DEVELOPMENT AND EVALUATION OF MODEL-BASED OPERATIONAL YIELD FORECASTS IN THE SOUTH AFRICAN SUGAR INDUSTRY

CAREL NICOLAAS BEZUIDENHOUT

Submitted in fulfilment of the academic requirements for the degree of Ph.D. in the Discipline of Hydrology

School of Bioresources Engineering and Environmental Hydrology
University of KwaZulu-Natal
Pietermaritzburg

2005

Abstract

South Africa is the largest producer of sugar in Africa and one of the ten largest sugarcane producers in the world. Sugarcane in South Africa is grown under a wide range of agro-climatic conditions. Climate has been identified as the single most important factor influencing sugarcane production in South Africa. Traditionally, sugarcane mill committees have issued forecasts of anticipated production for a region. However, owing to several limitations of such committee forecasts, more advanced technologies have had to be considered. The aim of this study has been to develop, evaluate and implement a pertinent and technologically advanced operational sugarcane yield forecasting system for South Africa. Specific objectives have included literature and technology reviews, surveys of stakeholder requirements, the development and evaluation of a forecasting system and the assessment of information transfer and user adoption. A crop yield model-based system has been developed to simulate representative crops for derived Homogeneous Climate Zones (HCZ). The system has integrated climate data and crop management, soil, irrigation and seasonal rainfall outlook information. Simulations of yields were aggregated from HCZs to mill supply area and industry scales and were compared with actual production. The value of climate information (including climate station networks) and seasonal rainfall outlook information were quantified independently. It was concluded that the system was capable of forecasting yields with acceptable accuracy over a wide range of agro-climatic conditions in South Africa. At an industry scale, the system captured up to 58% of the climatically driven variability in mean annual sugarcane yields. Forecast accuracies differed widely between different mill supply areas, and several factors were identified that may explain some inconsistencies. Seasonal rainfall outlook information generally enhanced forecasts of sugarcane production. Rainfall outlooks issued during the summer months seemed more valuable than those issued in early spring. Operationally, model-based forecasts can be expected to be valuable prior to the commencement of the milling season in April. Current limitations of forecasts include system calibration, the expression of production relative to that of the previous season and the omission of incorporating near real-time production and climate information. Several refinements to the forecast system are proposed and a strong collaborative approach between modellers, climatologists, mill committees and other decision makers is encouraged.

The author hereby declares that, unless where it has been acknowledged, the research results reported in this thesis are original and were conducted by the author in a personal capacity.

____
Carel Nicolaas Bezuidenhout

____
Supervisor: Prof. R.E. Schulze

Date: ___________________

University of KwaZulu-Natal, Pietermaritzburg, South Africa

Table of Contents

Acknowledgements

The author would like to thank his project supervisor, Dr. Abraham Singels from the South African Sugarcane Research Institute (SASRI) and his academic supervisor, Prof. Roland Schulze of the University of KwaZulu-Natal (UKZN), for their committed guidance and assistance during the course of this study. In addition, the following persons are thanked for their time and comments during research steering committee meetings; Mr. Erik Schmidt (formerly of SASRI), Mr. Eric Hulbert (formerly of SASRI), Dr. Kathy Hurley (SASRI) and Dr. Greg Kiker (formerly of UKZN). Direct and indirect funding towards this project was made available by SASRI and UKZN. Mr. Rob Donaldson and Mr. Doug Hellman from SASRI are thanked for their valuable support and comments during the course of the study. The Agricultural Engineering and Biometry Departments at SASRI are acknowledged for capturing and processing the climate data used in this study. Additional climate data were made available through Water Research Commission funded projects at the School of Bioresources Engineering & Environmental Hydrology at UKZN. Mr. Steven Lynch and Mrs. Manju Maharaj from UKZN kindly assisted in the provision of these datasets. The following persons provided valuable agronomic and milling information of sugarcane regions and grower groups: F. Botha, B. Swart, S. v.d. Walt, T. Culverwell, A. Jones, R. Fortmann, P. Brenchley, G. Lagerwall, R. Stranack, E. Hulbert, G. Maher, D. Hellmann, D. McElligott, G. Buchanan, M. Eweg, D. McArthur, A. Simpson, W. Gillham, J de Lange and G. Groom. Mr. Q. Hildebrand, Mr. D. Rossler and Dr. D. Nixon from the South African Sugar Association and Mr. J. de Lange from the Illovo Sugar Ltd milling company provided information on the potential value of yield and production forecasts to the sugar industry. Mr. D. Whitfield (Industry Liason Manager, South African Sugar Association) provided a history of Mill Group Board forecasts, as well as estimates of areas under cane in certain mill regions. Dr. A. Wynne (S.A. Canegrowers Association) supplied a wide range of historical industry production data since 1980. Mr. T. Lumsden (UKZN) released data from his MSc study for re-analysis. The Department of Water Affairs and Forestry, Impala Water Board and Transvaal Suiker Beperk supplied river flow and water restriction information for the Pongola and Mpumalanga regions. The author would like to acknowledge Mr. C. Gers (SASRI) for valuable comments, GIS assistance and assisting with the derivation of Homogeneous Climate Zones. Mr. J.

Lambrecht (Transvaal Suiker Bpk.) assisted with GIS mapping and spatial data for the Mpumalanga region. Mr. N. Govender (SASRI) assisted with some of the statistical approaches followed in the study. The South African Weather Service, in particular Mrs. Anna Bartman, is thanked for the provision an archive of historical seasonal climate outlooks. Ms. S. Bunge (SASRI) carried out the operational yield forecasts that were used for the 2003 milling season. Bronwyn Bezuidenhout assisted with the preparation and proof reading of this thesis.

List of Symbols and Abbreviations

ϕ	Latitude (°N)
ACRU	Agricultural Catchments Research Unit
AMD	Absolute mean deviation
AVHRR	Advanced Very High Resolution Radiometer
AWC	Plant available water holding capacity (mm)
AWS	Automatic weather station
BEEH	School of Bioresources Engineering and Environmental Hydrology
CCD	Cold Cloud Duration
CGMS	Crop Growth Simulation System
CSAG	Climate Systems Analyses Group (University of Cape Town)
$\overline{CV}$	Mean coefficient of spatial variation (%) in one homogeneous climate zone. $\overline{CV} = \left(CV_{R_S} + CV_{MAP} + CV_{HU}\right)/3$
CVAP	The Climate Variability in Agriculture R&D Program
DSP	Decision support program
DSSAT	Decision Support System for Agrotechnology Transfer
$\overline{D_X}$	The mean relative discrimination of Zone X
$\bar{\partial}_X$	The mean distance between normalised data points in Zone X
$D_{X,Y}$	The relative discrimination between Zone X and Zone Y
$\bar{\partial}_{X,Y}$	The mean distance between normalised data points in Zone X and Zone Y combined
ECMWF	European Centre for Medium-Range Weather Forecasts
EC_{ref}	Daily potential cane reference evapotranspiration ($mm.d^{-1}$)

ENSO	El Niño-Southern Oscillation
ET_O	Daily reference evapotranspiration over short grass ($mm.d^{-1}$)
FAO	Food and Agriculture Organization of the United Nations
GCM	General Circulation Model
GIS	Geographical Information System
H&S	Hargreaves and Samani (1985) Equation
HCZ	Homogeneous Climate Zone
HU	Long-term mean thermal time (base 10°C, $°C.d.an^{-1}$)
$\overline{HU}$	Mean HU within one HCZ
IRI	International Research Institute for Climate Prediction
LANDSAT	Land Satellite
MAP	Long-term Mean Annual Precipitation ($mm.an^{-1}$)
$\overline{MAP}$	Mean of MAP within one HCZ
MGB	Mill Group Board
NOAA	National Oceanic and Atmospheric Administration
PAR	Photosynthetic active radiation ($MJ.m^{-2}.d^{-1}$)
QC	Quaternary Catchment
R_a	Extraterrestrial solar radiation ($MJ.m^{-2}.d^{-1}$)
$RMSE$	Root Mean Square Error
R_S	Long-term mean annual solar radiation ($MJ.m^{-2}.an^{-1}$)
$\overline{R_S}$	Mean R_S within one HCZ
SASA	South African Sugar Association
SASRI	South African Sugarcane Research Institute
SAWS	South African Weather Service

SMD	Squared Mean Difference
SOI	Southern Oscillation Index
SPOT	Systeme Pour l'Observation de la Terre
SST	Sea surface temperature
T_{dew}	Dew point temperature (°C)
T_{max}	Daily maximum temperature (°C)
T_{mean}	Mean daily temperature (°C)
T_{min}	Daily minimum temperature (°C)
T_{ra}	The range between long-term mean air temperature of the hottest and coldest months in the year (°C)
u	Mean daily wind speed at 2 m ($m.s^{-1}$)
VR	Variance ratio
z	Altitude above mean sea level (m)

1 Introduction

1.1 Problem Statement in Summary

South Africa is the largest producer of sugar in Africa and is one of the ten largest sugarcane producers in the world. Sugarcane is an economically important crop and is grown under a wide range of agroclimatic conditions in the eastern regions of the country. These regions are subject to high climatic variability, resulting in large fluctuations in annual production. Sugar industry stakeholders, who include the growers, millers and marketers, are therefore compelled to engage in decision making under high levels of climatic uncertainty. This, at times, evokes sub-optimal risk alleviating management approaches, which may significantly reduce profitability. Current crop forecasts are insufficiently accurate and neither tailored to assist with the wide range of decisions stakeholders need to make. A new forecasting system based on the most appropriate technology and sensitive to the requirements of industry stakeholders therefore had to be developed.

1.2 Problem Statement

1.2.1 Background to the South African Sugar Industry

Sugarcane (*Saccharum* species hybrids) flourishes under a long and warm summer growing season with a high incidence of radiation and with adequate soil moisture. This needs to be followed by a dry, sunny and fairly cool, but frost-free, winter to promote high sucrose quantities before harvesting (Smith, 1992). Although relatively far south of the equator, some areas in eastern South Africa are well suited for commercial sugarcane production. In these areas sugarcane is produced under a wide range of climatic, agronomic and socio-economic conditions.

The South African sugar industry comprises of approximately 430 000 ha under cane[1], with 72% of the annual production cultivated by 2 000 large-scale commercial growers. Forty eight thousand small-scale subsistence growers cultivate an additional 15% of the country's annual crop, while the remaining 13% is grown by milling companies (Isaacs, 2003). The industry extends latitudinally from 25°S to 31°S and

[1] In this study, for quantifiable attributes of sugarcane, the shortened term "cane" (*e.g.* cane yield, cane quality) was preferred as by general practice in the South African sugar industry.

altitudinally from sea level to approximately 1 100 m above sea level. Of the industry's cane, 68% is grown within 30 km of the coast in the KwaZulu-Natal province, while 17% of the cane production is found in higher altitude, frost prone, but high rainfall areas of the KwaZulu-Natal midlands. The remaining 15% is produced under irrigation in the drier areas in the northern KwaZulu-Natal and Mpumalanga provinces (Isaacs, 2003).

From an agronomical point of view, each sugarcane field is uniquely cultivated to suit the specific climatic, soil and socio-economic conditions present. This results in an extended range of management practices over the industry. Approximately 40 cultivars that are specially adapted to South African growing conditions are available. Sugarcane is harvested at ages ranging from 12 to 24 months, depending on climatic conditions, and is cultivated under different irrigation and rainfed scenarios.

Annual sugar production in South Africa has increased from 500 000 tons in 1950 to more than 2.5 million tons in 2001. In the past decade, sugar production has been approximately 2 million tons per annum and annual exports often exceeded 50% of the sugar produced. However, inter-annual variability of production has been around 25%, mainly attributed to a high variability in rainfall.

Climate, and especially rainfall, is probably the single most important factor that influences sugarcane production in South African. Inter-annual rainfall variability in the sugarcane growing belt ranges from 20% to 35% (Schulze, 1997) and the area is typically subject to relatively frequent severe and wide-spread droughts (e.g. 1983, 1992, 2003), occasional flood producing tropical cyclones (e.g. 1984, 2000) and less frequent mid-latitude cut-off low pressure systems producing excessive heavy rainfalls (e.g. 1987).

1.2.2 Yield[2] Forecasting in the South African Sugar Industry

Official communication channels have been implemented to forecast and convey estimates of the sugarcane crop size in South Africa. Prior to, and during, the milling season (normally April to December) growers survey the status of their crops and

[2] The term "yield" in this study always refers to the tons sugarcane, or tons sucrose, per hectare harvested.

issue a formal estimate of the total tonnage of sugarcane expected to be delivered to their representative mill. A committee at that mill, known as the Mill Group Board (MGB), collates this information. These committees are constituted of elected grower representatives and members of the particular milling company. Once a month, starting in March of each year, the MGB reviews the information at hand and forecasts the expected size of the season's crop. Mill Group Boards from the 15 mills in South Africa function independently and have adopted a range of forecasting techniques. These include simulation modelling (de Lange and Singels, 2003), empirical relationships (Lumsden *et al.*, 1998) and field scouting. Mill Group Board forecasts are then forwarded to the Industrial Affairs Division of the South African Sugar Association (SASA), where a national forecast is compiled and distributed to marketers, milling companies, government and other interested stakeholders.

There are several limitations to the forecasting process described above. First, growers have an incentive to over-estimate yields for delivery scheduling purposes (Wynne, 2001). Secondly, MGBs often base their forecasts on either subjective or non-representative information. Thirdly, there is no benchmark (or second opinion) available for their respective forecasts and, fourthly, MGBs cannot quantify the level of accuracy of their forecasts. Inaccuracies in MGB forecasts, resulting in sub-optimal decision making, have prompted the need for further research into forecasting of the sugarcane crop in South Africa.

Several researchers have highlighted potential benefits that could stem from more accurate sugarcane yield forecasts in Southern Africa. These include better irrigation management, improved control over the crop removal operation, more efficient mill operations and the ability to tailor agronomic practices to suit the expected climate of a specific season (Hildebrandt, 1998; Lumsden *et al.*, 1998; Schmidt, 1998; Ahmadi *et al.*, 2000; Lumsden, 2000). In addition to these, national scale marketing and sales activities, such as price fixing and exchange rate risk management, could also benefit from accurate yield forecasts. Few researchers in Southern Africa have, however, investigated alternative yield forecasting techniques to the MGB forecasts. Lumsden *et al.* (1999) used crop yield models and climate outlooks to forecast yields for the Eston mill in the KwaZulu-Natal midlands. McGlinchey (1999) undertook a similar, but more operational, study which demonstrated the modelling of sugarcane yields in

Swaziland, while Smith (1992) reviewed empirical methods to forecast yields in South Africa.

1.2.3 Benefits from Climate and Yield Forecasting

Hansen (2002) produced Figure 1.1 to illustrate the determinants on which benefits from seasonal climate outlooks may be derived by tailoring information for decision making. It is evident that a forecast system should have three distinctive characteristics. First, there needs to be scope for improvement within the industry if more information was available at a specific point in time (human vulnerability). Secondly, a prognostic skill of external forces, such as climate, needs to exist, and thirdly, participants in the system should have the capacity to identify and change certain properties to alleviate system vulnerabilities. The simple nature of Figure 1.1, however, overshadows many complexities that exist in the fields of climate forecasting, decision making under uncertainty and mitigating against vulnerabilities.

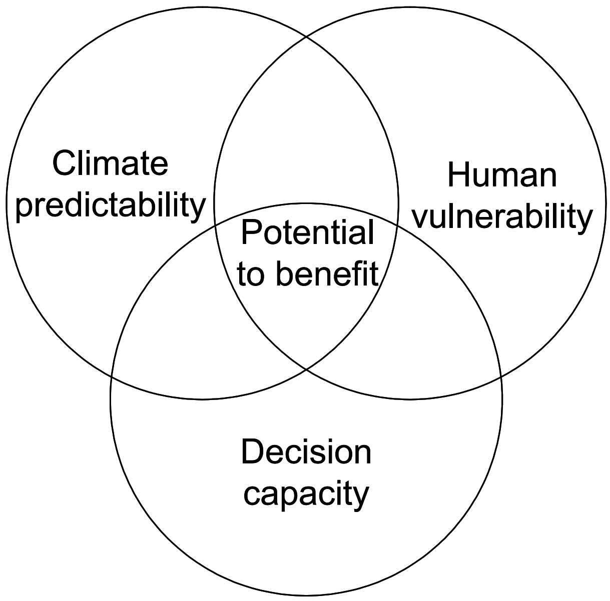

Figure 1.1 Determinants of the potential for humans to benefit from seasonal climate outlooks (from Hansen, 2002)

The skills associated with climate forecasts are increasing and Hansen (2002) noted that substantial portions of climate variability throughout much of the world can be explained by changes in the El Niño-Southern Oscillation (ENSO). The growing understanding of ocean and atmosphere systems has resulted in a degree of climate predictability with lead times of several months in advance (Cane, 1999; Hansen, 2002). Hansen and Jones (1999) concluded that the ability to skilfully forecast climate at lead times of several months raises the prospect for improving agricultural decisions at all levels, ranging from field-scale crop management to regional and global food security forecasting and mitigation.

Translating seasonal climate outlooks to production or economic outcomes is not a straightforward task (Hansen, 2002). Not only is this a relatively new and rapidly changing field of research, but climate outlooks remain probabilistic because of chaotic processes inherent in the atmosphere (Stern and Easterling, 1999). By implication, therefore, as noted by Horie *et al.* (1992), the most representative yield forecasts should be made after assuming the most probable future climate scenarios. Stone *et al.* (2000) emphasised the use of agricultural models to convert general climate outlook information into specific agricultural parameters, such as yield. Nevertheless, the probabilistic nature of climate outlooks and translated into yield forecasts necessitates risk-associated decision making capabilities among stakeholders (Thornton and Wilkes, 1998; Hammer, 2000a).

Stern and Easterling (1999) found that, as a consequence of the probabilistic nature of climate forecasts, stakeholder decision making capabilities are often ill-suited for using the information. They suggested that although scientific capabilities to produce accurate forecasts are limited, there is good reason to believe that much benefit can be gained by appropriately linking this capability to the practical needs in society. The management of climatic risk has always been an intrinsic and critical component of agriculture and farmers are inevitably required to make decisions under uncertainty (Eakin, 2000; Hammer *et al.*, 2001; Meinke *et al.*, 2001). Climatic uncertainty requires decision makers to anticipate a range of possibilities, often leading to risk alleviating management strategies that will reduce negative impacts in low production years. Unfortunately, risk alleviating management is often implemented at the expense of sustainability, inefficient resource utilisation and reduced productivity and profitability (Hansen, 2002).

There is nevertheless considerable scope to continue with research which aims to convey yield forecasts based on climate outlooks in a decision friendly manner. Marquis and Ray (1981) showed that crop forecasts could be expected to improve the domestic value of a crop under certain supply demand scenarios. Hansen (2002) confirmed that agricultural decision makers will realise the potential benefits of climate forecasting once the point of providing climate information alone has been exceeded. Hammer *et al.* (2001) also suggested that the most relevant information to decision makers will be the likely outcomes (e.g. production or income) of viable decision options within the system being managed.

1.3 Aim and Objectives of this Study

The aim of this study was to develop, evaluate and implement an appropriate and technologically advanced operational yield forecasting system for the South African sugar industry. Yield forecasting in this context was defined as the ability to capture, or explain, in advance a certain degree of the natural variability in local and regional sugarcane production.

With a relatively large number of related studies documented in the literature, it was considered vital for this study to align itself with recommendations and warnings issued by previous researchers. Stone *et al.* (2000) identified the following key considerations on the development of applications of climate outlooks for agricultural utilisation:

- Understand the system and its management.
- Understand the impact of climate variability.
- Determine opportunities for tactical management in response to forecasts.
- Evaluate the value of tactical decision options.
- Participate during implementation and evaluation.
- Provide feedback to climate forecasters.

