.

b) The correlation between current broad scale rainfall forecasts and subsequent rainfall observed at local mill supply area scale could be investigated.

c) The skills attained by climatologists in the field of seasonal rainfall forecasting should be monitored periodically, as new techniques or refinements to existing techniques are investigated.

d) The use of streamflow forecasting to assess water supply during dry years could be investigated for areas dependent on irrigation.

11 REFERENCES

Ångström, A., 1924. Solar and terrestrial radiation. *Quarterly Journal of the Royal Meteorological Society*. 50: 121-126.

Baier, W., 1979. Note on the terminology of crop-weather models. *Agricultural. Meteorology* 20:137-145.

Berri, G., 1995. Personal communication. IRICP pilot project leader, Lamont-Doherty Earth Observatory of Columbia University, Palisades, New York, USA.

Boote, K.J. and Loomis, R.S., 1991. The prediction of canopy assimilation, 1991. In *Modelling Crop Photosynthesis from Biochemistry to Canopy*. CSSA Special Publication No. 19:109-140.

Brodskil, MV and Kostyukov, V.V., 1992. Complexation of various models for agricultural prediction. *Soviet Meteorology and Hydrology*. 1:65-68.

Brüggemann, E.A., 1998. *Crop models pertinent to sugarcane production in South Africa*. Unpublished seminar. Department of Range and Forage Resources, University of Natal, Pietermaritzburg, RSA.

Cane, M.A., Eshel, G. and Buckland, R.W., 1994. Forecasting Zimbabwean maize yield using eastern equatorial Pacific sea surface temperature. *Nature*. 370:204-205.

Carter, J. and Brook, K., 1996. Regional applications, managing with climate variability - Developing a national drought alert framework. *Proceedings of Managing with Climate Variability Conference 'Of droughts and flooding rains'*, Canberra, Australia, 16-17 Nov. 1995. Occasional Paper CV03/95, Land and Water Resources Research and Development Corporation. 53-60.

Cousens, D., 1998. Personal communication. Illovo Sugar Limited, Mount Edgecombe, RSA.

De Jager, J.M., Potgieter, A.B. and van den Berg, W.J., 1998. Framework for forecasting the extent and severity of drought in maize in the Free State Province of South Africa. *Agricultural Systems*. 57:351-365.

Dent, M.C., Lynch, S.D. and Schulze, R.E., 1989. *Mapping mean annual and other rainfall statistics over southern Africa*. Water Research Commission Report 109/1/89. Pretoria, RSA. pp 198 plus appendices.

Du Pisani, A.L., 1987. The CERES-MAIZE model as a potential tool for drought assessment in South Africa. *Water SA*. 13:159-164.

Dubey, R.P., Ajwani, N., Kalubarme, M.H., Sridhar, V.N., Navalgund, R.R., Mahey, R.K., Sidhu, S.S., Jhorar, O.P., Cheema, S.S. and Narang, R.S., 1994. Pre-harvest wheat yield and production estimation for the Punjab, India. *International Journal of Remote Sensing.* 15:2137-2144.

Duchon, C.E., 1986. Corn yield prediction using climatology. *Journal of Climate and Applied Meteorology.* 25:581-590.

Durling, J.C., Hesterman, G.B. and Rotz, C.A., 1995. Predicting first cut alfalfa yields from preceding winter weather. *Journal of Production Agriculture.* 8:254-259.

Gardner, B.R., Blad, B.L., Garrity, D.P. and Watts, D.G., 1981. Relationships between crop temperature, grain yield, evapotranspiration and phenological development in two hybrids of moisture stressed sorghum. *Irrigation Science.* 2:213-224.

Glover, J. 1972. Practical and theoretical assessments of sugarcane yield potential in Natal. *Proceedings of the South African Sugar Technologists' Association.* 46:138-141.

Hammer, G.L., Holzworth, D.P. and Stone, R., 1996a. The value of skill in seasonal climate forecasting to wheat crop management in a region with high climate variability. *Australian Journal of Agricultural Research.* 47:717-737.

Hammer, G.L. and Nicholls, N.N., 1996. Managing for climate variability - The role of seasonal climate forecasting in improving agricultural systems. *Invited paper to Conference on Agricultural Meteorology*, Brisbane, Australia.

Hammer, G.L., Stephens, D.J. and Butler, D., 1996b. *Development of a National Drought Alert Strategic Information System: Wheat Modelling Sub-project*, Volume 6. Land and Water Resources Research and Development Corporation. 1-41.

Hellmann, D.B., 1993. The use of FRS data to interpret the effect of different growth cycles on the yield performance of variety N12. *Proceedings of the South African Sugar Technologists' Association.* 67:88-93.

Hellmann, D.B., 1997. Personal communication. Midlands South Extension Office, South African Sugar Association Experiment Station, Eston, RSA.