The **first objective** of this study was to review previously documented forecasting methods. This was undertaken not only to consider the recommendations and warnings emanating from these studies, but also to explore different technologies employed in the field, such as remote sensing, simulation modelling and empirical methods. A synthesis was compiled that underlines the potential use of these technologies in the context of the South African sugar industry. The **second objective** was to conduct a stakeholder survey and a series of consultations to identify important issues of yield forecasts in the South African sugar industry. These included issues on management requirements, potential benefits, points of vulnerability and ways to alleviate certain problems. As a **third objective**, a reasonably flexible crop simulation environment had to be configured. This was done not only to allow the execution of operational yield forecasts, but also to allow evaluation exercises, such as comparing a history of production information with simulated values. The simulation environment had to be representative of the entire sugar industry in South Africa, providing enough information to delineate cane production regions into logical and manageable sub-units. The **fourth objective** of this study was to incorporate the use of climate forecasts in the yield forecasting system and to evaluate the system and some of its sub-components, such as the importance and quality of climate data. As a **fifth objective**, information transfer and user adoption of the above-mentioned forecasting system had to be evaluated against existing forecasts. Finally, the **sixth objective** was to recommend on technology adoption and to highlight further research in this field.

2 Review of Yield Forecasting Methods

2.1 Introduction

Prior to the development of any new sugarcane yield forecasting system, it is considered important to first review previous studies. This is undertaken in order to consider any recommendations and warnings arising from these studies and to explore different technologies employed in the field.

Technologically, crop yield forecasting methods can be subdivided into four main categories. These are statistical methods, crop yield modelling, remote sensing and methods combining more than one of the afore-mentioned categories. The objective of this chapter is to synthesise the potential use of different forecasting technologies in the context of decision making in the South African sugar industry.

2.2 Statistical Forecasting Methods

Statistical forecasting includes methods of regression modelling and neural networks. Historical information is used to establish relationships between yield and other frequently observed variables. These relationships are then used with currently available variables to forecast yields. The methods do not attempt to explain the dynamics prevalent within a system and can only be developed if a relatively long history of crop performance in relation to certain input variables is available for a particular area of interest. Nevertheless, Lumsden *et al.* (1999) and Matthews *et al.* (2000) noted that statistical forecasting methods are currently the most widely used numerical methods for operational crop forecasting. Input variables could include simple derivatives of temperature and rainfall, more complex integrated drought and crop moisture indices and pre-season indicators, such as ocean surface temperature indices.

For sugarcane, Lumsden *et al.* (1999) reported good performances using a locally derived statistical method to forecast sugarcane production from accumulated rainfall at the Eston mill in South Africa. Kuhnel (1994) investigated statistical relationships between ENSO and sugarcane production for different regions in Australia. He found that yields for certain sugarcane districts in Queensland were inversely related to the

Southern Oscillation Index (SOI) recorded in the year prior to harvest. In a similar study, Jury (1998) derived a sugarcane yield model for South Africa using the North Indian Ocean surface meridional wind, Southern Indian Ocean air pressure, East Atlantic Ocean surface meridional wind and South Indian Ocean outgoing longwave radiation as inputs for his model. Jury (1998) reported a coefficient of determination (R^2) of 0.69.

In various studies on other crops, such as wheat (e.g. Chmielewski and Potts, 1995), rice (e.g. Naylor *et al.*, 2002), maize (e.g. Cane *et al.*, 1994; Podesta *et al.*, 1999) and soyabean (Hansen *et al.*, 1998), researchers have also derived statistical relationships between crop yields and different climatic variables. Moreno *et al.* (2001) noted that the use of simple variables, such as solar radiation, temperature and precipitation to forecast the size and quality of crops has resulted in inconsistent results. The reason for this is that important driving factors, such as the soil water balance, were often neglected. Stephens *et al.* (2000) overcame this problem by developing a soil moisture stress index for wheat by using daily rainfall and temperature data and weekly radiation to drive a simple water balance algorithm. Likewise, Meyer *et al.* (1991) linked crop yields with integrated climate indices, only in their case using existing indices from other applications.

Integrated climate indices assimilate large quantities of data, such as rainfall, evapotranspiration and soil properties, into a single number representative of the "big picture" (Hayes, 2001). These indices might be suitable for statistical yield forecasting (Meyer *et al.*, 1991). Hayes (2001) highlighted various general indices, such as the crop moisture index (CMI), the surface water supply index (SWSI), the Palmer drought severity index (PDSI) and the drought reclamation index, which are generated for the USA by their National Drought Mitigation Center. After assessing one of the indices, *viz*. the PDSI, Alley (1984) and Karl and Knight (1985) stated the following advantages and limitations: First, it provided decision makers with a measurement of the abnormality of recent weather in a region. Secondly, it provided an opportunity to place prevailing conditions in an historical perspective and, thirdly, it provided spatial and temporal representations of historical droughts. Limitations included the fact that the index might require local calibration and that input variables describing the soil might be over simplified. The indices also do not perform well in regions where there are extremes in the variability of rainfall or runoff, such as Australia and large parts of South Africa (Smith *et al.*, 1993). Integrated climate indices are not readily available in South Africa. However, generating and using such indices for yield forecasting should nevertheless be an important consideration for the future.

2.3 Yield Forecasts using Crop Yield Models

Crop yield models provide a conceptual / mechanistic explanation of a crop's response to external factors such as the climate and soil. These models have been developed and applied in many areas of research, including yield benchmarking (e.g. Inman-Bamber, 1995; Nielsen *et al.*, 2002), irrigation scheduling (e.g. McGlinchey *et al.*, 1995; Steele *et al.*, 1996) and crop nutrition (e.g. Thorburn *et al.*, 2002; Gungula *et al.*, 2003). Matthews and Stephens (2002) provided a comprehensive overview on the development and applications of crop models in developing countries. Crop yield models are becoming increasingly important in translating information on climate variability into forecasts and recommendations tailored to the needs of decision makers (Hansen and Jones, 1999). Bannayan and Crout (1999) emphasised that one of the most important potential applications of crop modelling is in crop forecasting. A regional crop forecasting system integrates crop specific modelling with actual climate to date and projects the crop condition forward using probable future climate scenarios (Stephens *et al.*, 2000; Potgieter *et al.*, 2002). Complex crop yield models quantify the contribution of various physiological processes and climatic elements to yield (Cheeroo-Nayamuth *et al.*, 2000). Hansen and Jones (1999) and Stone *et al.* (2000), however, argue that crop yield models do not need to be comprehensive or complex to be useful, mainly because simpler models require less input parameters.

Various researchers have developed model-based yield forecast systems, both for sugarcane and for other crops. McGlinchey (1999) combined recorded climate data and climate forecasts issued by the South African Weather Service (SAWS) and the University of Zululand Climate Impact Predictions (Anon., 1999a) to serve as input for the CANEGRO model (Inman-Bamber, 1991). Nine sugarcane crops with different harvest months were simulated with future climate scenarios selected from historical climate records according to the climate outlook. This resulted in a probabilistic range of yields. Promburom *et al.* (2001) used a modified version of the CANEGRO model to forecast regional sugarcane production in Thailand. An accuracy of 4.8% in yield at field and provincial scales was reported. Potgieter *et al.* (2003) used the APSIM-Sugar model (Keating *et al.*, 1999) to assess yield forecasts at a field scale in Bundaberg, Australia. Different seasons with similar SOI values were selected from a 110 year climate record and simulation results were statistically evaluated. Lumsden *et al.* (1999) evaluated the use of different types of models to forecast sugarcane yields in the KwaZulu-Natal midlands in South Africa. They concluded that crop yield models outperformed statistical methods, such as linear regression models. In addition, they also supported the use of a less complex crop yield model, *viz.* the *ACRU*-Thompson model (Schulze, 1995), as opposed to the more complex CANEGRO model.

For other crops, Supit (1997) and Stephens *et al*. (2000) demonstrated the use of crop models and seasonal forecasts for operational crop forecasts in the Australian and European wheat industries, respectively. De Jager *et al.* (1998) proposed a framework for forecasting maize yields in South Africa. Supit (1997) also referred to the extensive use of crop yield models from the Crop Growth Simulation System (CGMS) for forecasting agricultural yields in the European Union.

Thornton and Wilkes (1998) mentioned the necessity to re-simulate crops at regular intervals after updating climate data with the most recent records when using crop yield models to forecast yields on a regional scale. They also recognised the necessity to simulate different future scenarios in order to generate statistical distribution functions of probable yields. The spread of these distribution functions can be expected to decrease as the season progresses, because of increased certainty of the climate to date and reduced uncertainty in the future climate (Thornton and Wilkes, 1998). In contrast to simulating future scenarios probabilistically, Supit (1997) and Roebeling *et al.* (1999) defined two indices termed the yield indicator and the yield difference indicator. These indices are comparisons between yields of incomplete crops simulated up to the date of the last climate record and equivalent yields simulated under conditions of no water stress, and simulated for the previous season, respectively.

Hansen and Jones (1999) noted that crop yield models have been developed using the information from management intensive trial experiments. Commercial production, on the other hand, is hampered by many external factors, such as sub-optimal management and incidences of pests and diseases, which are not normally included in the models. Day (2001) mentioned that the exclusion of some of these factors have been a primary reason for some forecast studies to have failed when attempting to simulate large scale yield variability. The aggregation of model outputs to regional scales therefore often result in significant over-estimates of production and might require additional empirical corrections (Hansen and Jones, 1999). These empirical corrections should, however, not be seen as replacements for process-level crop models where physiological descriptions capture many mechanisms of crop response to weather variability and which have proven useful for regional applications (Hansen and Jones, 1999). Crop yield models should, therefore, not be discarded without substantial justification (Hansen and Jones, 1999).

Hansen and Jones (1999) identified two opposing schools of thought regarding the required complexity of crop yield models when results are to be scaled up to regional levels. Rabbinge (1993), cited by Hansen and Jones (1999), suggested that model complexity should increase with spatial scale, while other authors such as Addiscott (1993) and Heuvelink (1998), cited by Hansen and Jones (1999), have argued the opposite. Hansen and Jones (1999) believe that a hybrid approach would be most applicable.

Struzik (2001) noted the importance of satisfying both the spatial and temporal data requirements at which agricultural models operate. Temporal scales (i.e. time steps) of crop yield models are normally in the order of days, while spatial scales are normally in the order of one hectare. Horie *et al.* (1992) and Lumsden *et al.* (1999) emphasised the importance for climate input to be timely in the case of forecasting yields. Liu and Scott (2001) noted that a lack of solar radiation data is common in many countries and can be a major limitation to regional applications of crop yield models. The provision of timely and suitable high resolution inputs for crop models when conducting yield forecasts could, therefore, be problematic. However, various solutions to these problems have been investigated. These include spatialisation methods based on ground measurements (Struzik, 2001), satellite information (Struzik, 2001), automatic weather station networks (AWS, e.g. Georgiev and Hoogenboom, 1999; Singels *et al.*, 1999b) and estimating / deriving missing input variables, such as solar radiation, from other measurements (e.g. Bristow and Campbell, 1984; Hunt *et al.*, 1998; Liu and Scott, 2001).

Possible disadvantages of crop yield models are the lack of genetic coefficients to simulate yields from different cultivars (Ogoshi, 1995) and a lack of credibility among industry stakeholders (Meinke *et al.*, 2001). Bannayan and Crout (1999) note that an important advantage of crop models is the ability to quantify risk under uncertain conditions after conducting frequency analyses on outputs from repeated simulations.

2.4 Remote Sensing for Crop Yield Estimation and Forecasting

Remote sensing is a grouping of usually airborne techniques used for gathering information about an object or an area without coming into physical contact with it (Anon., 1997). Measurements of crop reflectance are used to derive estimates of photosynthesis, water stress, pests and diseases (Wisiol, 1987) and levels of management (Maselli *et al.*, 1993). Remote sensing can be an attractive alternative to traditional field scale scouting because of its ability to cover large areas at relatively low cost (Anon., 1997). Gadekar (1998), for example, reported a 60% reduction in the cost of a crop forecast application when remote sensing technologies were employed in place of conventional field scouting.

Three satellite missions are commonly used to provide information for agricultural purposes. They are the Land Satellites (LANDSAT), Systeme Pour l'Observation de la Terre (SPOT) and the NOAA Advanced Very High Resolution Radiometer (AVHRR). Table 2.1 displays some important image properties from these satellites.

Previous researchers have utilised different remote sensing techniques for yield estimates. These may be grouped into three main categories, namely the derivation of vegetation indices (e.g. King and Meyer-Roux, 1990), estimates of radiation interception (e.g. Jaggard and Clark, 1990) and estimates of canopy temperature (e.g.

Gardner *et al.*, 1981). Schmidt *et al.* (2000) concluded that vegetation index data from the NOAA-AVHRR satellite are likely to give good indications of mill average yields in the South African sugar industry. Higher resolution vegetation index information, such as that obtained from LANDSAT and SPOT, could theoretically also be used to estimate sugarcane conditions at farm and field scales (Schmidt *et al.*, 2001). Lumsden *et al.* (1999), Schmidt *et al.* (2000) and Schmidt *et al.* (2001) concluded that there is significant scope to use remote sensing information for yield forecasting in the South African sugar industry. In a follow-up assessment, Gers (2003) found that a yield estimating application using a single-date LANDSAT image was not sufficiently accurate, suggesting that temporal information between consecutive images needs to be integrated in order to generate reliable yield estimates.

Table 2.1 Important image properties of different satellites frequently used for remote sensing in agriculture (*pers comm.* Dr. A. Sand, CNES - Program Directorate, Toulouse, France)

Satellite	Resolution (m^2 per pixel)	Flyover frequency (days)
LANDSAT	900	16
SPOT	400	26
NOAA-AVHRR	1 210 000	9

Remote sensing holds large potential benefits to crop yield forecasting applications. The technology is cost effective, timely, accurate, flexible and information is spatially consistent (Gadekar, 1998). A constraint, however, may be frequent cloud encumbrances (Smith *et al.*, 1995).

2.5 Combined Methods for Crop Yield Forecasting

Roebeling *et al.* (1999) estimated relative evapotranspiration from Meteosat satellite information. This information was used in a statistical exponential linear model to forecast maize and sorghum yields in the Horn of African region. The models, especially those for maize, performed reasonably well (R^2 values between 0.70 and 0.85). Roebeling *et al.* (1999) pointed out that the model was more accurate for drought sensitive crops, such as maize, as opposed to more robust crops such as sorghum.

Various studies have resulted in improved crop forecasts after combining remote sensing technologies with crop modelling (Maas, 1988; Horie *et al.*, 1992; Roebeling *et al.*, 1999). Promburom *et al.* (2001) coupled simulation models with remote sensing technologies to forecast sugarcane yields over large areas in Thailand. By using satellite information, Maas (1988) managed to reduce the yield forecast error from 30% to 2% after adjusting the simulated green leaf area index. Maas (1988) confirmed that relatively simple crop yield models performed well, but noted that the combination of remote sensing information and crop modelling compensated for each other's weaknesses. Horie *et al.* (1992) also concluded that this technique resulted in the most effective yield forecast method investigated.

2.6 A Synthesis of Yield Forecasting Methods for the South African Sugar Industry

Statistical yield forecasting still seems to be the most commonly used method in agriculture. This is so despite strong indications by Maas (1988) and Horie *et al.* (1992) that more accurate methods had been in existence for more than a decade. A possible reason for the lack of technology uptake could be that statistical methods are easy to use, require few input data and can easily be developed by industry stakeholders. If this hypothesis is true, then it implies that tangible shortcomings exist between industry stakeholders and scientists who are developing more advanced yield forecast systems. It is notable that few scientific descriptions of yield forecast systems have ventured into a discussion on possible information transfer and technology adoption. This would confirm some of the issues highlighted from the literature in Section 1.2.3 and emphasises the need to not only develop a scientifically acceptable yield forecast system, but to also assess the concerns and reservations of people intending to use the information.

There seems to be elucidating evidence that statistical methods are less accurate for crop yield forecasting than either crop yield modelling or remote sensing forecasting techniques. Fundamentally, statistical methods are based on a re-occurrence philosophy, suggesting that if enough historical data exist, the future will be largely explainable from the past. While this approach is disputable, especially under conditions of possible climate change and genetic improvement, cognisance should be taken that important driving parameters, such as ENSO and other integrated indices, are imbedded in statistical models. It would, however, be more appropriate to link sea surface temperature indices with climate responses rather than directly with yields. Hansen and Jones (1999) pointed out that the historical yield data often used to develop statistical relationships are representative of non-optimal management conditions from administrative production regions and that they will inevitably include intrinsic trends caused by other driving factors, such as area expansions, trends in yield decline and cultivar and agronomic changes. Unlike crop yield models, which often require daily climate input data, most statistical methods require input at a courser resolution. However, many scientific and remote sensing techniques exist for the very purpose of estimating variables in data poor areas. Some statistical methods have been developed using integrated indices, such as drought and / or soil moisture indices. This affirms the need for a mechanistically rigorous systems approach. It is concluded from this discussion that statistical methods should only be used if insufficient information exists to use crop yield models.

Crop yield models used in isolation do capture direct crop responses to climate, but often omit many compelling driving factors such as incidences of pests and diseases and changes in cultivars. Because of these limitations, models often over-predict yields, resulting in large inconsistencies when extrapolated over wider areas. Remote sensing techniques seem the most appropriate way to overcome some of the above limitations of crop yield models. Various remote sensing techniques exist to establish values related to a crop's status with respect to its cultivar, water stress, rate of growth, leaf area index, pest and disease incidence and yield. These values can be used to reset the crop's status in the models. A technique that does not require remote sensing is to express simulated yields in relative terms by comparing them with simulated yields from another season or with simulated potential yields. This technique, however, assumes that parameters not simulated, such as pests and diseases, will have a uniform effect over all the simulations. It can be assumed that the aforementioned technique will be inferior to a well developed method which combines a crop yield model with remote sensing information.

Crop yield model applications are well suited for decision making under conditions of uncertainty. Uncertainty can be addressed by repeated simulations and expressing outputs as frequency distributions, *i.e.* Monte Carlo approach. The future climate is one of the main uncertainties in yield forecasting. Some studies disregard climate outlooks entirely and aim only at providing single non-probabilistic values of future yields (e.g. Supit, 1997; Roebeling *et al.*, 1999). Stone *et al.* (2000), however, emphasised the importance of forecasting both yield shift and yield uncertainty and Potgieter *et al.* (2002) showed significant added value when including some knowledge of the future climate in a yield forecast.

From the above discussion it may be concluded that the integration of a crop yield model forecast system with a strong remote sensing component would be well suited for the South African sugar industry. The incorporation of climate outlooks in a probabilistic manner is also supported. Sufficient climate data for crop yield modelling exist in the South African sugar industry and the industry has also engaged in a comprehensive crop modelling research programme. A simple, although mechanistically-based, locally calibrated crop model seems most appropriate for a yield forecasting application. Remote sensing technologies in the industry are, at present, still under-utilised and have therefore been excluded in this study. The development of a model-based yield forecast system would, however, allow for subsequent integration with remote sensing technologies.

A wide range of issues has to be addressed for the development of a model-based yield forecast system. These include establishing industry stakeholder requirements, selecting a suitable crop yield model and configuring and preparing input data. The following three chapters report on these facets of the research.

3 Industry Stakeholder Requirements for Sugarcane Yield Forecasting

3.1 An Overview of Stakeholder Requirements for Yield Forecasts

The previous two chapters highlighted some issues related to information transfer and the adoption of yield forecasts in the South African sugar industry. In Chapter 2 it was hypothesised that a lack of communication and understanding between industry stakeholders and scientists may be a major attributing factor for the failure of more technologically advanced yield forecast systems. It was, therefore, necessary to assess the requirements and reservations of a range of stakeholders before another yield forecast system was to be developed.

Fischhoff (1994) and Meinke *et al.* (2001) suggested that, in order to avoid failure, scientists and stakeholders should form partnerships in acknowledging the complexity and determining the demands of a forecasting system. Stone *et al.* (2000) warned that, although scientists are capable of structuring and quantifying their thoughts about system dynamics, they are often isolated from the requirements of analysts and practitioners. Likewise, Stone *et al.* (2000) also noted that decision makers and practitioners may not always know what to ask for, since they might be unaware of the possibilities of a scientific system.