Heskith, J.D., Baker, D.N. and Duncan, W.G., 1971. Simulation and growth and yield in cotton: respiration and the carbon balance. *Crop Science.* 11:394-398.

Hildebrand, Q.L., 1998a. Personal communication. South African Cane Growers Association, Durban, RSA.

Hildebrand, Q.L., 1998b. Maximising profits from sugar through optimising the length of milling season. Proceedings of the Annual General Meeting of the South African Sugar Industry

Agronomists Association (26 November 1998). SA Sugar Association Experiment Station, Mount Edgecombe, RSA.

Horie, T., Yajima, M. and Nakagawa, H., 1992. Yield forecasting. *Agricultural Systems*. 40:211-236.

Hughes, A.D., 1992. *Sugarcane yield simulation with the ACRU modelling system*. Unpublished MSc Dissertation. Department of Agricultural Engineering, University of Natal, Pietermaritzburg, RSA.

Idso, S.B., Jackson, R.D. and Reginato, R.J., 1977. Remote sensing of crop yield. *Science*. 196:19-25.

Inman-Bamber, N.G., 1991a. A growth model for sugarcane based on a simple carbon balance and the CERES-Maize water balance. *South African Journal of Plant and Soil*. 8:93-99.

Inman-Bamber, N.G., 1991b. Some physiological factors affecting the optimum age and season for harvesting sugarcane. *Proceedings of the South African Sugar Technologists' Association*. 65:103-108.

Inman-Bamber, N.G., 1994. Temperature and seasonal effects on canopy development and light interception of sugarcane. *Field Crops Research*. 36:41-51.

Inman-Bamber, N.G., 1995. Climate and water as constraints to production in the South African Sugar Industry. *Proceedings of the South African Sugar Technologists' Association*. 69:55-59.

Inman-Bamber, N.G., Culverwell, T.L. and McGlinchey, M.G., 1993. Predicting yield responses to irrigation of sugarcane from a growth model and field records. *Proceedings of the South African Sugar Technologists' Association*. 67:66-72.

Inman-Bamber, N.G. and Kiker, G.A., 1997. Personal communication. South African Sugar Association Experiment Station, Mount Edgecombe and Department of Agricultural Engineering, University of Natal, Pietermaritzburg, RSA.

Inman-Bamber, N.G., Singels, A. and Muchow, R.C., 1998. A systems approach to benchmarking for sugarcane production in Australia and South Africa. *Proceedings of the South African Sugar Technologists' Association*. 71:117-122.

Inman-Bamber, N.G. and Thompson, G.D., 1989. Models of dry matter accumulation by sugarcane. *Proceedings of the South African Sugar Technologists' Association*. 63:212-216.

Jackson, R.D., Idso, S.B., Reginato, R.J. and Pinter, P.J. Jr., 1981. Canopy temperature as a crop water-stress indicator. *Water Resources Research*. 17:1133-1138.

Jaggard, K.W. and Clark, C.J.A., 1990. Remote sensing to predict the yield of sugar beet in England. In *Applications of Remote Sensing in Agriculture*. Ed. M.D. Steven and J.A. Clark. University Press, Cambridge, UK. 201-206.

Jones, CA and Kiniry, J.R., 1986. *CERES - Maize: A Simulation Model of Maize Growth and Development*. Texas A & M University Press. pp 194.

Kalma, J.D., Lyons, T.J., Nunez, M. and Pitman, A.J., 1991. Advances in understanding meso-scale climatic variability In *Climatic Risk in Crop Production: Models and Management for the Semiarid Tropics and Subtropics*. Ed. R.C. Muchow and J.A. Bellamy. CAB International, Wallingford, UK. 445-466.

King, C. and Meyer-Roux, J., 1990. Remote sensing in agriculture: from research to applications. In *Applications of Remote Sensing in Agriculture*. Ed. M.D. Steven and J.A. Clark. University Press, Cambridge, UK. 377-395.

Kuhnel, I., 1993. Impacts of extreme climatic conditions on sugar cane production in northeastern Australia. In *Extreme Hydrological Events: precipitation, floods and droughts*. Ed. I. Kuhnel and Z.W. Kundzewicz. Proceedings of International Symposium, Yokohama, Japan. 20-23 July, 1993. IAHS Publ. No. 213. 157-163.

Landman, W.A., 1995. *Bulletin of the Research Group for Statistical Climate Studies*, South African Weather Bureau, Pretoria, RSA. Vol 3 No. 3. pp 32.

Landman, W.A. and Bartman, A.G., 1998. Personal communication. Research Group for Seasonal Climate Studies, South African Weather Bureau, Pretoria, RSA.

Lecler, N.L., Schulze, R.E. and Pike, A., 1996. Initial assessment of methodologies to generate agricultural and hydrological forecasts from seasonal categorical rainfall forecasts in South Africa. In *Workshop on Reducing Climate-Related Vulnerability in Southern Africa.* Victoria Falls, Zimbabwe. Published by SADC, NOAA and NASA. 113-130.

Linacre, E.T., 1991. Unpublished manuscript. School of Earth Sciences, Macquarie University, Sydney, Australia.

Lindesay, J.A., Harrison, M.S.J. and Haffner, M.P., 1986. The Southern Oscillation and South African rainfall. *South African Journal of Science*. 82:196-198.

Lorber, M.N., Fluck, R.C. and Mishoe, J.W., 1984. A method for analysis of sugarcane (*Saccharum sp.*) biomass production systems. *Transactions of the American Society of Agricultural Engineers.* 27:146-158.

Lourens, U.W., 1995. *A system for drought monitoring and severity assessment.* Unpublished PhD Thesis. Department of Agrometeorology, University of the Orange Free State, Bloemfontein, RSA.

Lumsden, T.G., Lecler, N.L. and Schulze, R.E., 1998. Simulation of sugarcane yield at the scale of a mill supply area. *Proceedings of the South African Sugar Technologists' Association.* 72:12-17.

Lumsden, T.G., Schulze, R.E., Lecler, N.L. and Schmidt, E.J., 1999. *An Assessment of the Potential for Sugarcane Yield Forecasting using Seasonal Rainfall Forecasts and Crop Yield Models. A Case Study for the Eston Mill Supply Area.* Report to SA Sugar Association Experiment Station. School of Bioresources Engineering and Environmental Hydrology, University of Natal, Pietermaritzburg, RSA. pp 135.

MacDonald R.B. and Hall, F.G., 1980. Global crop forecasting. *Science.* 208:670-679.

Macvicar, C.N., De Villiers, J.M., Loxton, R.F., Verster, E., Lambrechts, J.J.N., Merryweather, F.R., Le Roux, J., Van Rooyen, T.H. and Harmse, H.J. von M., 1977. *Soil Classification - A Binomial System for South Africa.* Department of Agricultural Technical Services, Soil and Irrigation Research Institute, Pretoria, RSA. pp 150.

Mann, Q., Meyer, J. and Hellmann, D.B., 1997. Personal communication. South African Sugar Association Experiment Station, Mount Edgecombe, RSA.

Martin, R., Washington, R. and Downing, T.E., 2000. Seasonal maize forecasting for South Africa and Zimbabwe derived from an agroclimatological model. *Journal of Applied Meteorology* (in press).

Maselli, F., Conese, C., Petkov, L. and Gilabert, M.A., 1993. Environmental monitoring and crop forecasting in the Sahel through the use of NOAA NDVI data. A case study: Niger 1986-89. *International Journal of Remote Sensing.* 14:3471-3487.

Mason, S.J., 1996. Personal communication. Climatology Research Group, University of Witwatersrand, Johannesburg, RSA.

McBride, J.L. and Nicholls, N.N., 1983. Seasonal relationships between Australian rainfall and the Southern Oscillation. *Monthly Weather Review.* 111:1998-2004.

McCree, K.J., 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, The Netherlands.

McGlinchey, M.G. and Inman-Bamber, N.G., 1996. Effect of irrigation scheduling on water use efficiency and yield. *Proceedings of the South African Sugar Technologists' Association.* 70:55-56.

McGlinchey, M.G., Inman-Bamber, N.G., Culverwell, T.L. and Els, M., 1995. An irrigation scheduling method based on a crop model and an automatic weather station. *Proceedings of the South African Sugar Technologists' Association.* 69:69-73.

Meyer, S.J., Hubbard, K.G. and Wilhite, D.A., 1993. A crop-specific drought index for corn: II. Application in drought monitoring and assessment. *Agronomy Journal.* 85:396-399.

Monteith, J.L., 1965. Evaporation and the environment. In *The State and Movement of Water in Living Organisms.* Ed. A.E. Frogg. Academic Press, New York, USA. 205-234.

Monteith, J.L. and Unsworth, M.H., 1990. *Principles of Environmental Physics.* 2nd Edition. Ed. E. Arnold. Hodder and Stoughton, London, UK. pp 291.

Motha, R.P. and Heddinghaus, T.R., 1986. The joint agricultural weather facility's operational assessment program. *Bulletin of the American Meteorological Society.* 67:1114-1122.

Pandey, P.C., Dadhwal, V.K., Sahai, B. and Kale, P.P., 1992. An optimal estimation technique for increasing the accuracy of crop forecasts by combining remotely sensed and conventional forecast results. *International Journal of Remote Sensing.* 13:2735-2741.