A wide range of factors have previously been attributed to the non-adoption and misuse of climate forecasts. Eakin (1999) noted that small-scale farmers might not find climate forecast information useful. Likewise, Stern and Easterling (1999) found that those with the most education and money, such as large-scale farmers, will benefit from useful information before small-scale enterprises do. Pulwarty and Redmond (1997) and Nicholls (2000) conducted surveys in Columbia and Australia, respectively, and found that decision makers raised the following reservations about the adoption of seasonal climate outlooks:

- Forecasts are difficult to interpret.
- Models do not integrate all the relevant information.
- Uncertainty and dissatisfaction exist over the accuracy of forecasts.
- Large fluctuations exist between successive forecasts.

- Additional information is necessary before decisions could be made.
- Insufficient procedures exist to integrate forecasts with decision making.
- The value of the forecast is not obvious.
- It is difficult to assess forecasts.
- Contradictions exist between competing forecasts.
- A history of previous forecasts is not available.
- There is a lack of access to expertise.
- There is a lack in communicating the forecast to end-users.

Although these points refer specifically to the adoption of seasonal climate outlooks, most concerns are generic and should caution developers of yield forecast systems. Forecasts will only become useful once they have met the decision maker's requirements in terms of attributes such as timing, lead time, accuracy, currency, parameters and spatial and temporal resolution (de Jager *et al.*, 1998; Stern and Easterling, 1999; Nicholls, 2000). General principles of communication could also increase user adoption. These include

- Presenting information in short and simple terms,
- Giving additional guidance on how to take advantage of the information,
- Allowing frequent repetition of the information over different channels, and
- Transferring information through people who are trusted by the users (Stern and Easterling, 1999).

Previous studies have highlighted aspects of spatial and temporal resolutions that should be considered to augment the usefulness of forecasts. Spatially, Lumsden *et al.* (1999) noted the importance of forecasting yields on a mill supply area scale, while temporally de Jager *et al.* (1998) and Lumsden *et al.* (1999) recommended updated forecasts on a monthly basis. Gadekar (1998) and Everingham *et al.* (2002a) emphasised the strategic importance to provide stakeholders with forecasts with long lead times.

Thornton and Wilkes (1998) and Hammer (2000b) stressed the importance to communicate yield forecasts probabilistically. The application of seasonal forecasts in agriculture concerns risk management, which is only possible under a clear understanding of the likelihood of certain outcomes (Hammer, 2000a). Communicating uncertainty and risk in a simple manner and with clarity does, however, present some difficulties. Stone *et al.* (2000) suggested that the use of simple bar graphs (histograms) showing the occurrences of different likely outcomes and their relation to the long-term mean or median might be appropriate. Likewise, plotting graphs of probabilities of non-exceedance, as done by Singels and Bezuidenhout (1999), might circumvent opportunities for users to misinterpret information.

Stone *et al.* (2000) indicated that more sophisticated management capabilities are required to deal with risk-associated decision making. In climate forecasting, for example, it has been shown that users might prefer to adopt inferior outlooks, with scientifically unjustified narrow ranges of uncertainties, if the possible outcomes of other outlooks are too wide (Nicholls, 2000). Users should also not become disheartened, or reckless, if they "lose" or "benefit" from the forecast in any particular season. Rather, the information should be viewed in a statistically consistent manner over many seasons in order to benefit (Meinke and Hochman, 2000).

There are numerous areas in which decision making can be enhanced once yield forecasts have been transferred and interpreted in a correct way (Horie *et al.*, 1992; Wood, 1995; de Jager *et al.*, 1998; Everingham *et al.*, 2002a). Everingham *et al.* (2002a) subdivided stakeholders in the Australian sugar industry into four sectors, *viz.* farming, harvesting / transport, milling and marketing.

- In the farming sector, seasonal climate outlooks and yield forecasts can enhance decisions on the timing and methods of planting, selection of herbicides, decisions on irrigation strategies and the determination of fertiliser quantities (Horie *et al.*, 1992; Lumsden *et al.*, 1998; Schmidt, 1998; Everingham *et al.*, 2002a).
- The harvesting and transport sectors can anticipate benefits from using climate outlooks and yield forecasts when planning harvesting and haulage schedules, pre-empting capacity requirements and planning and directing strategies for machinery sales (Gadekar, 1998; Lumsden *et al.*, 1998; Everingham *et al.*, 2002a).

- Millers can use seasonal outlooks and yield forecasts to determine the number of staff required as well as to establish the opening date, crush rate and length of the milling season (Hildebrandt, 1998; Everingham *et al.*, 2002a).
- In the marketing sector, yield forecasts can be used to plan an overall marketing strategy, including forward selling, shipping and warehouse requirements (Gadekar, 1998; Everingham *et al.*, 2002a).

An assessment of stakeholder requirements for yield forecasts should incorporate all the issues discussed up to this point. These include the spatial and temporal resolution for forecasts, format of information, coping with uncertainty and understanding the stakeholders' management skills as well as incorporating areas of potential benefits. Hansen (2002) emphasised the use of exploratory surveys, which are designed to establish stakeholders' perceptions and perspectives. It was noted that few surveys have been documented in the literature. One such survey was the Climate Variability in Agriculture R&D Program (CVAP) in the Australian sugar industry, whereby key decisions, financial implications and preferred mechanisms of information transfer among stakeholders were established (Anon., 1999b). Hansen (2002) warned that a distinction should be made between what stakeholders want and what they actually need. Three important characteristics of yield forecasts were highlighted by Hansen (2002). These are

- Site specificity (e.g. importance of scale, spatial variability and spatial representivity of climate outlooks and data),
- Temporal specificity (e.g. timing of forecasts to suit decisions and intra-seasonal characteristics, such as start of rainy season and duration of dry spells), and
- Forecast skill (for users concerned with managing risk).

The aim of this section is to establish vulnerabilities in the South African sugar industry that could be alleviated by using yield forecasts. Specific objectives are first, to identify key areas where yield forecasts could enhance decision making and secondly, to establish their financial impacts, strategic timing, spatial and temporal resolution requirements, desired reporting formats, required parameters and preferred communication media.

3.2 A Questionnaire Survey for Industry Stakeholders

A questionnaire (see Appendix A) was distributed to a wide range of stakeholders in the South African sugar industry, who elaborated on the potential expected benefits of yield forecasts in their respective sectors. They were requested to state their preferences in the yield forecast's time of issuing, its spatial and temporal resolution, output parameters, reporting format and means of communication. The questionnaire mainly comprised of multiple-choice questions and stakeholders were often allowed to make more than one choice per question, if they so preferred.

The outcomes of the questionnaire, in conjunction with previous literature, were used to identify key issues where yield forecasts could have significant potential impacts in the industry. Individual stakeholders working in these sectors were consequently consulted for further information. All outcomes are reported in this chapter.

3.3 Results from the Questionnaire Survey

Thirty nine completed questionnaires were returned after the survey was widely broadcasted. No record was kept of the number of persons who did not respond to the questionnaire. Respondents were representative of commercial- and small-scale grower groups, miller-*cum*-growers, national and international millers, marketers, transporters and exporters. Figure 3.1 displays the relative contribution of representatives from these different sectors. It should be noted that respondents often felt that they were representing more than one industry sector.

Figures 3.2 and 3.3 summarise various time-related issues on yield forecasts that were addressed during the survey. Figure 3.2 captures stakeholder preferences in (a) temporal resolutions and (b) forecast frequencies. It is evident from Figure 3.2(a) that 96% of representatives would be satisfied if the milling season (i.e. April – December) were subdivided and reported upon a monthly basis. Users requiring lower temporal resolutions, such as seasonal production, could calculate the mean over the above-mentioned monthly values. Figure 3.2(b) shows that 96% of the representatives required at most monthly updates of yield forecasts. Figure 3.3 summarises the time of the year when stakeholders require yield forecasts. It is evident that the demand for forecasts increases towards the opening date of the milling season in April (milling season shaded in grey). Most representatives also expressed a preference for obtaining updates at the commencement of each month.

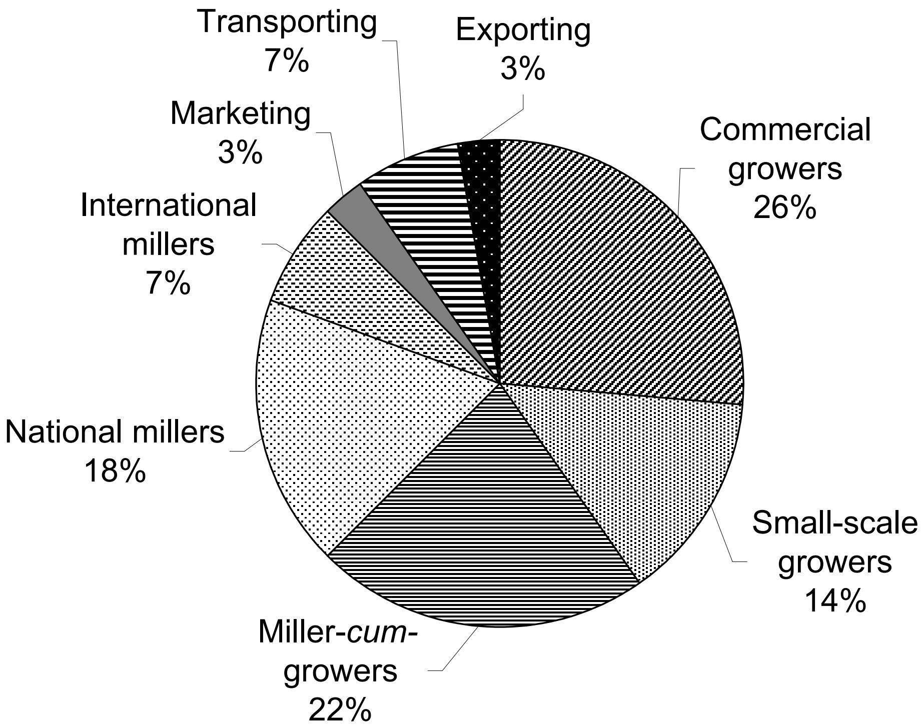

Figure 3.1 Distribution of stakeholder sectors which responded to the questionnaire survey

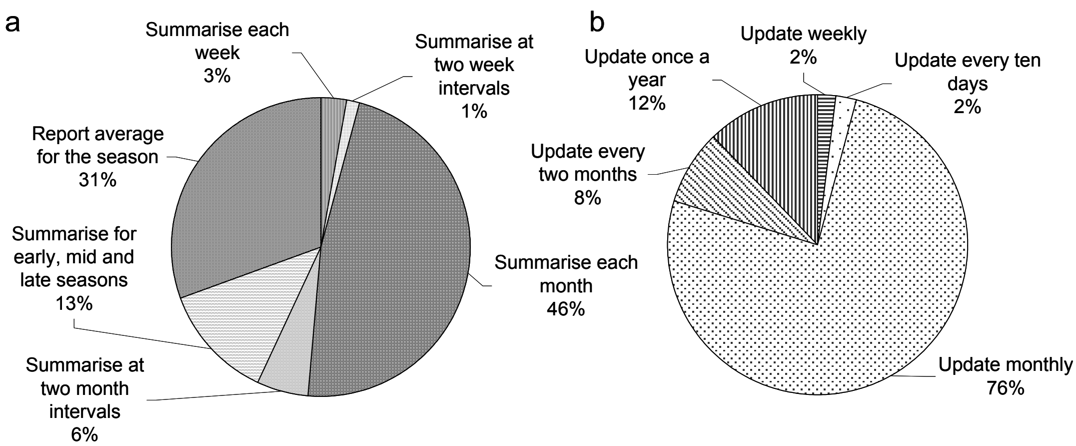

Figure 3.2 Stakeholder preferences (a) in temporal resolutions of reporting and summarising information and (b) in frequencies of forecast updates

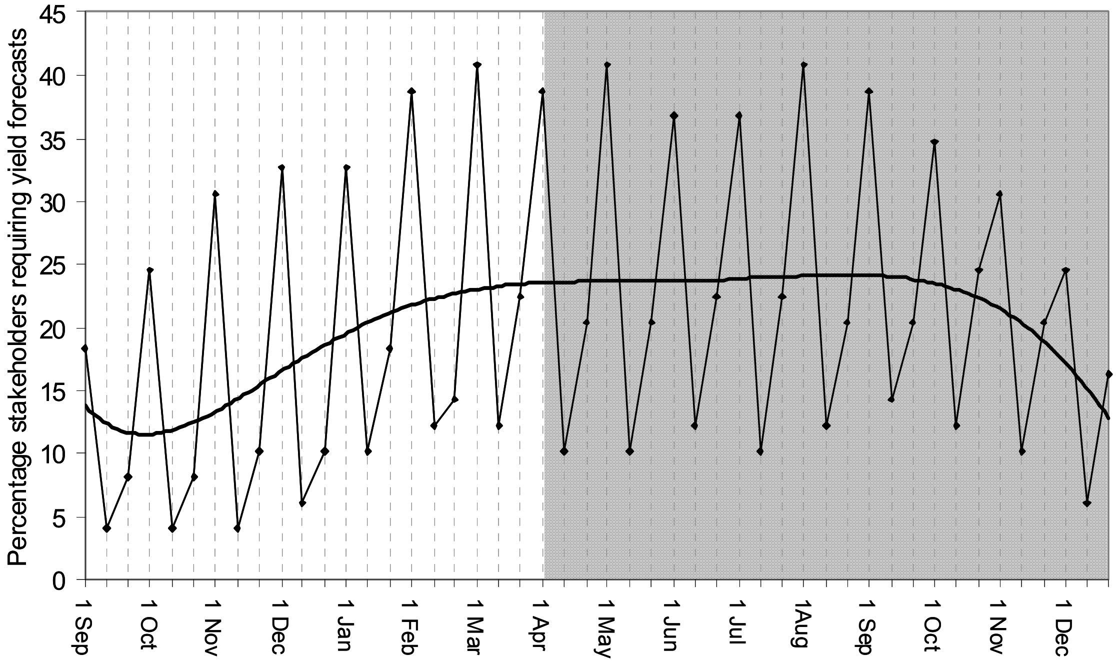

Figure 3.3 A time chart indicating the percentage of stakeholder respondents who required yield forecasts at certain times in the year, with the thicker line indicating a general trend of increasing demand towards the opening of the milling season (area shaded in grey)

Figure 3.4 summarises a range of preferences relating to the format, contents and methods of information transfer of the yield forecast. A small majority of respondents requested yield forecasts to be issued with confidence bands (Figure 3.4a). This denotes that stakeholders might be aware of the complexities of risk-based decision making. Most stakeholders showed interest in obtaining forecasts of all the parameters suggested, *viz.* cane yield, sucrose yield, fibre content and reduced sugar (non-sucrose) content (Figure 3.4b). A slight majority of stakeholders proposed that yields be expressed as a percentage of an equivalent crop grown in the previous season (Figure 3.4c). Although not reflected in Figure 3.4(c), it should be noted that a majority of grower representatives would also be interested in a comparison between the current crop and the previous crop grown on the same field. A large majority of stakeholders in this survey preferred email as a means of communication (Figure 3.4d). A relatively large number of stakeholders also expressed an interest to conduct further post-forecast calculations using a decision support program (DSP).

Figure 3.5 shows the preferred spatial resolution at which yield forecasts should be made. It is noted that a farm scale, homogeneous climate zone scale and mill supply area scale were strongly supported.

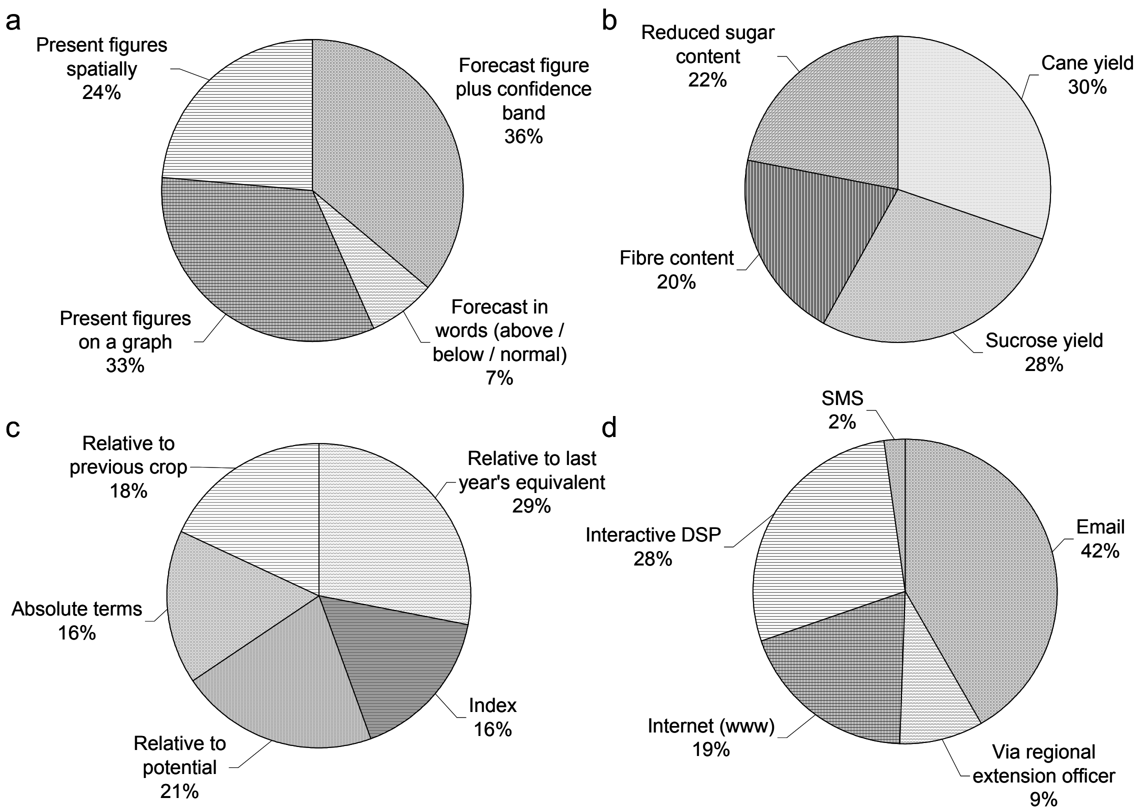

Figure 3.4 Stakeholder preferences in (a) the manner by which yield forecasts should be communicated, (b) specific parameters to be reported, (c) the way by which yields should be expressed and (d) what communication medium should be used

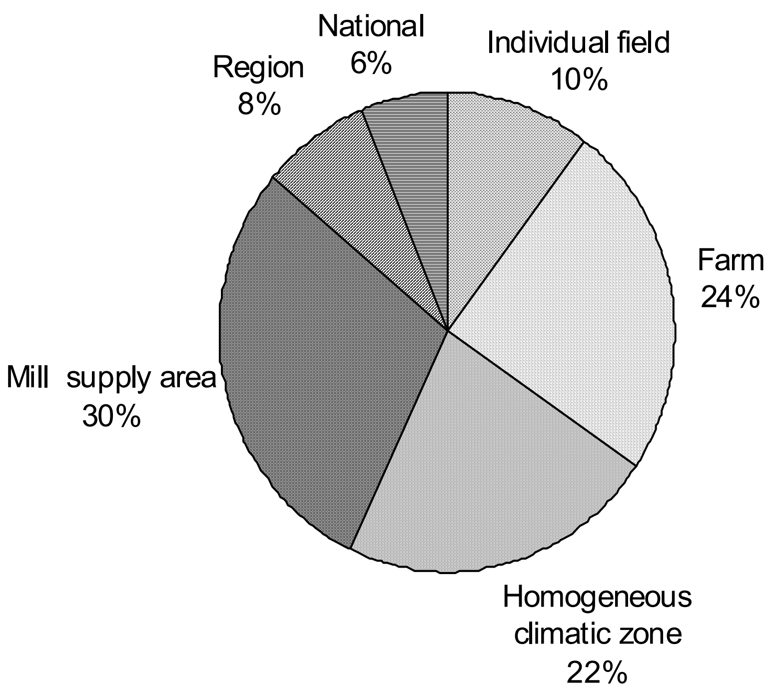

Figure 3.5 Stakeholder preferences with regards to the physical size of yield forecast units

Subsequent to the stakeholder survey, a list of four key issues, at different scales, where yield forecasts could significantly benefit the industry was compiled. These issues were

- Adjusting international marketing strategies,
- Enhancing national financial systems,
- Fine tuning milling operations, and
- Adjusting fertilizer recommendations.

International marketers intend to use yield forecasts prior to the start of the milling season. Accurate forecasts of production would enable marketers to exploit price hedging, to forward-sell anticipated sugar surpluses at higher profit margins and to cut expenditure on freight fixtures. Strategic marketing prior to the milling season could increase profits by as much as R18 million per annum if forecast errors could be reduced from 7% to 2.5% (*pers comm.* Mr. Q.L. Hildebrandt, International Marketing Director, SASA).

On a national scale, yield forecasts are required, first to estimate the amount of income to be disbursed from millers to growers and, secondly, to estimate tax rebates for communal small-scale growers. In both these cases, forecasts would not improve primary productivity or profitability, but would assist in anticipating and streamlining large financial transactions (*pers comm.* Mr. D.P. Rossler, General Manager, Umthombo Agricultural Finance, SASA).

Determining the commencement and closure of a mill, as well as proposing crush rates for the mill prior to April, can significantly increase mill productivity. Hildebrandt (1998) illustrated how between R120 000 and R800 000 could be lost per season at the Noodsberg mill alone (KwaZulu-Natal midlands) if production forecasts contained an error of between 2% and 11%. These losses were ascribed to lower cane qualities, less favourable cropping cycles and increased grower and miller expenses. Lumsden *et al.* (1999) also estimated an approximate R800 000 economic benefit per season for the Eston mill (KwaZulu-Natal midlands) by adopting a model-based yield forecast system. At the Umfolozi mill (Northern KwaZulu-Natal) the length of the milling season may vary from 30 to 37 weeks. If the milling season commences too early, sugarcane supplies will deplete too early, resulting in the unnecessary crushing of low quality cane at the beginning of the season. An improvement of 0.05% in annual cane quality could be anticipated if the opening of the milling season were correctly adjusted by one week (*pers comm.* Mr. J. de Lange, Cane Procurement Manager, Umfolozi mill, Illovo Sugar Ltd). This results in higher production of sucrose and molasses valued at approximately R500 000 per annum. It is likely that the above-mentioned benefits would also pertain to the other 12 mills in South Africa. Accurate yield forecasts could, therefore be expected to increase mill productivity by approximately R9 million per annum.

On a farm scale, stakeholders have indicated that the adjustment of fertilizer applications based on yield forecasts would enhance profitability and sustainability. Assuming 300 000 ha of cane harvested per annum in South Africa, a savings of just 5% to standard fertilizer recommendations could reduce expenditure by as much as R21 million (*pers comm.* Dr. D.J. Nixon, South African Sugarcane Research Institute). This would, however, involve high accuracies in field scale yield forecasts between 10 and 24 months in advance, which implies unrealistically high skilled climate outlooks, extensive field scale data collection and near-perfect crop model output.

3.4 Discussion and Conclusions

Although speculative and often more intricate than discussed, the survey's results point to potentially large savings from yield forecasts for some sectors in the South African sugar industry. Stakeholders generally envisage that yield forecasts could significantly enhance decision making across a range of scales in the industry. International marketing, national financing and mill operations are likely candidates to benefit from yield forecasts. At a farm level, growers require longer lead times and higher accuracies before agronomic benefits can be expected. Yield forecasts may, however, assist growers in general business issues such as managing cash flows, investments and purchasing new capital.

Strong information transfer guidelines were compiled from the results of this survey. Yield forecasts should be issued at the commencement of each month, beginning in the September prior to the next milling season. Reports containing summaries of cane yield, sucrose yield, reduced sugar content and fibre content should be communicated via email. Different crops, each harvested in a different month of the milling season, should be reported. Certain stakeholders would also require a single production value for the entire season. Results should be summarised for homogeneous climate zones, but should also be aggregated at a mill and national scale. All forecasted parameters should include confidence bands and values should be compared with those of the previous season. A DSP enabling users to aggregate information and to express results relative to different reference crops would assist a wide range of stakeholders. The DSP should, however, not replace a concise report on forecasted yields. Based on general principles of communication, it is strongly encouraged that yield forecasts be communicated through the South African Sugar Association and the South African Cane Growers Association. Local extension officers and agricultural economists should be encouraged, where possible, to integrate yield forecasts with agronomic, business and risk management approaches.

4 A Review of Sugarcane Crop Yield Models for Yield Forecasting in South Africa

4.1 Introduction

In the previous two chapters it was concluded that crop model-based yield forecasts would enhance decision making across various sectors of the South African sugar industry. This chapter briefly considers different sugarcane crop yield models for yield forecasting purposes in South Africa.

Crop yield models perform simulations by numerically integrating fundamental processes with the aid of computers (Sinclair and Seligman, 1996). A crop's response at a given location is simulated using measured or estimated input parameters that describe soil and atmospheric conditions. Three suitable models were identified for conducting yield forecasts in the South African sugar industry. These are, CANEGRO (Inman-Bamber, 1991), *ACRU*-Thompson (Schulze, 1995) and Canesim (formerly known as IRRICANE, Singels *et al.*, 1998). These models have been tested in South Africa and have been used previously in crop forecasting applications (Lumsden *et al.*, 1999; McGlinchey, 1999; de Lange and Singels, 2003). In this chapter each model's history, level of complexity, required input parameters and accuracy is reviewed with the aim of identifying the most suitable model for yield forecasting.

4.2 CANEGRO

The CANEGRO model is based on the concepts and structure of the CERES-Maize model (Jones and Kiniry, 1986) and was developed for a single cultivar (NCo376) to assist in quantifying the impacts of a local pest in the South African sugar industry, *viz. Eldana saccharina* (Inman-Bamber, 1991). The model has since been included as the official sugarcane model in the Decision Support System for Agrotechnology Transfer (DSSAT, Inman-Bamber and Kiker, 1997). Since its development in 1991, some improvements have been made to the model's mass balance (Singels and Bezuidenhout, 2002) and its ability to accommodate more cultivars (Cheeroo-Nayamuth *et al.*, 2003; Zhou *et al.*, 2003). Adapted versions of the CANEGRO model have been used in yield forecast applications in Swaziland (McGlinchey, 1999) and Thailand (Promburom *et al.*, 2001).

CANEGRO simulates plant, atmospheric and soil properties with relatively high levels of complexity on a daily time step. Different mechanistic processes, such as radiation driven biomass accumulation, foliage development, canopy expansion and soil water movement are modelled (Bezuidenhout, 2000). The model requires a wide range of input variables, from cultivar and soil specific coefficients to daily climate and irrigation records.

Later versions of CANEGRO also generate output of sucrose yields (Singels and Bezuidenhout, 2002). Reduced sugar and fibre content are not explicitly simulated. Singels and Bezuidenhout (2002) compared model output with observed data from a wide range of experiments situated within the South African sugar industry. They reported root mean square errors (*RMSE*) of 5.48 $t.ha^{-1}$ and 2.60 $t.ha^{-1}$ and R^2 values of 0.82 and 0.86 for cane biomass and sucrose yield, respectively. These results equate to an approximate 82% accuracy and were based on data that were not annualised. Accuracies of the model compared well with an equivalent Australian sugarcane model, *viz.* APSIM (Keating *et al.*, 1999).

4.3 ACRU-Thompson

Thompson (1976; 1978) derived a linear relationship between evapotranspiration and sugarcane yield for the NCo376 cultivar. The Thompson equation was first imbedded into the 1984 version of the *ACRU* agrohydrological model (Schulze, 1995). *ACRU* is a daily time-step and multi-purpose physical conceptual model that deterministically simulates hydrological and crop responses to water movements through a multi-component natural or agricultural system (Schulze, 1995; Schulze and Smithers, 2004). Figure 4.1 illustrates schematically the conceptual components of the model. The *ACRU* sugarcane yield model requires fewer input data than CANEGRO and has been used extensively in regional applications of water demand and crop production (Schulze and Smithers, 2004).

Hughes (1992) evaluated the original Thompson model as imbedded in *ACRU*, for sugarcane production at both coastal and inland sites in KwaZulu-Natal. The model yielded R^2 values of 0.53 and 0.67 for the coast and inland, respectively. Model enhancements by Hughes (1992) improved these R^2 values by 35%. Schulze *et al.*

(1999) added a degree-day driven crop coefficient development routine to the *ACRU*-Thompson model during a sugarcane yield forecast study for the Eston mill in the KwaZulu-Natal midlands. Lumsden (2000) reported higher accuracies for *ACRU*-Thompson than for DSSAT-CANEGRO during the same investigation (R^2=0.78 *vs*. R^2=0.62). This was attributed mainly to CANEGRO's sensitivity to site specific input variables as opposed to the more robust simulation philosophy in *ACRU*.

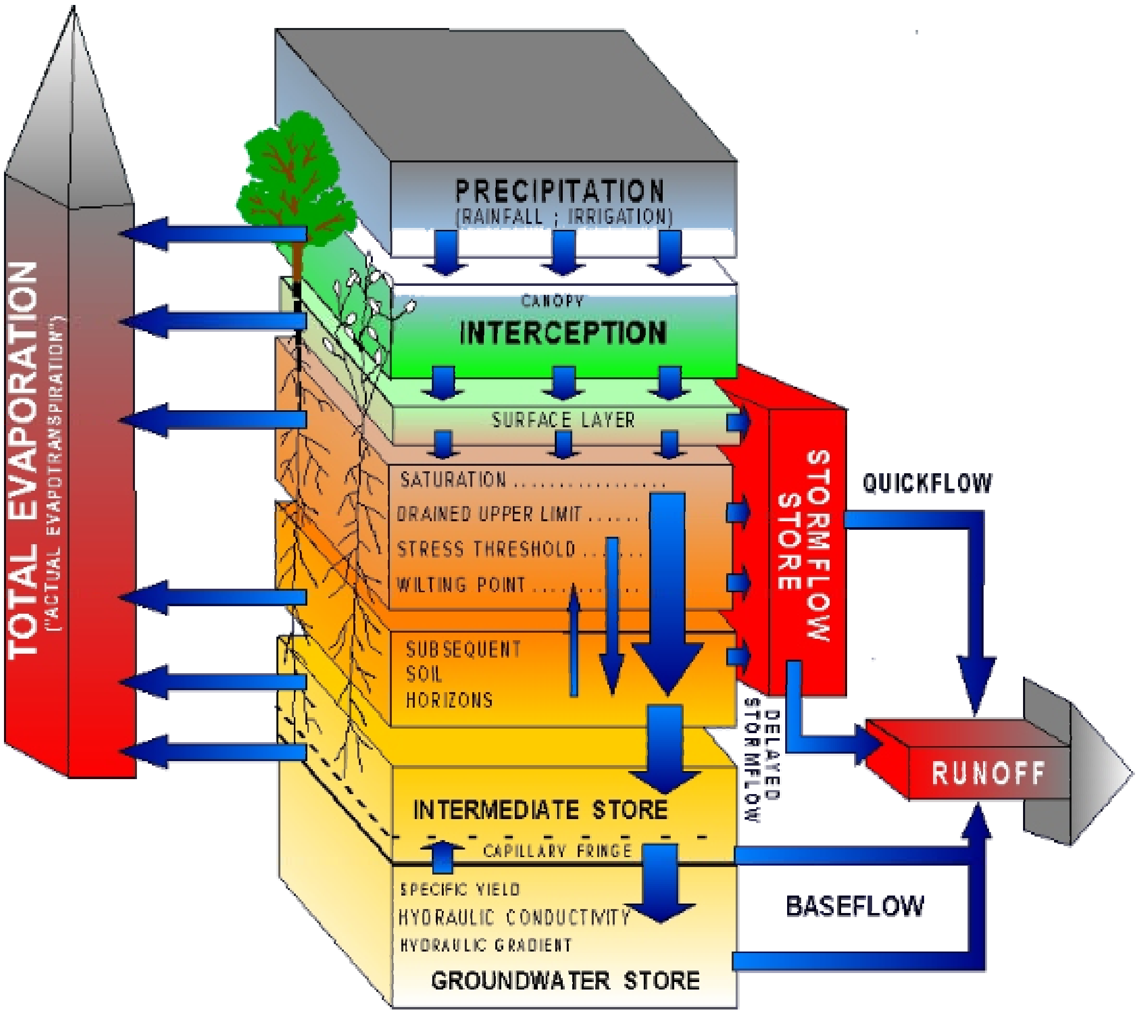

Figure 4.1 A schematic diagram of conceptual components in the ACRU agrohydrological modelling framework (from Schulze, 1995)

4.4 Canesim

The Canesim model, formerly called IRRICANE, was initially developed as a simple, user-friendly irrigation scheduling tool (Singels *et al*., 1998). Singels *et al.* (1999a) subsequently included an empirical cane yield equation, derived from the cane yield to transpiration relationship for the NCo376 cultivar in the CANEGRO model. In addition, Singels and Donaldson (2000) also incorporated a cultivar-sensitive canopy expansion algorithm. De Lange and Singels (2003) have used the Canesim model to forecast sugarcane production for the Umfolozi mill supply area in northern KwaZulu-Natal.

Canesim simulates soil water processes and plant development on a daily time step at a semi-mechanistic level of complexity. It requires daily climate data, irrigation information and a single input for the soil, *viz*. the soil's water holding capacity. Currently, the model simulates cane yield, although attempts are underway to include the sucrose yield algorithms developed by Singels and Bezuidenhout (2002).

As yet unpublished model verifications using a wide range of independently observed cane yield data display a *RMSE* of 4.88 t.ha^{-1} and a R^2 value of 0.88 (*pers comm.* Dr. A. Singels, South African Sugarcane Research Institute, Mt. Edgecombe). Figures 4.2 and 4.3 illustrate the model's accuracy when used to forecast point specific (Bezuidenhout and Singels, 2001) and regional (Gers *et al.*, 2001) cane yields, respectively. Both these comparisons yielded R^2 values of 0.87.

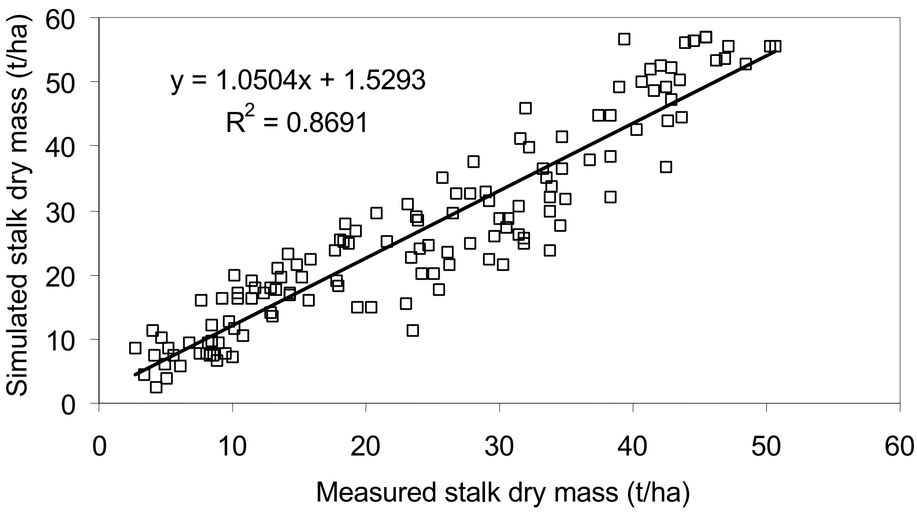

Figure 4.2 Simulated Canesim and measured cane yields from 26 crops of a single cultivar grown experimentally under widely different agronomic conditions in South Africa (from Bezuidenhout and Singels, 2001)

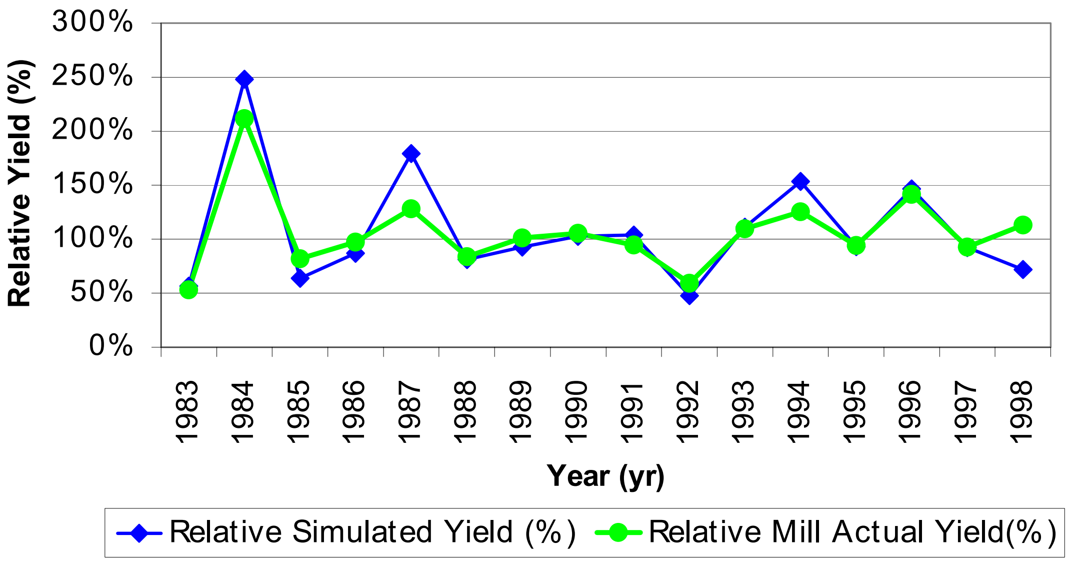

Figure 4.3 Simulated Canesim and measured cane yields (expressed as a percentage of the preceding season's yield) for the Darnall mill supply area on the KwaZulu-Natal north coast (from Gers *et al.*, 2001)

4.5 Discussion

CANEGRO is the only model currently used in the South African sugar industry for simulating sucrose yield. None of the above-mentioned models simulate fibre and reduced sugar contents or the effect of nutrients, weeds, flowering, pests and diseases on sugarcane. Unfortunately CANEGRO requires a large number of sensitive input variables, which makes it less suitable for wide-scale regional applications. In a direct comparison in the KwaZulu-Natal midlands, the *ACRU*-Thompson model produced better regional simulation results than CANEGRO (Lumsden *et al.*, 1999).

The Canesim model seems currently the most appropriate for an industry-wide yield forecasting application. Canesim has been verified at both point and regional scales with accuracies that exceed those of the *ACRU*-Thompson and CANEGRO models (Bezuidenhout and Singels, 2001; Gers *et al.*, 2001). Canesim also requires fewer input parameters and work is currently underway to include the simulation of sucrose yield by the method of Bezuidenhout and Singels (2001).

The next chapter will attempt to establish spatial simulation units for the Canesim model and to derive adequate model inputs for each of these units.

5 Homogeneous Climate Zones and Spatial Information for Sugarcane Yield Forecasting

5.1 Introduction

Chapter 3 provided guidelines on preferred formats, resolutions and methods of information transfer from yield forecasts. In addition, Chapter 4 reviewed and presented a viable sugarcane yield model for forecasting purposes. This chapter provides spatial model simulation units and derives additional information that is required for forecasting purposes.

Model applications are inherently suitable for small-scale experimental studies which are accompanied by intensive soil and atmospheric measurements. However, large-scale model applications are more meaningful, as they may address commercial production issues. The aggregation process during which model outputs are spatially extrapolated over larger areas is generally known as *up-scaling*. Commercially, crops are grown in an environment which varies both spatially and temporally and up-scaling becomes imperative owing to the high cost associated with measuring input variables at the desired density (Hansen and Jones, 1999).

King (1991, cited by Hansen and Jones, 1999) noted that up-scaling entails first, an accurate description of the landscape's spatial and temporal heterogeneity and secondly, the correct integration of this heterogeneity. Up-scaling processes may include

- The subdivision of the landscape into smaller more homogeneous simulation units or representative grid points,
- Deriving model input not measured directly from available surrogate data, and
- Stochastic sampling of input variables.

Hansen and Jones (1999) warned that new properties and processes, such as lateral water flows, competition for water allocation and inconsistent farm resource allocations, emerge at regional scales. At the same time agricultural systems at larger scales could be expected to become less sensitive to high-frequency variability (O'Neill and Deangelis, 1986, cited by Hansen and Jones, 1999; Müller, 1992).