Pasov, V.M. and Yatsalo, B.L., 1992. Application of a complex of models to agrometeorological forecasting. *Soviet Meteorology and Hydrology.* 12:71-76.

Phillips, J.G., Rosenzweig, C. and Cane, M., 1996. Exploring the potential for using ENSO forecasts in the U.S. corn belt. In *Drought Network News.* Volume 8. A Newsletter of the International Drought Information Center and the National Drought Mitigation Center. Published in cooperation with the World Climate Program of the WMO. 6-10.

Pike, A. and Schulze, R.E., 1995. AUTOSOIL Version 3: A Soils Decision Support System for South African Soils. Department of Agricultural Engineering, University of Natal, Pietermaritzburg, RSA.

Popov, G.F., 1990. Agrometeorological crop monitoring and yield forecasting. In *Using Meteorological Information and Products.* Ed. G.F. Popov and A. Price-Budgen. 427-435.

Rimmington, G.M. and Nicholls, N.N., 1993. Forecasting wheat yields in Australia with the Southern Oscillation Index. *Australian Journal of Agricultural Research.* 44:625-632.

Ritchie, J.T., 1983. Reduction of stress related to soil water deficit. In *Crop Reactions to Water and Temperature Stresses in Humid Temperate Climates.* Ed. C.D. Raper and P.J. Kramer. Westview Press, Boulder, Colorado, USA. 329-340.

Ritchie, J.T., 1985. The CERES maize model. In *CERES Maize: A Simulation Model of Maize Growth and Development. Users guide.* Ed. C.A. Jones. Blackland Research Center, Texas A and M University, Temple, Texas, USA.

Sakamoto, C.M., 1989. Tools for an operational drought impact assessment system. In *Workshop on Drought*. Ed. A.L. Du Pisani. Southern African Regional Commission for the Conservation and Utilization of the Soil, Pretoria, RSA. 12-22.

Schmidt, E.J., 1998. Forecasting the sugarcane crop. In *The Link*, September 1998. South African Sugar Association Experiment Station, Mount Edgecombe, RSA. 7:11.

Schulze, R.E., 1995. *Hydrology and Agrohydrology: A Text to Accompany the ACRU 3.00 Agrohydrological Modelling System*. Water Research Commission, Pretoria, RSA. Report TT69/95. pp 552.

Schulze, R.E., 1997. *South African Atlas of Agrohydrology and -Climatology*. Water Research Commission, Pretoria, RSA. Report TT82/96. pp 276.

Schulze, R.E., Domleo, F.B., Furniss, P.W. and Lecler, N.L., 1995. Crop Yield Estimation. In *Hydrology and Agrohydrology: A Text to Accompany the ACRU 3.00 Agrohydrologcal Modelling System*. Ed. R.E. Schulze. Water Research Commission, Pretoria, RSA. Report TT69/95. AT19-1 to AT19-14.

Schulze, R.E., Hallowes, J., Lynch, S.D., Perks, L.A. and Horan, M., 1998. Forecasting seasonal streamflow in South Africa: A preliminary investigation In *Forecasting Rainfall and Streamflow over South and Southern Africa*. Ed. A.M. Joubert. ESKOM Technology Group, Rocheville, RSA. Research Report No. TRR/T98/046.

Schulze, R.E., Hutson J.L. and Cass, A., 1985. Hydrological characteristics and properties of soils in southern Africa 2 : Soil water retention models. *Water* SA. 11:129-136.

Schulze, R.E. and Kunz, R.P., 1995. Reference Potential Evaporation. In *Hydrology and Agrohydrology: A Text to Accompany the ACRU 3.00 Agrohydrologcal Modelling System*. Ed. R.E. Schulze. Water Research Commission, Pretoria, RSA. Report TT69/95. AT4-1 to AT4-38.

Schulze, R.E. and Maharaj, M., 1998. Personal communication. Department of Agricultural Engineering, University of Natal, Pietermaritzburg, RSA.

SCWG, 1991. *Soil Classification - Taxonomic System for South Africa*. Soil Classification Working Group, Department of Agricultural Development, Soil and Irrigation Research Institute, Pretoria, RSA. pp 257.

Singels, A., Kennedy, A.J. and Bezuidenhout, C.N., 1998. Irricane: A simple irrigation scheduling procedure for sugarcane using automatic weather stations. *Proceedings of the South African Sugar Technologists' Association*. 72:117-122.

Singels, A., Kennedy, A.J. and Bezuidenhout, C.N., 1999. Weather based decision support through the internet for agronomic management of sugarcane. *Proceedings of the South African Sugar Technologists' Association.* 73:30-32.

Singh, B.H. and Bapat, S.R., 1988. Pre-harvest forecast models for prediction of sugarcane yield. *Indian Journal of Agricultural Sciences.* 58:465-469.

Smith, R.C., Adams, J., Stephens, D.J. and Hick P.T., 1995. Forecasting wheat yield in a Mediterranean-type environment. *Australian Journal of Agricultural Research.* 46:113-125.

Smithers, J.C. and Schulze, R.E., 1995. *ACRU Agrohydrologcal Modelling System : User Manual Version 3.00.* Water Research Commission, Pretoria, RSA. Report TT70/95. pp 368.

Sridhar, V.N., Dadhwal, V.K., Chaudhari, K.N., Sharma, R., Bairagi, G.D. and Sharma, A.K., 1994. Wheat production forecasting for a predominantly unirrigated region in Madhya Pradesh (India). *International Journal of Remote Sensing.* 15:1307-1316.

Stephens, D.J., Lyons, T.J. and Lamond, M.H., 1989. A simple model to forecast wheat yield in Western Australia. *Journal of the Royal Society of Western Australia.* 71:77-81.

Stephens, D.J., Walker, G.K. and Lyons, T.J., 1994. Forecasting Australian wheat yields with a weighted rainfall index. *Agricultural and Forest Meteorology.* 71:247-263.

Stone, R. and Auliciems, A., 1992. SOI phase relationships with rainfall in eastern Australia. *International Journal of Climatology.* 12:625-636.

Thompson, G.D., 1976. Water use by sugarcane. *South African Sugar Journal.* 60:593-600 and 627-635.

Thompson, G.D., 1986. *Agrometeorological and Crop Measurements in a Field of Irrigated Sugarcane.* Mount Edgecombe Research Report No. 5. South African Sugar Association Experiment Station, RSA.

Tsuji, G.Y., Jones, J.W., Hoogenboom, G., Hunt, L.A. and Thornton, P.K., 1994. Introduction In *DSSAT Version 3.* Ed. G.Y. Tsuji, G. Vehara and S. Balas. International Benchmark Sites Network for Agrotechnology Transfer, University of Hawaii, Hawaii, USA. Vol. 1 pp 164.

Uhrig, J.W., Engel, B.A. and Baker, W.L., 1992. An application of neural networks: Predicting corn yields. In *Computers in Agricultural Extension Programs.* Ed. D.G. Watson, F.S Zazueta and A.B. Bottcher. American Society of Agricultural Engineers, St. Joseph, Michegan, USA. 348-353.

Van Antwerpen, R., McGlinchey, M.G., Inman-Bamber, N.G. and Bennie, A.T.P., 1996.

Estimating root water use in sugarcane. *Proceedings of the South African Sugar Technologists' Association.* 70:57-58.

Van Antwerpen, R., Meyer, J.H. and Inman-Bamber, N.G., 1993. Cane root development and crop modelling. *Proceedings of the South African Sugar Technologists' Association* 67: 73-77.

Van Dillewijn, C., 1952. *Botany of Sugarcane.* The Chronica Botanica Co., Waltham, Mass., USA.

Van Heerden, J., Terblanche, D.E. and Schulze, G.C., 1988. The Southern Oscillation and South African summer rainfall. *Journal of Climatology.* 8:577-597.

Walker, G.K, 1989. Model for operational forecasting of western Canada wheat yield. *Agricultural and Forest Meteorology.* 44:339-351.

Wiegand, C.L. and Richardson, A.J., 1990. Use of spectral vegetation indices to infer leaf area, evapotranspiration and yield: I. Rationale. *Agronomy Journal.* 82:623-629.

Wilkens, P.W., Thornton, P.K. and Bowen, W.T., 1994. Real-time yield forecasting in Burkina Faso. In *Applying Crop Models and Decision Support Systems.* Ed. P.K. Thornton, W.T. Bowen, P.W. Wilkens, J.W. Jones, K.J. Boote and G. Hoogenboom. University of Florida, Gainesville, Florida, USA. 56-57.

Willmott, C.J., 1981. On the validation of models. *Physical Geography.* 2:184-194.

Wisiol, K., 1987. Choosing a basis for yield forecasts and estimates. In *Plant Growth Modelling for Resource Management Volume 1.* Ed. K. Wisiol and J.D. Hesketh. CRC Press, Boca Raton, Florida, USA. 75-103.

APPENDIX A: SPATIAL TEMPERATURE ESTIMATION TECHNIQUE

The spatial temperature estimation technique (Schulze and Maharaj, 1998) involves consideration of the proximity and relative altitude of surrounding climate stations available for the estimation of daily temperatures at a point of interest, for example a farm. Weightings related to proximity and relative altitude are assigned to the available stations with the sum of these weightings being used to form a ranking of station suitability. The assignment of weightings is shown in Figure Aa) and b) for proximity and altitude related weightings respectively. The proximity of stations is expressed in terms of distance apart in geographic map co-ordinates. Stations that are closer in proximity and at a more similar altitude to the point of interest, are given a higher weighting and thus greater suitability for estimation of temperatures.