Spatially, Supit (1997) and Thornton *et al.* (1997) used a grid approach to simulate variable crop production within a wider region. This technique is associated with raster configurations in GIS (e.g. Piwowar *et al.*, 1990), which is a suitable platform for remote sensing technologies and general circulation models (GCM, e.g. Hewitson and Crane, 1996). As an alternative to a grid approach, regions have also been subdivided into small, relatively uniform but spatially irregular simulation units, *i.e.* polygons (e.g. van Lanen *et al.*, 1992; de Jager *et al.*, 1998; Rosenthal *et al.*, 1998; Promburom *et al.*, 2001). Polygons are normally larger than grid cells and are demarcated more logically than grids, being based on topography, soil and climate characteristics.

Agronomically, Hansen and Jones (1999) noted that the use of more than one planting date has improved yield forecasts. Similarly, Lumsden *et al.* (1998) stated that a crop forecast system should be sensitive to prevailing sugarcane cropping cycles in a region. Promburom *et al.* (2001) also highlighted the importance of ensuring regionally representative soil descriptions and irrigation strategies.

The aim of this chapter is to acquire the necessary spatial information and input data to warrant the development of a crop forecast system for the South African sugar industry. Specific objectives are to

- Subdivide the South African sugarcane growing region into smaller uniform simulation units, and to
- Gather the relevant agronomic and aggregation information for each unit.

These objectives were conducted independently and are reported separately. A final conclusion of the chapter consolidates the outcomes.

5.2 Climate Regionalisation

Climate does not display well-defined natural boundaries. Nevertheless, the spatial correlation between variables within a region has formed the basis for partitioning regions into smaller, relatively homogeneous spatial units (Hansen and Jones, 1999). Hansen and Jones (2000) pointed out that even if models were completely error-free and climate data represented the absolute average conditions for a region, simulated yields would not generally represent the region’s spatial average or inter-annual variability as a result of the integration of non-linear response functions. Unfortunately, climate data frequently lacks sufficient spatial resolution (de Wit and van Keulen, 1987; Hansen and Jones, 1999; Matthews *et al.*, 2000) and often represent sites that are convenient to measure (e.g. at administrative centres, close to the office or near a dam) as opposed to sites that are more representative of the larger district (Wörlen *et al.*, 1999). Climate regionalisation may, therefore, be a valuable technique to enable generalisations to be made about areas on the bases of spatially and temporally varying parameters, such as precipitation (Comrie and Glenn, 1998). By identifying and simulating homogeneous areas, model applications are simplified and regional accuracy could be expected to increase.

Several previous studies have regionalised parts of the South African sugar producing region according to climate and soil. Welding and Havenga (1974) and Schulze (1983) subdivided South Africa and KwaZulu-Natal, respectively into rainfall zones. Dent *et al.* (1989) noted reasonable agreement between zones derived for KwaZulu-Natal in the above-mentioned two studies. Dent *et al.* (1989) continued by subdividing South Africa into 712 relatively homogeneous regions for hydrological response purposes (*cf.* Figure 5.1). Subsequent more detailed delimitations have given rise to the 712 regions now being regarded as dated (*pers comm*. Prof. R.E. Schulze, University of KwaZulu-Natal, Pietermaritzburg). Camp (1999) subdivided KwaZulu-Natal into 195 bio-resource units based on climate, soil and topography (*cf.* Figure 5.2). Although very comprehensive, Camp's study excluded the Mpumalanga province, therefore not providing a consistent approach throughout the entire area in South Africa in which sugarcane is produced. Midgley *et al.* (1994) derived the boundaries of 1947 Quaternary Catchments (QCs) in South Africa. These QCs were produced according to topographical and hydrological, rather than climate characteristics. However, good relationships may be expected between topography and climate. The QCs have formed the basis of many hydrological research studies (e.g. Kienzle *et al.*, 1997; Perks, 2001; Taylor, 2001; Chetty *et al.*, 2003), which have resulted in high quality and representative climate data for each zone.

The sugarcane producing areas extend over 118 QCs (*cf.* Figure 5.3). With regard to sugarcane related studies, the QCs were the spatial unit for yield simulations (Schulze,

1997) as well as for streamflow reduction, irrigation and water use efficiency studies (Schulze *et al.*, 1999).

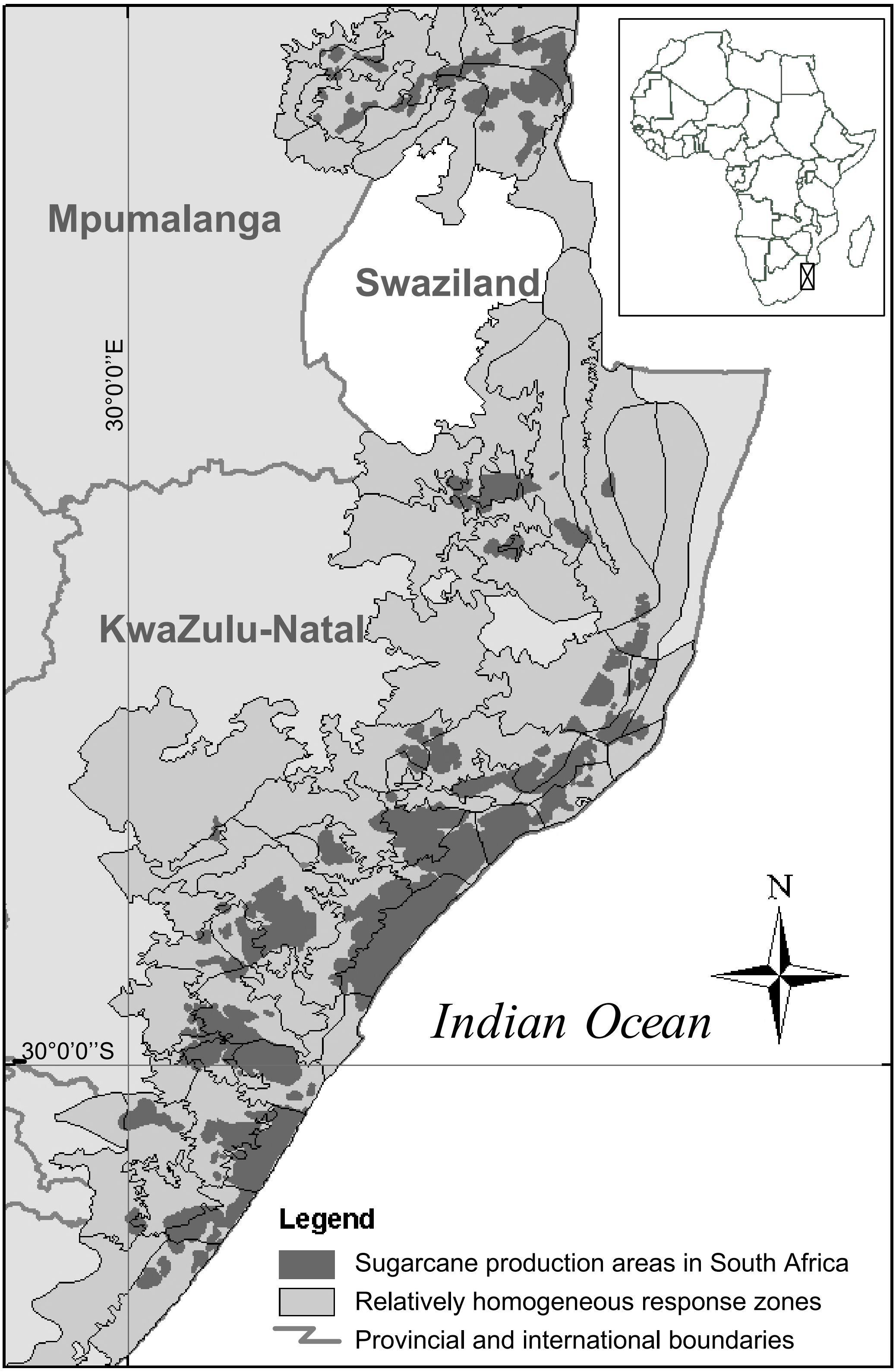

Figure 5.1 Relatively homogeneous hydrological response zones according to Dent *et al.* (1989) within the South African sugar industry

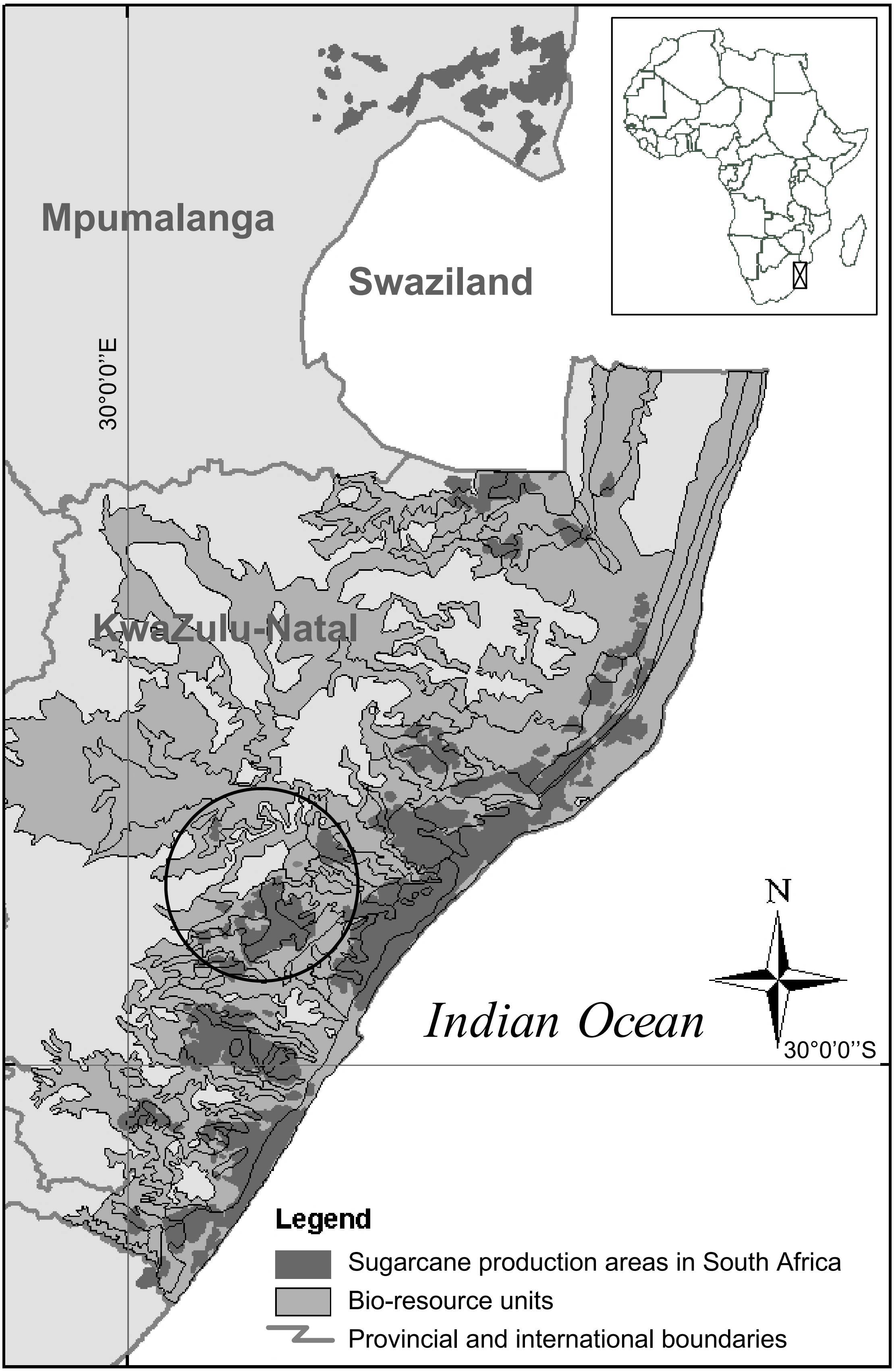

Figure 5.2 Bio-resource units based on similar soils and climates (Camp, 1999) for the sugarcane producing areas in KwaZulu-Natal. The encircled area depicts the Midlands North region, where revised homogeneous climate zones were based on these units

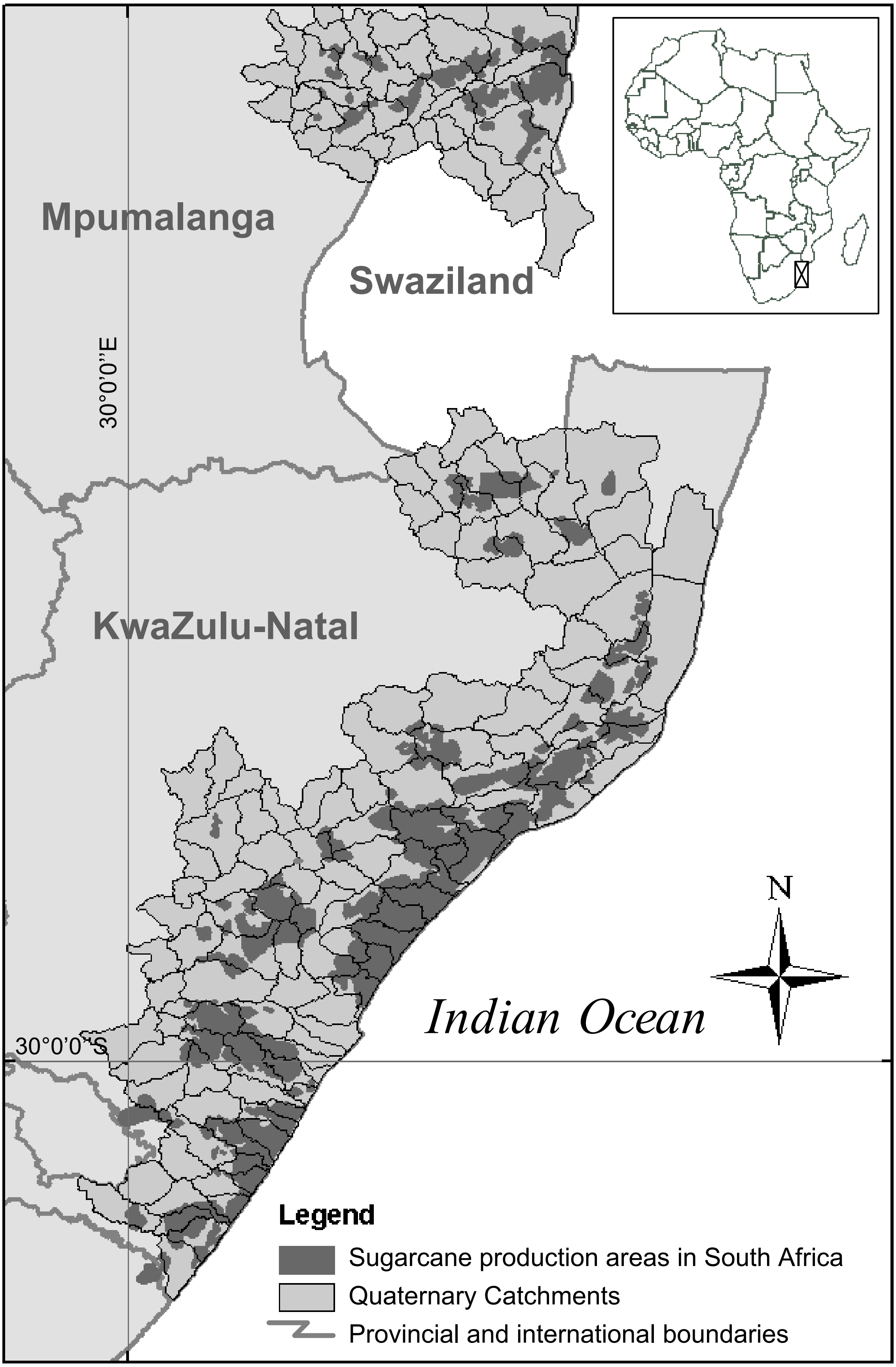

Figure 5.3 Quaternary Catchments (Midgley *et al.*, 1994) within the South African sugarcane producing areas

The aim of this section of the study is to establish reasonably homogeneous climate zones for the sugarcane producing areas of South Africa. The term *homogeneous* implies that zones should, first, be uniform within themselves and, secondly, be distinctly different from their neighbours. Specific objectives were to:

- Derive boundaries for homogeneous climate zones by assessing previously derived zones (Figures 5.1 – 5.3) and by consulting with regional extension officers, and
- Quantify the homogeneity of these zones.

Soils within the South African sugar industry are highly variable over short distances (Meyer, 1984; Meyer and Wood, 1990) and soil variability was therefore omitted from this regionalisation exercise. It was assumed more realistic to represent soils through statistical distributions within zones (e.g. Hansen and Jones, 1999).

5.2.1 Methods of Delineating New Climate Zones for the Sugarcane Belt

A template of new zones was based on the boundaries of the relatively homogeneous hydrological response zones of Dent *et al.* (1989), shown in Figure 5.1. These zones, however, were not specific to the sugar industry and were cropped to exclude those areas where sugarcane was not grown. Zones comprising altitude ranges in excess of 400 m were further subdivided. This was done to distinguish between temperature sub-zones, which may be assumed to be determined mainly by altitude. The climate zones were subsequently presented to area extension officers during an intensive series of consultations. Extension officers had the opportunity to critique and refine subdivisions within their areas of jurisdiction. In the course of refinements they were carefully guided not to base their decisions on soils and management issues. Generally, the proposed zones were confirmed, but sometimes were subdivided into smaller areas. The Midlands North (encircled area in Figure 5.2) and Mpumalanga regions were exceptions, as extension officers felt the proposed zones were inappropriate. In collaboration with the relevant extension officers, new zones were proposed for Midlands North using the KwaZulu-Natal bio-resource units (Figure 5.2, Camp, 1999) and zones in Mpumalanga were trimmed to exclude close neighbouring mountain ranges after obtaining high resolution production maps (J. Lambrechts, Transvaal Suiker Beperk Milling Company, Malelane).

The newly delineated homogeneous climate zones (HCZ) and the QCs (Figure 5.3, Midgley *et al.*, 1994) were then tested for homogeneity. The climatic uniformity within HCZs and QCs and the climatic contrast (discrimination) between neighbouring HCZs were quantified. This was done by using 1’ × 1’ (*i.e.* ~ 1.6 × 1.6 km) longitude / latitude estimates of long-term mean annual solar radiation (R_S in MJ.m^{-2}), mean annual precipitation (*MAP* in mm) and temperature (Schulze, 1997). Temperature was represented by long-term mean annual thermal units (*HU* in °C.d) with a base temperature of 10°C. Sugarcane production is known to correlate well with *HU* (Inman-Bamber, 1994; van Antwerpen, 1998), R_S and *MAP* (Inman-Bamber, 1995).

The climatic uniformity within zones was expressed as the average of the coefficients of spatial variation ($\overline{CV}$ in %) of the 1’ × 1’ *MAP*, *HU* and R_S gridded points. A zone was regarded suitably uniform when the $\overline{CV}$ of the gridded values was less than 10% (after Bezuidenhout and Gers, 2002).

The contrast between neighbouring HCZs was quantified by the mean relative discrimination of any particular zone *X* ($\overline{D_X}$ in %, Eq. 5.1). A $\overline{D_X}$ value of 0% would imply no climatic difference between a zone and its neighbours, while $\overline{D_X}$ will asymptotically approach 100% as the degree of contrast gets higher. Again, the criterion used by Bezuidenhout and Gers (2002) was adopted, which stated that a zone was suitably unique if its $\overline{D_X}$ value exceeded 18%.

$$\overline{D_X} = \frac{1}{N}\sum_{Y=1}^{N} D_{X,Y} \tag{5.1}$$

where Y (1..N) are neighbouring zones to Zone X and $D_{X,Y}$ is a discrimination index between Zone X and Zone Y (Eq. 5.2).

$$D_{X,Y} = \left[1 - \frac{\bar{\partial}_X}{\bar{\partial}_{X,Y}}\right] \times 100 \tag{5.2}$$

where $\bar{\partial}_X$ is the mean distance (Eq. 5.3) between normalised data points in Zone X and $\bar{\partial}_{X,Y}$ is the mean distance between normalised data points in Zones X and Y combined.