The observed data from the most suitable station are extracted to develop the temperature data set at the point of interest, with any corrections for altitudinal differences being made according to regional temperature lapse rates (Schulze and Kunz, 1995). Checks are performed on the estimated temperatures to ensure a minimum range of 1.5 °C between maximum and minimum temperatures. If the range is less than 1.5 °C the following corrections are applied:

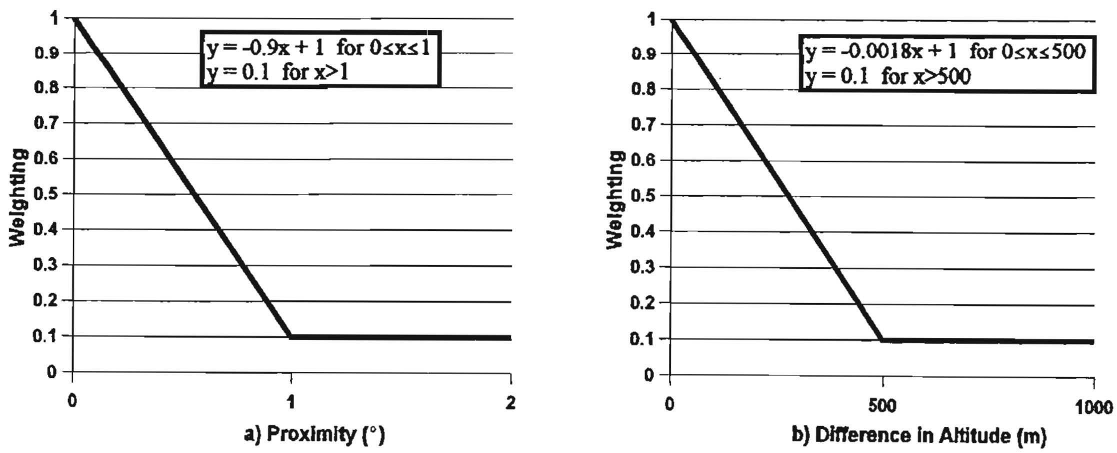


Figure A Weightings according to a) proximity and b) difference in altitude, assigned to climate stations surrounding a point of interest (Schulze and Maharaj, 1998)

$$Tmax_{new} = Tave + (0.5*1.5) + 0.05$$
$$Tmin_{new} = Tave - (0.5*1.5) + 0.05$$

$$\text{where} \quad Tave = (Tmax_{old} + Tmin_{old}) / 2$$

For any days of missing temperature data in the station record, data from the next most suitable station are extracted and corrected, if necessary, in order to infill the record. If necessary, the process of data extraction is continued successively through all stations, until observed data are found to represent temperatures on that day. If, for a particular day, no observed data can be found at any station, daily temperatures are derived from a harmonic analysis of the relevant monthly long term means of daily maximum and minimum temperatures of the most suitable station.

APPENDIX B: WORKING RULES USED IN THE PREDICTION OF SOIL TYPE AND DEPTH BASED ON SOIL PARENT MATERIAL AND OTHER INFORMATION (MANN, MEYER AND HELLMANN, 1997)

Parent Material	Slope Position	Slope Gradient (%)	MAP (mm)	Soil Type	Topsoil Depth (m)	Effective Subsoil Depth (m)
TMS (mist belt)	Topslope	5 - 12	< 800	Gs 14	0.4	0.5
			800 - 900	No 10 (60%) / Ia 11 (40%)	0.5 / 0.5	0.3 / 0.8
			> 900	Ia 11 (70%) / No 10 (30%)	0.5 / 0.5	0.8 / 0.3
		< 5	< 800	Gs 14	0.4	0.5
			800 - 900	Ia 11	0.5	1.0
			> 900	Ia 11 (60%) / Kp 11 (40%)	0.5 / 0.5	1.0 / 1.0
	Midslope	> 12	< 800	Gs 14	0.4	0.5
			800 - 900	Ia 11 (70%) / No 10 (30%)	0.5 / 0.5	0.8 / 0.3
			> 900	Ia 11	0.5	1.0
		5 - 12	< 800	Gs 14	0.4	0.5
			800 - 900	Kp 11	0.5	1.0
			> 900	Kp 11	0.5	1.0
		< 5	< 800	Gs 14	0.4	0.5
			800 - 900	Kp 11	0.5	1.0
			> 900	Kp 11	0.5	1.0
	Bottomslope	5 - 12	< 800	Gs 14	0.4	0.5
			800 - 900	Ma 11	0.5	1.0
			> 900	Ma 11	0.5	1.0
		< 5	< 800	Gs 14	0.4	0.5
			800 - 900	Du 10 (30% clay)	0.3	0.8
			> 900	Du 10 (30% clay)	0.3	0.8