Mean distance between normalised data points ($\bar{\partial}$), whether contained in one zone (*i.e.* $\bar{\partial}_X$) or from two zones (*i.e.* $\bar{\partial}_{X,Y}$), were calculated using Eq. 5.3.

$$\partial_{i,j} = \sqrt{\left(MAP'_i - MAP'_j\right)^2 + \left(HU'_i - HU'_j\right)^2 + \left(R_S{'}_i - R_S{'}_j\right)^2} \tag{5.3}$$

where i and j refer to different normalised data points within one or two zones and MAP', HU' and R_S' are normalised values according to their range in Zones X and Y combined (*cf.* Eq. 5.4).

For example,

$$MAP'_i = \frac{MAP_i}{\left(MAP_{\max} - MAP_{\min}\right)} \tag{5.4}$$

where $MAP_{\max}$ and $MAP_{\min}$ are the maximum and minimum mean annual precipitation data points within Zones X and Y, respectively.

5.2.2 Results and Discussion

Forty-eight HCZs (*cf.* Table 5.1 and Figure 5.4) were derived and tested for homogeneity and 118 QCs (*cf.* Appendix B) were tested for uniformity. Testing for homogeneity included the calculation of both $\overline{CV}$ and $\overline{D_X}$ values, while uniformity testing only included the calculation of $\overline{CV}$ values. The mean $\overline{CV}$ over all HCZs was 4.5%, while the mean $\overline{CV}$ over all QCs was 6.1%. The fact that QCs subdivided the industry into a greater number of, but less uniform, zones signified the importance of developing the HCZ approach. It is also noted from Table 5.1 that no $\overline{CV}$ value exceeded 10% and that no $\overline{D_X}$ value was below 18%. All the zones were therefore concluded to be sufficiently uniform within themselves, but also adequately different from their neighbours.

The HCZs that were derived in this may have value beyond yield forecast applications. They may apply to a wide range of climate related applications and research efforts, such as extrapolating recommendations from experimental results, identifying impact zones of pests and diseases and climate change impacts research.

This study allowed for an expert based approach to delineate HCZs derived primarily from the work by Dent *et al.* (1989). Homogeneous Climate Zones were independently verified against interpolated datasets. These interpolated datasets were derived using independent variables, such as distance from the sea and altitude. It is believed that alternative regionalisation methods other than an expert based method would have been based on the same information, therefore limiting an independent verification exercise. As in any large industry, land use within the sugarcane production areas is not uniform and deep, inaccessible valleys as well as urban areas and timber plantations often fragment the sugarcane belt in South Africa. Such areas were automatically excluded by the extension officers. Uniformity of areas under cane within zones may, therefore, be higher than that established during the assessment. The author believes that this approach could only be superseded if detailed land use maps and higher resolution climate data were available.

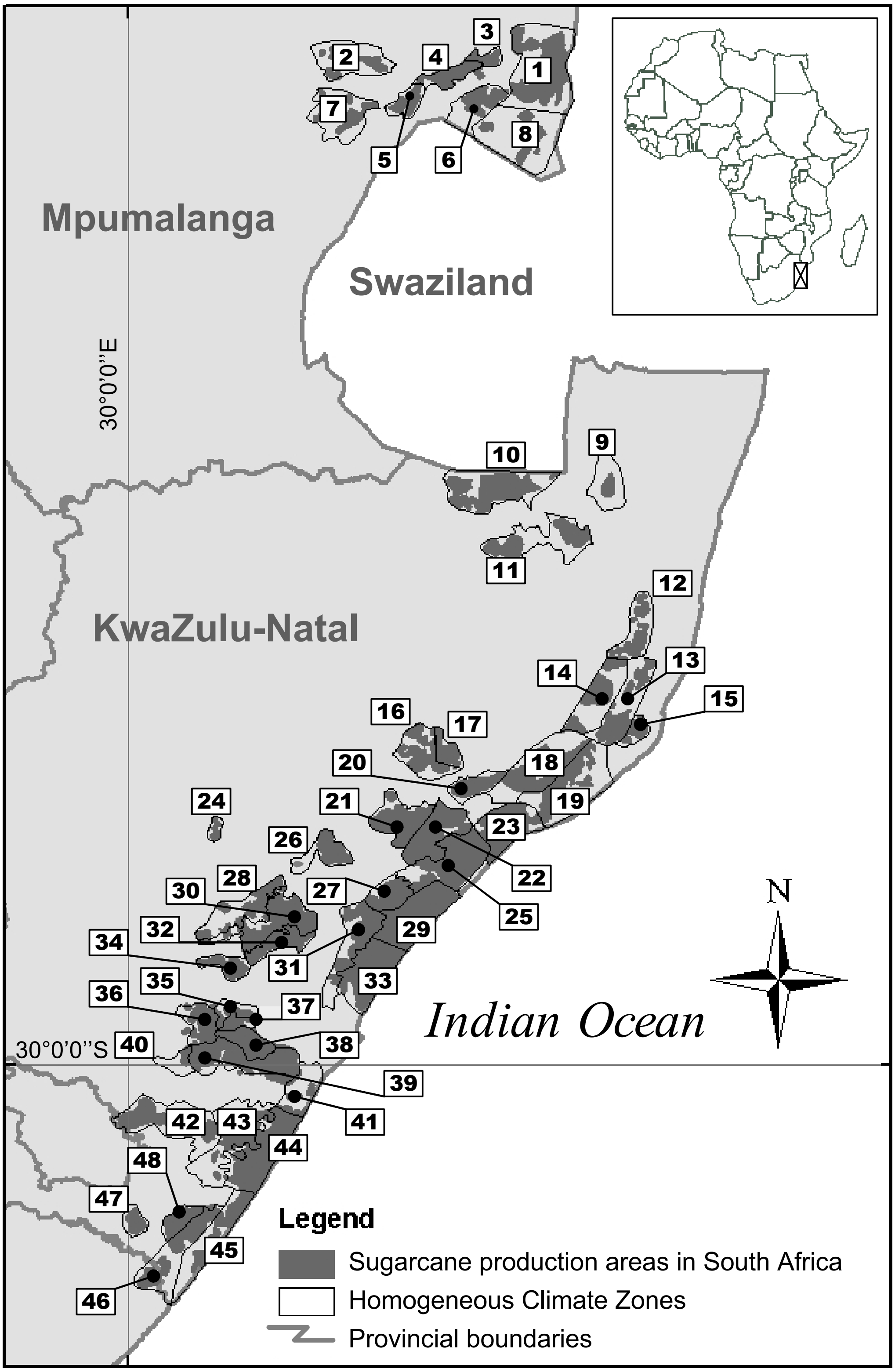

Figure 5.4 Homogeneous Climate Zones for the South African sugar industry, with zone numbers corresponding with those in Table 5.1

Table 5.1 Homogeneous Climate Zones for the sugarcane producing areas of South Africa, their areas, the spatial means of gridded values of annual thermal time ($\overline{HU}$), precipitation ($\overline{MAP}$) and solar radiation ($\overline{R_S}$), mean coefficient of variance ($\overline{CV}$), mean relative discrimination ($\overline{D_X}$), mean plant available soil water holding capacity (*AWC*) and mean age of cane at harvest

No.	Name	Area (km^2)	[1] $\overline{HU}$ (°C.d)	$\overline{MAP}$ (mm)	$\overline{R_S}$ ($MJ.m^{-2}$)	$\overline{CV}$ (%)	$\overline{D_X}$ (%)	[3] *AWC* (mm)	[3] Age at Harvest (months)
1	Komati	1150	4457	584	8599	4.1	42.3	(50)[2]	(12)[2]
2	Nelspruit	523	3493	798	8442	5.3	49.8	(50)	(12)
3	Hectorspruit	99	4310	647	8602	4.2	49.0	(50)	(12)
4	Malelane / Kaapmuiden	198	4130	734	8568	6.0	52.5	(50)	(12)
5	Louws Creek	181	3854	757	8504	5.9	47.8	(50)	(12)
6	Kaalrug / Inala	378	4207	814	8558	7.4	39.3	(50)	(12)
7	Barberton	554	3435	783	8555	6.4	31.6	(50)	(12)
8	Komati Projects	937	4219	671	8475	4.7	35.4	(50)	(12)
9	Makatini Flats	335	4419	576	8208	2.8	49.0	(50)	(13)
10	Pongola	767	4151	627	8496	7.2	45.1	(50)	(12)
11	Mkuzi	611	4145	624	8308	4.9	41.8	(50)	(12)
12	Hluhluwe	364	4348	743	7809	2.8	55.9	100 (50)	13 (12)
13	Mtubatuba	541	4312	948	7655	2.9	61.0	100 (60)	13 (12)
14	Mzondeni Strip	517	4267	843	7854	1.8	58.4	100 (60)	13 (12)
15	Umfolozi Flood Plains	121	4317	1142	7503	2.3	61.8	110 (60)	13 (12)
16	Melmoth Mist Belt	419	2916	938	7776	6.2	65.4	104	19
17	Melmoth Hinterland	154	3255	873	7872	4.0	52.1	94	19
18	Heatonville	746	4117	956	7631	3.4	62.5	110 (50)	13 (12)
19	Empangeni	827	4180	1168	7384	4.2	61.4	110 (60)	14 (12)
20	Zululand River Valley	223	3978	792	7696	2.9	73.2	(60)	(12)
21	Entumeni	427	3324	937	7798	7.6	60.3	92	17
22	Eshowe	474	3541	1079	7530	3.9	69.9	92	14
23	Emoyeni	485	3975	1234	7303	4.2	62.9	85	15
24	Muden	62	3068	639	8607	5.8	91.2	(60)	(13)
25	Amatikulu	663	3919	1018	7350	3.5	57.8	75 (60)	13 (12)
26	Kranskop Mist Belt	278	2505	920	7663	5.9	76.0	104	23
27	Doornkop	257	3473	1047	7416	5.0	58.9	104	14
28	New Hanover	458	2764	983	8095	7.8	62.0	110	23
29	Upper North Coast	616	3830	996	7239	2.4	55.2	85 (60)	14 (12)
30	Wartburg / Fawnleas	561	2769	870	7884	4.9	59.0	94	23
31	Upper Tongaat	525	3429	955	7371	5.3	59.9	112	14
32	Windy Hill Mist Belt	244	2679	985	7728	4.6	56.5	104	23
33	Lower North Coast	586	3750	974	7206	2.2	50.8	85 (60)	14 (12)
34	Hilton / Umgeni Valley	174	2967	895	8002	7.0	59.3	97 (60)	(12)
35	Umlaas Road	139	2908	786	7850	2.9	55.1	92	23 (12)
36	Bainsfield / Richmond	260	2675	891	7892	4.8	55.4	127	18
37	Tala Valley	86	2986	722	7770	3.7	59.4	(60)	(12)
38	Eston	224	2894	827	7586	3.7	53.3	92	23
39	Mid Illovo	572	3027	846	7570	6.0	59.4	125	18
40	Umkomaas	130	3316	774	8066	5.3	70.7	(60)	(12)
41	Illovo	315	3655	969	7106	2.1	65.7	64	14
42	High Flats	730	2584	817	7681	5.6	70.5	84	20
43	Dumisa	708	3307	883	7435	4.6	62.0	72	20
44	Sezela	668	3664	980	7269	3.9	53.4	64	14
45	Umzimkulu Coastal	527	3597	1079	6920	2.7	57.2	74	14
46	Paddock / North Bank	557	3279	950	7058	6.1	53.9	82	22
47	Hlaku / Nqabeni	125	2956	776	7528	2.2	67.3	100	23
48	Oribi / North Paddock	162	3134	877	7239	5.6	56.7	65	23
	Mean	**430**	**3552**	**869**	**7805**	**4.5**	**57.4**		

[1] Thermal time was calculated with a base daily mean temperature of 10°C
[2] Values in brackets refer to irrigated crops
[3] *AWC* and age at harvest are discussed in Section 5.3.2

5.3 Soil and Crop Management Information for Homogeneous Climate Zones

5.3.1 An Introduction to the Spatial Collation of Soil and Crop Management Information

The Canesim model requires a value for *AWC* and information on irrigation scheduling as well as crop start and end dates. In addition, information on relative areas under certain management conditions, as well as areas delivering sugarcane to certain mills is required for aggregation purposes. De Jager *et al.* (1998), Lumsden *et al.* (1998) and Hansen and Jones (1999) point out a wide range of issues that need to be considered when information is collated to represent and aggregate regional agricultural production. These include the following:

- Collecting representative model input data for large areas may be costly.
- Spatial differences in management (e.g. cultivars and planting dates) and heterogeneity of soils need to be captured in an attempt to avoid aggregation bias.
- Relative areas of crops under different management (e.g. rainfed *vs*. irrigated) need to be established.
- Initial soil water content may be based on the end result of a soil water balance simulation prior to the planting date.
- Simulations based on certain input parameters (e.g. *AWC* and rooting depth) may be calibrated against historical regional production to empirically improve the model's accuracy.
- Regional model applications may often treat individual soil map units as being homogeneous, describing the soil by a single set of parameters. This is done even though soils within individual map units can often vary remarkably (Reybold and TeSelle, 1989) and their model input parameter values are often not normally distributed (e.g. Young *et al.*, 1998).
- Soil specifications derived by experts may be sufficient for purposes of forecasting yields, especially if yield is expressed in relative terms (e.g. as a percentage of the previous season's yield).
- Simulations based on regional mean input parameters may be unrepresentative of mean regional production (e.g. Luxmoore *et al.*, 1991).

- An alternative to using mean input parameters is to perform Monte Carlo simulations using stochastically sampled statistical distributions assumed to be representative of input parameters (e.g. Shaffer, 1988; Haskett *et al.*, 1995; Hansen and Jones, 2000).

Little of the information required by Canesim exists for HCZs. Regional extension officers were therefore consulted to establish estimates of mean crop age at harvest, mean *AWC*, the percentage of total area under irrigation, typical irrigation practices and the area within each zone supplying sugarcane to a particular mill. It has been acknowledged that simulations based on average management practices and crop conditions may be an over-simplification of the real system. Hansen and Jones (1999) highlighted many shortcomings in this approach, but also mentioned the cost of establishing more representative estimates. Input variables are interdependent. For example, irrigation practices may be correlated to the *AWC*. The above-mentioned potential refinements were omitted in this study owing to the magnitude of the data gathering exercise required. The potential improvement to forecast accuracy resulting from such refinements needs further investigation, but falls outside the scope of this study.

5.3.2 Soil and Harvest Age Information

Lumsden *et al.* (1998) used estimates of mean *AWC* to represent soils over relatively large areas under sugarcane. In a similar fashion, estimates of mean *AWC* values were collated from regional extension officers for rainfed and irrigated crops within each HCZ (Table 5.1). Extension officers in irrigated regions noted that sugarcane rooting depths were often limited owing to problems such as soil compaction and water tables caused by over-irrigation. Although soils in these regions may have higher *AWC* values than those reflected in Table 5.1, values in the vicinity of 50 mm and 60 mm were assumed to be more representative of the true rooting zones of sugarcane under irrigation.

Information on typical crop harvest ages at different times of the milling season were also collected from extension officers. Crops harvested in the early season (i.e. April and May) are normally older than crops harvested towards the end of the season. This is attributed mainly to the fact that early season crops need to wait until the milling season has opened, while late season crops are often harvested slightly prematurely to avoid the milling off-season. Table 5.1 reflects the mean harvest age of cane over the milling season for each HCZ.

5.3.3 Irrigation Information

Numerous irrigation practices are followed in the South African sugar industry. Few of these practices are optimal and crops are often over- or under-irrigated. An irrigation strategy that is representative of most practices was carefully established after in-depth consultations with extension officers and irrigation experts. According to extension officers, a representative strategy would have to allow for frequent over-irrigation, especially during relatively wet periods in the north (Zones 1 – 18), but would also have to allow under-irrigation during hot and dry periods. This was represented by an assumed irrigation schedule by which 30 mm of effective irrigation would be applied once the soil moisture deficit falls below 30 mm. A minimum irrigation cycle of seven days was assumed. This would allow soils with *AWC* values of 50 mm to be over-irrigated at times, but would also allow a gradual depletion of soil moisture during hot and dry periods when the net water requirement is higher than 30 mm per week. Extension officers were satisfied that, on a large scale, the above-mentioned strategy was representative of sugarcane irrigation practices in South Africa.

In addition to the above-mentioned irrigation strategy, the effect of water restrictions during droughts was also incorporated in some areas. Historic events of water restrictions were expressed as a percentage of the non-restricted water quota. This percentage was subsequently used to reduce the 30 mm irrigation event on a pro rata basis.

Water restrictions in the Pongola area (Zone 10) were determined from river flow measurements of the South African Department of Water Affairs and Forestry. The Pongola river flow meter is situated downstream of the bulk of irrigated sugarcane in this catchment. However, assumptions emanating from the local water board could still be applied (*pers comm*. Mr. F. Cronje, Impala Water Board, Pongola). Water restrictions were assumed using a linear relationship with the river flow rate. Available water for irrigation was assumed at 100% while the flow rate exceeded 10 $m^3.s^{-1}$. Below this value, a linear reduction from 100% to 0% was assumed between 10 $m^3.s^{-1}$ and 4 $m^3.s^{-1}$.

Historic records on enforced water restrictions are available for the Mpumalanga province for the period 1994 to 2002. This region is served by various rivers and non-uniform water allocations within homogeneous climate zones may exist. No information on water restrictions prior to 1994 is available and, unlike the Pongola river, flow data could not be used to estimate water restrictions. No water restrictions were consequently assumed before 1994.

5.3.4 Sugarcane Delivery Information

Table 5.2 A summary of mills in the South African sugar industry and a list of Homogeneous Climate Zones (HCZs) supplying sugarcane to each mill, with mill numbers and HCZs corresponding with those in Figure 6.1

No.	Mill name	HCZ
1	Komati	1, 8
2	Malelane	1, 2, 3, 4, 5, 6, 7
3	Pongola	9, 10, 11
4	Umfolozi	12, 13, 14, 15, 18
5	Entumeni	21, 22
6	Felixton	11, 18, 19, 20, 23
7	Amatikulu	16, 17, 22, 23, 25
8	Darnall	25, 27, 29
9	Gledhow	27, 29, 31, 33
10	Union Co-op	24, 26, 28, 30, 32, 34, 36
11	Noodsberg	24, 26, 28, 30, 32, 34
12	Maidstone	31, 33
13	Eston	35, 36, 37, 38, 39, 40
14	Sezela	41, 42, 43, 44
15	Umzimkulu	45, 46, 47, 48

Estimates of annual areas harvested within each HCZ and allocations of these areas to different mills in the vicinity were established from local extension officers and historic harvesting records. Areas within each HCZ were also subdivided into irrigated and rainfed crops. Table 5.2 summarises the allocation of HCZs to different mills in the industry. It is noteworthy that some HCZs deliver sugarcane to more than one mill.

5.4 Discussion and Conclusions

The South African sugarcane production areas were subdivided into 48 homogeneous climate zones to provide simulation units. Climatically, all zones displayed less than 10% internal variability, but differed by more than 18% from neighbouring zones. The new homogeneous climate zones were allocated to traditional mill supply areas to enable the aggregation of simulation results to mill scales.

Additional information on mean cane age at harvest and mean total plant available soil water holding capacity were collated for each homogeneous climate zone. Likewise, a single, but versatile, irrigation strategy was assumed for all sugarcane crops under irrigation. The use of means, as opposed to statistical distributions for input variables, may, however, over-simplify model results (*cf.* Hansen and Jones, 1999) and although it did not form part of this study, further work in this area should be encouraged. Information on irrigation water restrictions was included to temporarily adjust irrigation strategies in certain areas. This information is still limited because of simple conversion assumptions based on streamflow measurements and limited lengths of actual records. It may be expected that attempts to simulate river flow and water use could enhance the accuracy of simulations in irrigated areas.

Further to the above-mentioned information, the Canesim model also required daily climate data for each homogeneous climate zone. The climate data are discussed in Chapter 6.

6 Climate Data for Homogeneous Climate Zones

6.1 Introduction

The previous two chapters provided a crop yield model and homogeneous climate zones (HCZ) for sugarcane yield forecasting purposes in the South African sugar industry. In addition, agronomic and aggregation information was gathered for each HCZ. In Chapter 5 suitable climate data for model simulations were neither assessed nor collated. The aim of this chapter was to review the origin of different climate datasets, assess the usefulness of these datasets for yield forecasting purposes and collate suitable climate data for each HCZ.

Horie *et al.* (1992) successfully implemented a regional yield forecast system for rice in Japan using climate data from 860 stations. While capturing spatial and temporal climate variability to this level of detail is ideal, Matthews *et al.* (2000) warn that such intensive networks of climate stations are seldom available in developing countries. Alternative methods of capturing climate variability within crop production regions may include the use of interpolation techniques, meso-scale General Circulation Models (GCMs, e.g. Stern and Easterling, 1999; Eakin, 2000; Stone *et al.*, 2000) and remote sensing (e.g. Anon., 1996; Ba and Nicholson, 2001; Dyson *et al.*, 2002).

Point-scale climate data for individual stations should be accurate, representative of a larger HCZ and cover a recent period in time before reliable operational yield forecasts can be established. Many research studies (e.g. Horie *et al.*, 1992; Lumsden *et al.*, 1999) have stressed the importance to use the most recent climate data, since the alternative, *viz.* climate outlooks, are normally expressed in probabilistic terms and often have low forecast skills. High quality climate data, nevertheless, remain a considerable limitation in regional model applications (Hansen and Jones, 1999; Liu and Scott, 2001).

The objectives of this chapter are, first, to review different sources of climate data available to the South African sugar industry and, secondly, to compile datasets of suitable climate data for the Canesim model for the different Homogeneous Climate Zones (HCZ) over the period 1978 to 2002.

6.2 A Review of Climate Data Options Available to the South African Sugar Industry

Several methods of collecting, emulating and enhancing climate data exist. These include the use of

- Date from climate and rainfall stations,
- Empirical interpolation and substitution methods,
- Stochastic weather generators,
- GCM down-scaling and
- Remote sensing.

Each of these methods offer solutions to different problems, but also have inherit limitations. This introduction briefly examines the different climate data options available to the South African sugar industry.

6.2.1 Use of Data from Climate and Rainfall Stations

The South African Sugarcane Research Institute (SASRI) has a well established network of manual and automated climate and rainfall stations (e.g. Inman-Bamber, 1995; Singels *et al.*, 1999b). In most cases, daily climate data are processed on a monthly basis, causing a lag of up to six weeks before quality controlled data are available for modelling purposes (Bezuidenhout and Singels, 2001). In addition, other organisations such as the South African Weather Service and the Agricultural Research Council, as well as many smaller private corporations, also record climate data within, or in close proximity to, sugar producing areas. The School of Bioresources Engineering and Environmental Hydrology (BEEH) at the University of KwaZulu-Natal in Pietermaritzburg has collated temperature and rainfall data from all the above-mentioned organisations into the single largest and most representative climate database for South Africa (Lynch, 2004; Schulze and Maharaj, 2004). Extensive quality control has been performed and missing data infilling techniques applied (e.g. Smithers and Schulze, 2000; Lynch, 2004; Schulze and Maharaj, 2004) to enhance data for applications in models.

Climate station data usually form the basis for regional model applications, but may often be limited by their point-scale representivity, calibration inconsistencies and missing data (Downing and Washington, 1997; Hunt *et al.*, 1998; Hansen and Jones, 1999). Point-scale rainfall measurements may, in particular, be misrepresentative of a larger area. Boughton (1981) reported measurement errors of up to 20%, while Hansen and Jones (1999) and Wilby and Wigley (2000) indicated that, over an area with a sparse network of stations and especially in summer rainfall areas with strong convective activity, the direct use of rainfall station data was likely to underestimate magnitudes and also frequencies of individual events. This may result in incorrect estimates of soil moisture content (de Wit and van Keulen, 1987; Hansen and Jones, 1999), which affects the simulation accuracy of crop yields. A proposed method to alleviate the problem would be to increase the number of rainfall recording sites and to perform multiple simulations within each HCZ. Other climate parameters, such as temperature and solar radiation, may be assumed to be spatially more uniform than rainfall and a one station could be used to represent the whole HCZ.

6.2.2 The Application of Empirical Interpolation and Substitution Methods

Interpolation of climate values between stations with data and the use of surrogate values derived from other observed climate parameters present favourable opportunities to fill in missing data, both spatially and temporally. Spatial interpolation can be achieved through techniques such as principal components analyses (e.g. Boyer and Feldhake, 1994) and geostatistical methods (e.g. Bland and Clayton, 1994; Söderström and Magnusson, 1995). Hansen and Jones (1999) and Lynch (2004) concluded that inappropriate results were achieved during attempts to spatially interpolate daily rainfall data. However, while assessing mean and inter-annual variability in simulated yields, Hansen and Jones (1999) concluded that simulations based on interpolated climate values might be more representative in assessing regional crop yields. Schulze and Maharaj (2004) have interpolated daily temperature values over South Africa for a 50 year period to a 1’ × 1’ latitude × longitude resolution (*i.e.* on a grid of ~ 1.6 km × 1.6 km).

Techniques to calculate surrogate values for missing data have been developed for rainfall, temperature, solar radiation and evapotranspiration. These techniques intercalate data that are temporally fragmented. Lumsden *et al.* (1998) and Smithers and Schulze (2000) successfully filled in missing rainfall data using inverse distance weighting and other statistical methods, respectively. Wörlen *et al.* (1999) and Schulze and Maharaj (2004) discuss various issues such as adiabatic gradients and katabatic temperature drainage, as well as the role of cloud cover, that need to be considered when attempting to fill in missing temperature data. Bristow and Campbell (1984), Hunt *et al.* (1998) and Liu and Scott (2001), among others, assessed the use of empirical routines for estimating missing solar radiation data, using substitution from nearby stations and diurnal temperature values as main input parameter. Hunt *et al.* (1998) and Zelenka *et al.* (1998) identified a non-linear decline in accuracy when data from stations with increasing distance were used as substitutes. Liu and Scott (2001) emphasised the importance of using neighbouring stations in similar climatic zones for substituting solar radiation and also noted distinct differences in the dynamics at coastal sites as opposed to those at inland sites. Hargreaves and Samani (H&S, 1985) and Linacre (1991) derived routines to estimate reference potential evaporation using observed temperature, calculated extraterrestrial solar radiation, latitude and altitude as inputs. The H&S equation has received significant recognition, including support by the FAO (Allen *et al.*, 1998).

6.2.3 Emulating Climate Variables using Stochastic Weather Generators

Stochastic weather generators (e.g. Richardson, 1981; Zucchini and Adamson, 1984; Richardson, 1985; Hansen, 1999) can emulate climate statistics at a specific location and output will include typical frequencies of short and long-term dry and wet, as well as cold and hot spells. These tools are valuable for long-term model applications on risk assessments and benchmarking, but are unsuitable when specific historic periods need to be simulated (Sharpley and Williams, 1990; Schulze, 1995; Liu and Scott, 2001).

6.2.4 Use of Downscaled Results from General Circulation Models

Eakin (2000) and Stone *et al.* (2000) used GCM output to generate climate information for agricultural needs. GCMs have coarse spatial resolutions and perform reasonably well in simulating inter-annual climate variability at sub-continental scales (Bates *et al.*, 2000). To enhance the usefulness of GCM output, meso-scale atmospheric models and empirical downscaling techniques have been developed to estimate regional and local climate values (Wilby *et al.*, 1998; Stern and Easterling, 1999; Bates *et al.*, 2000). These techniques may rely on regression analyses, canonical correlation analyses, statistical analogues, artificial neural networks and the use of topography and weather classification schemes (Hewitson and Crane, 1996; Bates *et al.*, 2000). Unfortunately, as noted by Mearns *et al.* (1995, cited by Hansen and Jones, 1999), GCMs tend to over-predict rainfall frequency and under-predict rainfall intensity. Hewitson and Crane (1996) and Wilby and Wigley (2000) have identified fundamental limitations to downscaling techniques, which include assumptions that:

- Regional climate is exclusively driven by synoptic scale systems,
- GCMs provide enough variables at the right resolution to represent synoptic systems and
- Statistical relationships used for downscaling do not change over time (e.g. under conditions of climate change).

6.2.5 Remote Sensing Applications for Producing Climate Data

Remote sensing techniques have been used to estimate rainfall, solar radiation and evapotranspiration. Meteosat-based observations of Cold Cloud Duration (CCD) have been used to identify wet spells over large areas. The observations are, however, unsuitable for establishing daily rainfall patterns at resolutions less than 50 km^2 (Anon., 1996). Arkin and Meisner (1987) and Xie and Arkin (1997) developed improved techniques to estimate rainfall from GOES satellite observations of CCD. These routines did not perform well when thick cirrus clouds were present and when precipitation originated from warm clouds (Xie and Arkin, 1997). Nearly 70% of sugarcane production in South Africa occurs in close proximity to the KwaZulu-Natal coastline, which is an area synonymous with warm cloud precipitation. Therefore, it was assumed that GOES-based observations of CCD may not reflect actual rainfall satisfactorily within the study area. Another method of remotely sensing precipitation is through the use of radar. While a network covering 70% of South Africa exists (Banitz, 2001), only radar data collected at the Durban International Airport (29°59’5’’S, 30°57’30’’E) are applicable to the sugar producing areas in the country (van Heerden and Steyn, 1999). Similar to problems in CCD, radar observations are limited when rainfall originates from low coastal clouds on the KwaZulu-Natal coastline (*pers comm.* Dr. D.E. Terreblanche, South African Weather Service, Bethlehem).

Zelenka *et al.* (1998) and Roebeling *et al.* (1999) established methods of deriving estimates of solar radiation using satellite data. However, these methods are laborious and require high levels of calibration and image refinement and analyses before data can be made available for crop modelling purposes (*pers comm.* Dr. A. Zelenka, Swiss Meteorological Institute, Zürich, 2001). Roebeling *et al.* (1999) have also derived a method of estimating evapotranspiration from Meteosat data.

6.3 The Derivation of Climate Data for Crop Yield Modelling Purposes

6.3.1 Background

Two comprehensive climate datasets that may be suitable for yield forecasting purposes were identified. The first, named the BEEH climate dataset, was compiled using the rainfall and temperature databases housed at the University of KwaZulu-Natal (*cf.* Section 6.2.1). Although this dataset could be expected to contain data from a large network of climate stations in the sugarcane belt, lack of funding has prevented the BEEH database from being updated beyond 2000. Data from climate stations managed by SASRI (*cf.* Section 6.2.1), in contrast, are frequently updated as they are used in a wide range of operational applications (e.g. Singels *et al.*, 1999b). A second dataset, named the SASRI climate dataset, therefore contained data originating exclusively from the climate stations managed by SASRI. It was assumed that, because of its need to be available in near real-time for forecasting purposes, an operational yield forecast system would be based on the SASRI climate dataset. The BEEH dataset may, however, be used to point out limitations in the current SASRI climate station network.

6.3.2 The BEEH Climate Dataset

A single representative data record was derived from the BEEH climate database for each HCZ from January 1978 to December 1999. This was based on the most centrally located rainfall station and a spatially interpolated centroid temperature record. A BEEH climate record therefore did not reflect climate at any specific geographical point. All records were complete since infilling techniques had been applied prior to the data extraction (Lynch, 2004; Schulze and Maharaj, 2004).

The Canesim yield model requires reference potential evapotranspiration for sugarcane (EC_{ref} in mm.d^{-1}), which is defined as the daily evaporative demand for a three metre high sugarcane canopy under no water stress. This variable could not be calculated by the conventionally used Penman-Monteith method as applied by McGlinchey and Inman-Bamber (1996) since no estimates of solar radiation and wind speed were available in the BEEH climate database. The H&S equation (Eq. 6.1, Hargreaves and Samani, 1985; Allen *et al.*, 1998) and that by Linacre (Eq. 6.2, Linacre, 1991; Schulze, 1995), provide alternative ways of estimating reference evapotranspiration (ET_O in mm.d^{-1}). In the H&S equation

$$ET_O = 0.0023(T_{mean} + 17.8)(T_{mx} - T_{mn})^{0.5} R_a \qquad (6.1)$$

where R_a is extraterrestrial solar radiation (in MJ.m^{-2}.d^{-1}) and T_{mean}, T_{mx} and T_{mn} (in °C) are the daily mean, maximum and minimum temperatures, respectively. In the Linacre (1991) equation

$$ET_O = \left[0.015 + 4\times10^{-4} T_{mean} + 10^{-6} z\right] \times \left[\frac{380(T_{mean} + 0.006z)}{84 - \phi} - 40 + 4u(T_{mean} - T_{dew})\right] \qquad (6.2)$$

where z is altitude above sea level (in m), ϕ is latitude (in °, with positive values indicating the Northern Hemisphere), u is mean wind speed at 2 m (assumed constant at 2 m.s^{-1}) and

$$T_{mean} - T_{dew} = 0.0023z + 0.37T_{mean} + 0.53(T_{mx} - T_{mn}) + 0.35T_{ra} - 10.9 \qquad (6.3)$$

where T_{dew} is dew point temperature (in °C) and T_{ra} (in °C) is the range between long-term mean air temperature of the hottest and coldest months in the year.

Allen *et al.* (1998) noted that the H&S equation may need a linear adjustment at different locations. Likewise, an adjustment was required to convert potential evaporation to reference sugarcane evapotranspiration. It should be acknowledged that non-linear relationships exist when vegetation height changes via the theory of zero plane displacement (Calder, 1992; Shuttleworth, 1992). However, under fixed canopy height specifications, such as assumed for EC_{ref}, a linear relationship between potential grass and sugarcane evapotranpiration was assumed. Schulze *et al.* (1999)

derived adjustments to the Linacre equation in the sugarcane belt of South Africa. For this study data from 15 Automatic Weather Stations (AWS) situated in the sugarcane belt (*cf.* Table 6.2) were used to relate ET_O values derived by the H&S and Linacre equations to EC_{ref} values derived by the Penman-Monteith equation (McGlinchey and Inman-Bamber, 1996). A cross validation procedure was performed at each site. This consisted of

- an independent regression fit between the results of the H&S and Linacre equations and EC_{ref} using data records from 14 sites and
- subsequently performing a verification at the remaining site using the afore-mentioned regression coefficients.

A *RMSE* and bias error (in mm.d^{-1}) were calculated for each site (*cf.* Table 6.2).

6.3.3 The SASRI Climate Dataset

Climate stations managed by SASRI (*cf.* Table 6.1 and Figure 6.1) generally record daily rainfall, solar radiation (or sunshine hours), relative humidity at 8:00 and 14:00 (alternatively, dry and wet bulb temperatures), wind run and minimum and maximum temperatures. The methods described by Spitters *et al.* (1986) and Allen *et al.* (1998) are used to convert measurements of daily sunshine hours and dry and wet bulb temperatures to Photosynthetically Active Radiation (PAR in MJ.m^{-2}.d^{-1}) and relative humidity, respectively. Potential sugarcane reference evapotranspiration was subsequently calculated by the Penman-Monteith equation as used by McGlinchey and Inman-Bamber (1996). In addition, data from several rainfall stations within given HCZs where rainfed sugarcane was cultivated, were also included. A summary of these rainfall stations is supplied in Appendix C.

Data records from both climate stations and rainfall stations were often fragmented owing to one or more of the following reasons:

- Recordings at climate sites were discontinued,
- Theft and technical problems with instrumentation caused certain parameters to be incorrect or missing,
- Manual measurements were taken on weekdays (Monday – Friday) only, and
- Records for an entire month went missing in the mail.

Table 6.1 A summary of climate stations managed by the South African Sugarcane Research Institute at different locations in South Africa. Site numbers coincide with those in Figure 6.1 and provide approximate locations of each station. In some cases (e.g. Site 4), more than one station existed in close proximity to each other

Station name	Site	Station details (comment)
Tenbosch	1	1978 – 1995
Komati (AWS)	2	1996 – 2002 (data problems in 1997)
Amaxala (AWS)	2	2001 – 2002
Mhlati	3	1978 – 1998 & 2000 – 2002 (AWS)
Kaalrug	4	1978 – 1993 (data often fragmented)
Inala (AWS)	4	2000 – 2002
Makatini	5	1978 – 1999 (data often fragmented)
Pongola	6	1978 – 2002 & 1997 – 2002 (AWS)
Glenpark	7	1978 – 1996 (few solar radiation data) & 1997 – 2002 (AWS)
Riverview	8	1978 – 2002 (solar radiation terminated in 1992)
Monzi (AWS)	9	1997 – 2002
Dangu (AWS)	8	2000 – 2002
Entumeni	10	1978 – 2002
Mtunzini	11	1978 – 1999
Felixton	12	1987 – 2002
Heatonville (AWS)	13	1998 – 2002
Amatikulu	14	1998 – 2002
Doornkop	15	1978 – 1997
Glendale	16	1978 – 2002
Seven Oaks	17	1978 – 2002
Jaagbaan	18	1978 – 2002
Bruyns Hill (AWS)	19	1997 – 2002
Darnall	20	1978 – 2002
Gledhow	21	1996 – 2002
Tongaat	22	1978 – 2002
Mt. Edgecombe	23	1978 – 2002 (manual & AWS)
Crammond	24	1978 – 2002 (no wind speed)
Powerscourt	25	1978 – 1996
Umbumbulu (AWS)	25	1997 – 2002 (long period with missing data: 1999 – 2001)
Eston (AWS)	26	1997 – 2002
Beaumont	26	1984 – 1995
Thornville	27	1984 – 1996
Richmond (AWS)	28	1997 – 2002
Esperanza	29	1978 – 2002 (solar radiation data terminated in 1995)
Sezela	30	1978 – 2002
Umzimkulu	31	1990 – 2002
Paddock	32	1983 – 1994 (relative humidity problems)

AWS: Automatic Weather Station

A complete climate record for most HCZs from 1978 to 2002 was compiled using the data from different available SASRI climate stations shown in Table 6.1. In most cases, data from more than one climate station were carefully combined according to the procedure explained below. Special attention was given to representivity and notable trends between neighbouring climate stations were first removed before data were combined. For each HCZ a day-to-day procedure was used to assess pre-selected station records. Stations were ranked and data were selected from the station with the highest rank, *i.e.* the one containing the most complete record for the day. Rainfall station data were not combined in a similar fashion, but fewer simulations were carried out when rainfall data were unavailable.