Parent Material	Slope Position	Slope Gradient (%)	MAP (mm)	Soil Type	Topsoil Depth (m)	Effective Subsoil Depth (m)
DWYKA TILLITE	Topslope	5 - 12	< 800	Gs 16	0.3	0.2
			800 - 900	Gs 16	0.3	0.2
			> 900	Ia 11 (60%) / Gs 19 (40%)	0.5 / 0.3	0.5 / 0.3
		< 5	< 800	We 12	0.3	0.2
			800 - 900	We 13 (60%) / Gs 19 (40%)	0.4 / 0.3	0.2 / 0.3
			> 900	Ia 10 (40%) / Gs 19 (60%)	0.3 / 0.3	0.3 / 0.3
	Midslope	> 12	< 800	We 12	0.3	0.2
			800 - 900	We 12	0.3	0.2
			> 900	Gs 16	0.4	0.2
		5 - 12	< 800	Lo 12	0.3	0.3
			800 - 900	Lo 12	0.3	0.3
			> 900	Lo 12	0.3	0.3
		< 5	< 800	Lo 12	0.4	0.3
			800 - 900	Lo 12	0.4	0.3
			> 900	Lo 13	0.4	0.3
	Bottomslope	5 - 12	< 800	Kd 16	0.3	0.2
			800 - 900	Kd 16	0.3	0.2
			> 900	Kd 19	0.3	0.2
		< 5	< 800	Ka 20 (10 - 15% clay)	0.4	0.0
			800 - 900	Ka 20 (10 - 30% clay)	0.4	0.0
			> 900	Ka 20 (10 - 30% clay)	0.4	0.0

Parent Material	Slope Position	Slope Gradient (%)	MAP (mm)	Soil Type	Topsoil Depth (m)	Effective Subsoil Depth (m)
LOWER ECCA SHALE	Topslope	5 - 12	< 800	Ms 10 (50%) / Mw 11(50%)	0.3 / 0.4	0.0 / 0.0
			800 - 900	Gs 19	0.3	0.2
			> 900	Gs 19 (60%) / Cv 17 (40%)	0.3 / 0.3	0.3 / 0.5
		< 5	< 800	My 11 (60%) / Mw 11 (40%)	0.3 / 0.4	0.2 / 0.0
			800 - 900	Gs 19	0.3	0.3
			> 900	Gs 19 (60%) / Cv 17 (40%)	0.3 / 0.3	0.3 / 0.5
	Midslope	> 12	< 800	My 11 (60%) / Mw 11 (40%)	0.3 / 0.4	0.2 / 0.0
			800 - 900	Gs 19	0.3	0.3
			> 900	Gs 19	0.3	0.4
		5 - 12	< 800	My 11	0.3	0.2
			800 - 900	Gs 19	0.3	0.3
			> 900	Gs 19 (60%) / Cv 17 (40%)	0.3 / 0.3	0.4 / 0.5
		< 5	< 800	My 11	0.3	0.2
			800 - 900	Gs 19	0.3	0.2
			> 900	Cv 17	0.3	0.5
	Bottomslope	5 - 12	< 800	Wo 11	0.4	0.0
			800 - 900	Wo 11 (60%) / Bo 41 (40%)	0.5 / 0.3	0.0 / 0.4
			> 900	Wo 11	0.6	0.0
		< 5	< 800	Wo 11 (40%) / Rg 20 (60%)	0.4 / 0.4	0.0 / 0.0
			800 - 900	Wo 11 (40%) / Rg 20 (60%)	0.5 / 0.5	0.0 / 0.0
			> 900	Wo 11	0.6	0.0

Parent Material	Slope Position	Slope Gradient (%)	MAP (mm)	Soil Type	Topsoil Depth (m)	Effective Subsoil Depth (m)
DOLERITE and BASALT	Topslope	5 - 12	< 800	Sd 21	0.3	0.8
			800 - 900	Sd 22 (60%) / Hu 27 (40%)	0.3 / 0.3	0.8 / 1.0
			> 900	Ia 12 (50%) / Hu 18 (50%)	0.5 / 0.3	1.0 / 1.0
		< 5	< 800	Sd 21	0.3	0.8
			800 - 900	Sd 22 (60%) / Hu 27 (40%)	0.4 / 0.4	0.8 / 1.1
			> 900	Ia 12 (50%) / Hu 18 (50%)	0.5 / 0.4	1.0 / 1.1
	Midslope	> 12	< 800	Sd 21	0.3	0.8
			800 - 900	Sd 22 (60%) / Hu 27 (40%)	0.3 / 0.3	0.8 / 1.0
			> 900	Ia 12 (50%) / Hu 18 (50%)	0.5 / 0.3	1.0 / 1.2
		5 - 12	< 800	Sd 21	0.3	0.8
			800 - 900	Sd 22 (60%) / Hu 27 (40%)	0.3 / 0.3	0.8 / 1.0
			> 900	Ia 12 (50%) / Hu 18 (50%)	0.5 / 0.4	1.0 / 1.1
		< 5	< 800	Sd 21 (60%) / Bo 41 (40%)	0.3 / 0.3	0.8 / 0.4
			800 - 900	Sd 22 (60%) / Hu 27 (40%)	0.4 / 0.4	0.8 / 1.1
			> 900	Ia 12 (50%) / Hu 18 (50%)	0.5 / 0.4	1.0 / 1.1
	Bottomslope	5 - 12	< 800	Rg 20	0.5	0.0
			800 - 900	Rg 20	0.5	0.0
			> 900	Rg 20	0.5	0.0
		< 5	< 800	Rg 20	0.5	0.0
			800 - 900	Rg 20	0.5	0.0
			> 900	Rg 20	0.5	0.0

APPENDIX C: DETAILED ESTIMATED COSTS OF IMPLEMENTATION OF *ACRU*-THOMPSON AND SIMPLE RAINFALL MODEL BASED YIELD FORECASTING SYSTEMS

ACRU-Thompson

Initial Set-up

Task	Required Time (h)	Cost (R)
1) Obtain GIS coverage of farm boundaries within the MSA. Divide MSA into several sub-areas according to MAP.	12	1800
2) Assign driver climate stations to each sub-area as well as adjustment factors to ensure station data are representative of the sub-areas.	18	2700
3) Estimate an average soil TAM for each sub-area (in consultation with soils specialists).	4	600
4) Determine the dominant growth cycles in the MSA and the average proportion of each cycle harvested in a year.	60	9000
5) Obtain historical climate and cane yield data for past seasons. Perform data checks. Infill missing climate data. Simulate historical cane yields and verify against observed yield data.	250	37500
6) Selection of appropriate analogue years (corresponding to 15th, 50th and 85th percentile rainfall) for each category of rainfall possible (above, near and below normal rainfall) for all periods in a season where translation of categorical rainfall forecasts into daily rainfall values is required. The selection of appropriate analogue years (corresponding to 50th percentile rainfall) for the remaining periods leading up to harvest.	150	22500
7) Development of the capability to update the various climate records in near real-time (daily updates via communication links to automatic weather stations).	Assume already available	-
Total	494	74100

Monthly System Maintenance

Task	**Required Time per Month (h)**	**Cost (R)**
1) Checking and infilling of missing values, if necessary, of climate data from automatic weather stations	15	2250
2) Obtain seasonal rainfall forecasts for upcoming six month period	2	300
3) Generation of yield forecasts	12	1800
4) Present yield forecasts in graph or table form for the MSA (aggregated results) and for sub-areas of the MSA (grower estimation).	12	1800
Total	41	6150

Annual System Maintenance

Task	**Required Time per Year (h)**	**Cost (R)**
1) Post-season analysis of yield forecast accuracy to ensure optimal performance of yield forecasting system.	36	5400
Total	36	5400

Simple Rainfall Model

Initial Set-up

Task	Required Time (h)	Cost (R)
1) Assign a driver rainfall station to the MSA as well an adjustment factor to ensure station data are representative of the MSA.	3	450
2) Determine the average growth cycle length in the MSA	8	1200
3) Obtain historical rainfall and cane yield data for past seasons. Perform data checks. Infill missing rainfall data. Determine the MAP and average MSA cane yield of recent years (for example, the past 10 years). Calculate the mean rate of yield accumulation (t/ha/100mm). Predict historical yields and verify against observed yield data.	70	12600
4) Selection of appropriate analogue years (corresponding to 50th percentile rainfall) for each category of rainfall possible (above, near and below normal rainfall) for all periods in a season where translation of categorical rainfall forecasts into monthly rainfall values is required. The selection of appropriate analogue years (corresponding to 50th percentile rainfall) for the remaining rainfall periods leading up to harvest. Determination in both cases of the relevant 50th percentile rainfall totals.	24	3600
Total	105	15750

Monthly System Maintenance

Task	Required Time per Month (h)	Cost (R)
1) Checking and infilling of missing values, if necessary, of newly acquired rainfall data	3	450
2) Obtain seasonal rainfall forecasts for upcoming six month period	2	300
3) Generation of yield forecasts	1.5	225
4) Present yield forecasts in graph or table form for the MSA.	1.5	225
Total	8	1200

Annual System Maintenance

Task	**Required Time per Year (h)**	**Cost (R)**
1) Post-season analysis of yield forecast accuracy to ensure optimal performance of yield forecasting system.	5	750
Total	5	750