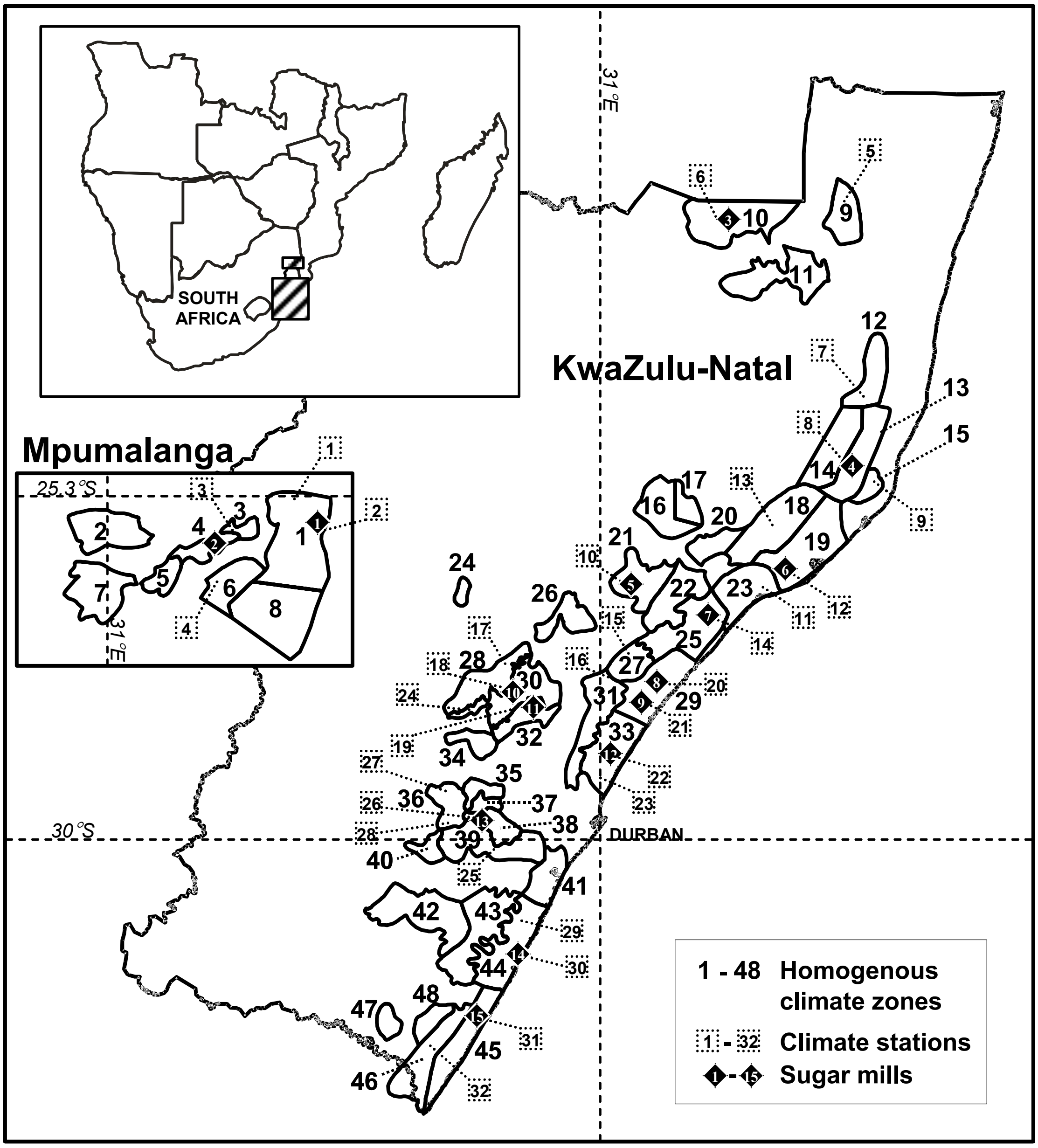

Figure 6.1 The distribution of homogeneous climate zones, climate stations and sugar mills in the South African sugarcane production areas

6.4 Results

Table 6.2 summarises results of the linear relationship (offset and slope coefficients) and cross validations that were performed to assess the suitability of the Hargreaves and Samani (H&S, 1985) and Linacre (1991) equations. Although unexplained, it is noted that the slope (*m*) and offset (*c*) coefficients deviated significantly from the 1:1 line for both equations. Root mean square error values indicate that EC_{ref} was estimated more accurately by the H&S equation at 11 of the 15 sites. The Linacre equation did, however, perform better at most inland sites. Bias errors associated with the two equations were less distinct. It was concluded that the H&S equation was generally more appropriate to estimate EC_{ref} at different locations in the South African sugar industry. The record set of each homogeneous climate zone was enhanced accordingly by applying the H&S equation in conjunction with a linear correction using values of -0.49 and 0.44 (*cf.* Table 6.2) as offset and slope coefficients, respectively.

Table 6.2 Cross validations reflecting the linear relationship (*c*=offset, *m*=slope) and independent verification results between reference sugarcane evapotranspiration (McGlinchey and Inman-Bamber, 1996) and reference short grass evapotranspiration according to Hargreaves and Samani (1985) and Linacre (1991), with site numbers corresponding with those shown in Figure 6.1

Site		Number of data points	Hargreaves and Samani (1985)					Linacre (1991)				
			Independent regression results excluding the particular site			Verification results using the independent regression (mm)		Independent regression results excluding the particular site			Verification results using the independent regression (mm)	
No.	Name		*c*	*m*	R^2	Bias	*RMSE*	*c*	*m*	R^2	Bias	*RMSE*
1	Tenbosch	918	-.44	.44	0.56	-.12	1.012	1.57	1.36	0.45	-.49	1.467
2	Amaxala	1877	-.45	.44	0.57	-.33	1.367	1.60	1.36	0.46	-.72	1.730
3	Mhlati	948	-.50	.44	0.57	.43	1.151	1.62	1.38	0.46	.17	1.419
4	Inala	754	-.48	.44	0.57	.74	1.105	1.61	1.38	0.46	.53	1.309
6	Pongola	1558	-.50	.44	0.57	-.05	1.088	0.65	1.39	0.46	-.11	1.362
7	Glenpark	2051	-.44	.44	0.57	.15	1.249	1.61	1.38	0.46	-.14	1.520
8	Dangu	1055	-.47	.44	0.57	.35	1.140	1.61	1.38	0.46	.32	1.406
9	Monzi	1713	-.47	.44	0.58	.28	1.263	1.61	1.38	0.46	.12	1.432
13	Heatonville	1785	-.48	.44	0.58	-.36	1.657	1.52	1.35	0.45	-.23	1.647
19	Bruyns Hill	1721	-.51	.44	0.59	-.55	2.007	1.63	1.38	0.46	.07	1.950
22	Tongaat	810	-.47	.44	0.58	.34	1.179	1.63	1.38	0.46	-.01	1.468
23	Mt. Edgecombe	1548	-.47	.44	0.57	-.20	1.049	1.66	1.38	0.46	-.31	1.423
25	Umbumbulu	1226	-.49	.44	0.57	.08	1.196	1.64	1.39	0.46	.23	1.349
28	Richmond	2085	-.58	.45	0.58	-.24	1.450	1.65	1.39	0.45	.13	1.347
38	Eston	1867	-.58	.45	0.59	-.12	1.504	1.66	1.40	0.46	.24	1.410
	Mean / Total	**21916**	**-.49**	**.44**	**0.58**	**.03**	**1.294**	**1.55**	**1.38**	**0.46**	**-.01**	**1.482**

Table 6.3 reflects the way in which climate records were compiled using the SASRI climate stations. Zones 10, 21 and 25 contained adequate data from a single representative climate station within the respective zone. Values for all other zones were compiled from data that were either incomplete or that originated out of neighbouring zones. No suitable data were collected by SASRI for Zones 24, 37 and

40. These zones represent sugarcane producing areas under irrigation that are situated in deep valleys in the midlands of KwaZulu-Natal.

Table 6.3 Time lines for each Homogeneous Climate Zone (HCZ), illustrating which climate stations' data were used to complete each record set, with climate station numbers corresponding with those given in Table 6.1

HCZ	'78 '80 '82 '84 '86 '88 '90 '92 '94 '96 '98 '00 '02
1	1 — 2
2	3 — 2 — 3
3	3 — 2 — 3
4	3 — 2 — 3
5	3 — 2 — 3
6	4 & 3 — 2 — 4 & 3
7	3 — 2 — 3
8	1 — 2
9	5 — 6*
10	6
11	6
12	7 & 6* — 7
13	8 — 8 & 11* — 9 — 8
14	8 — 8 & 11* — 9 — 8
15	8 — 8 & 11* — 9
16	10
17	10
18	10 — 10 & 12 — 13
19	10 — 10 & 12 — 13
20	10 — 10 & 12 — 13
21	10
22	10
23	11 — 14
24	No Data
25	14
26	10
27	15 & 16
28	17 & 18 — 17 & 19
29	20 — 20 & 16 — 21
30	18 — 18 & 19
31	16 & 15
32	18 — 19
33	22 — 19 & 23
34	24 & 18 — 24 & 19
35	25 — 26 & 27 — 26
36	25 — 26 & 27 — 28
37	No Data
38	25 — 26 — 27 — 26
39	25* — 27 — 25 — 26
40	No Data
41	29 — 29 & 30*
42	25* — 27 — 25 — 26
43	29 — 29 & 30*
44	30 & 29*
45	30 & 29* — 31 & 30 & 29*
46	25 — 32* — 25 — 31 — 25 & 26*
47	25 — 32* — 25 — 31 — 25 & 26*
48	25 — 32* — 25 — 31 — 25 & 26*

* Denotes that some parameters in the source data were linearly adjusted before adding it to the record.

6.5 Discussion and Conclusions

It may be concluded that various sources of climate data need to be evaluated and prioritised. Climate station data have relatively long records and therefore form a sound basis for long-term simulations. Empirical enhancements to climate station data, either through interpolation or substitution may, however, be more representative of larger areas. With the exception of radar-based rainfall information, GCM and satellite derived remote sensing technologies still seem more laborious and coarser in their assumptions to represent high spatial resolutions of climate and its variability (*cf.* Sections 6.2.4 & 6.2.5). These technologies should, nevertheless, not be discarded before a thorough investigation into their accuracy and representivity has been completed. That, however, falls outside the scope of this study. The use of stochastic weather generators is statistical and therefore inappropriate for forecasting yields for specific periods.

Two climate datasets for HCZs were compiled in this chapter. The first, *viz.* the BEEH climate dataset, contained rainfall and temperature records from a comprehensive database from 1978 to 1999. Reference potential evapotranspiration for sugarcane was calculated using the Hargreaves and Samani (1985) equation. The BEEH dataset has not been updated since 2000 and therefore lacks the ability to be used for operational yield forecasting purposes. The second dataset, *viz.* the SASRI climate dataset, has been kept up to date, but lacks representative climate stations in some HCZ. The above-mentioned data were obtained from climate station records that often required enhancement by interpolation and substitution techniques. Chapter 8 explains in more detail how climate data from these two sources were assessed.

In Chapter 7 climate forecasting and the adoption of seasonal rainfall outlook information for sugarcane yield forecasting are reviewed. The Canesim model-based yield forecast system for the South African sugar industry is also configured using information as described in Chapter 5, as well as the climate data collated in this chapter.

7 Seasonal Rainfall Outlooks and the Configuration of the Canesim Yield Forecast System

7.1 Introduction

A forecast of sugarcane yield relies heavily on two components. First, it is important to estimate the current status of the crop and secondly, it is important to project yields into the future using a number of probable trajectories. Remote sensing (e.g. King and Meyer-Roux, 1990; Gadekar, 1998), field scouting and climate data applied in conjunction with yield models (e.g. Lumsden, 2000; Matthews *et al.*, 2000; Promburom *et al.*, 2001) have been used to estimate the current status of the crop. In Chapters 4 and 6 a suitable South African sugarcane yield model was selected and the collection of climate data for the region was discussed, respectively. This chapter describes the configuration of crop yield simulations with the Canesim model in more detail and reports on the methods used to incorporate seasonal rainfall outlook information into these simulations.

Forecasts into the future can follow a long-term mean approach or, alternatively, additional knowledge may be used to assume future scenarios that may deviate from the long-term mean. The forecast skills of climate outlooks are likely to increase as a result of improved understanding and enhanced numerical modelling of prevailing ocean-atmosphere systems (Cane, 1999; Stern and Easterling, 1999). Various studies have shown the potential value of climate indicators and outlooks for agricultural yield forecasts (e.g. de Jager *et al.*, 1998; Jury, 1998; Singels and Bezuidenhout, 1999; Lumsden, 2000; Potgieter *et al.*, 2002). In all these cases climate uncertainty was, as can be expected, simplified to a manageable number of inputs and outputs. Researchers are likely to reduce the complexity of climate driving factors, such as trade winds and Sea Surface Temperatures (SST), to phases and indices, such as the SOI and anomalies thereof (Stone and Auliciems, 1992). Likewise, future climate scenarios are often reduced to a manageable number of discrete outcomes, such as selecting a few analogue seasons from the history of observations. These approaches are practical and valuable since they simplify a system which, to some extent, can not currently be simulated owing to its complexity. However, it is anticipated that different approaches may have different levels of success and a review of these methods is therefore presented.

Yield models used for yield forecasts, therefore, need to be configured in such a way that neither accuracy is sacrificed nor uncertainty is over- or under-estimated. Many different combinations of linking climate data, management data and future climate scenarios may be evaluated within a yield forecast system. Likewise, different seasonal climate outlooks and interpretations of these climate outlooks may influence yield forecast skills.

This chapter has two objectives. The first is to review general methodologies that form the basis of seasonal climate outlooks and to investigate the manner in which climate indices and outlooks are imported into yield forecast systems. The second is to derive a simple, but scientifically sound, configuration of input data and rainfall outlook information for a regional Canesim model-based sugarcane yield forecasting system in the South African sugar industry. It was envisaged that a thorough evaluation of a single, well contemplated configuration would illustrate certain advantages and constraints of the system developed, which could be further researched in future investigations.

7.2 A Review of Seasonal Climate Outlooks and their Adoption for Yield Forecasts

Monthly seasonal climate outlooks by the South African Weather Service (SAWS) have been compiled using a multi-faceted approach. A combination of outputs from statistical (Landman and Mason, 1999) and numerical models (Landman *et al.*, 2001) were compared and a careful expert-based synthesis was derived to produce final outlooks. Numerical model outputs include simulation results from the Climate Systems Analysis Group (CSAG, University of Cape Town, South Africa), the European Centre for Medium-Range Weather Forecasts (ECMWF, Reading, UK) and the International Research Institute for Climate Prediction (IRI, Columbia University, USA). SAWS outlooks are prepared for total rainfall and mean temperature for the ensuing three months and for three to six month lead times. Expected occurrences, derived by consensus using the combination of sources of information given above, are communicated by stating the probability of receiving above-, near- and below-normal rainfalls and temperatures for broad regions covering southern Africa.

Various other statistical climate forecasting models have been developed for southern Africa. Jury (1998) and Walker (1990), for example, derived statistical relationships between oceanic parameters and climate in South Africa. However, Downing and Washington (1997) regard statistical models as the most basic methods for making seasonal forecasts. General circulation models, which include more mechanistic descriptions of ocean-atmosphere dynamics, have become a norm for forecasting future climate (e.g. Stern and Miyakoda, 1995).

Nicholls (2000) notes that as a consequence of the increasing availability of climate forecasts from both statistical and modelling approaches, there is a necessity to adopt objective synthesising techniques. Although such techniques exist (e.g. Thompson, 1977; Winkler and Makridakis, 1983), Nicholls (2000) points out that few climate forecasters have been implementing them. In contrast, Mains (1996, cited by Nicholls, 2000) showed that subjective synthesising techniques often lead to conservative outlooks, and that individuals often did not handle dependence between different forecasts correctly. These authors suggest that consensus derived and subjectively synthesised climate outlooks, such as the SAWS outlook, may be sub-optimal.

In the South African sugar industry, Everingham *et al.* (2002b) used a case study to demonstrate the strengths and weaknesses of skillful rainfall outlooks in the context of yield forecasts. Lumsden *et al.* (1999) assessed the accuracy of SAWS outlooks by computing the frequencies of *hits* and *misses* of a forecast. Lumsden *et al.* (1999) concluded that the outlooks were accurate only 33% of the time. Potgieter *et al.* (2003), however, criticised this type of approach since climate outlooks should firstly not be categorised and secondly, should be regarded as probabilistic and therefore not be either right or wrong. Climate outlooks should not be taken as real images of the future, but as a means of statistically reducing *a priori* uncertainty (Todini, 1999).

Although climate outlook information has been adopted in numerous ways to enhance decision making, systems that link the future to likely analogue periods in the past have become increasingly utilised in agriculture (Everingham *et al.*, 2002b). This is mainly so because historical values of daily precipitation, solar radiation and temperature can be made available for model simulations. Examples of such systems have been described by Hodges *et al.* (1987), Meinke and Hammer (1997), de Jager *et* *al.* (1998), Lumsden *et al.* (1999), Hansen *et al.* (2001) and Potgieter *et al.* (2003). In these studies climate outlooks and indices, such as the SOI, were used to select more than one suitable analogue from the historical record. Model simulations based on several analogue seasons result in statistical distributions of simulated yields and economic outcomes, which facilitates a probabilistic interpretation of the forecast (Hansen, 2002). Although analogue-based yield forecast routines seem feasible, Hammer (2000b) and Antony *et al.* (2002) have emphasised the importance for researchers and climate forecasters to embark in participatory research in order to optimally enhance forecasts in the future.

7.3 System Configuration

An array of nine sugarcane crops, each harvested in a different month (April – December), was configured for a milling season. Collectively, these nine crops were subject to the typical management, soil and irrigation inputs derived for a particular HCZ in Chapter 5. Crops were of similar age and initiated in different consecutive months, thus reflecting a degree of inter-seasonal temporal variability. These arrays were configured for irrigated and rainfed crops, for each season from 1980 to 2002 and within each of the HCZs. This produced simulation configurations for 4 950 irrigated crops (9 crops × 25 zones × 22 seasons) and 6 534 rainfed crops (9 crops × 33 zones × 22 seasons).

The above-mentioned configurations were linked separately to the BEEH (1978 - 1999) and SASRI (1978 - 2002) climate datasets (*cf.* Chapter 6). For the SASRI climate dataset, additional crop arrays were configured for those HCZs in which more than one rainfall record existed (*cf.* Appendix C). In these cases temperature and evapotranspiration were assumed homogeneous within the HCZ. It should also be noted that Zones 24 (Muden), 37 (Tala Valley) and 40 (Umkomaas) had no representative climate data within the SASRI dataset and long-term mean yields of 108 $t.ha^{-1}$ were assumed for all these crops. This value was based on information gathered from regional extension officers and, based on model verifications, was inflated by 20% to include an average model bias.

The SAWS seasonal rainfall outlook was used to select either nine or ten suitable analogue seasons from each HCZ's climate history. Rainfall over a three-month period was accumulated for each season from 1980 to 1999. The three-month period coincided with the lead time of the SAWS outlook. Seasons were subsequently ranked from the driest to the wettest and were subdivided into three equal categories, namely below-normal (1st – 33rd percentile interval), near-normal (34th – 66th percentile interval) and above-normal (67th – 100th percentile interval). Thereafter, a nine or ten analogue seasons were selected from the midpoints of each category, these being at the 17th, 50th and 84th percentile, respectively. The number of seasons selected from each category coincided with the probability issued for the equivalent category in the SAWS seasonal rainfall outlook. In total, ten analogue seasons were selected when the rainfall outlook was not neutral, while only nine analogue seasons were selected with a neutral rainfall outlook. For example, if a neutral rainfall outlook was issued (33% probability of receiving above-normal, near-normal or below-normal rainfall), then the three closest analogue seasons to the tercile midpoints, *viz.* 17th percentile, the 50th percentile and the 84th percentile, respectively, would be selected. The simulation of the season was hence set to be completed along nine different anticipated future scenarios based on different sets of daily climate data from the selected analogue seasons.

The forecasted cane yield for a crop was calculated as the mean yield simulated over the nine or ten different analogue seasons. The standard deviation over these crops could be used to estimate the forecast uncertainty at the point of simulation. Field scale standard deviations can not, however, be simply extrapolated to larger production regions, such as HCZ, as a result of a strong spatial co-variation between production units (Górski and Górska, 2003). Forecast uncertainty at different scales was excluded from the scope of this research owing to the limited production data available during this study.

Everingham *et al.* (2002b) reviewed the above-mentioned routine. They argued that the use of ten analogue seasons may lead to unstable forecasts as a result of the relatively large weights attached to individual seasons. In this study, however, only 24 years of historic climate data were available, which restricted the selection of more than ten suitable analogue seasons. They also identified the fact that, while the three month lead time period is used as a determinant for analogue seasons, large proportions of crop growth may be governed by subsequent months in the analogue season, which were not assessed during the selection exercise. An increase in the number of selected analogue seasons and skillful rainfall outlooks with longer lead times may be useful in addressing these issues.

For hindcasting purposes, the simulation environment was configured in such a way that any date in the period 1978 – 2002 could be treated as the end of the record, after which data originating from analogue seasons could be used to complete the particular season. Therefore, a 22 year hindcast used for assessing the accuracy of crop forecasts, when assuming that data terminated on a specific day in the year, would result in more than 100 000 simulations ((4 950 irrigated crops + 6 534 rainfed crops) × 9 analogue assumptions). The data and simulation environment were designed in Microsoft Access®, which is a powerful data manipulation and storage platform.

7.4 Discussion and Conclusions

Generally, the use of climate outlook information for yield forecasts has been well recognised. This approach, even though some studies point to low skills and certain limitations, implies that researchers may be optimistic regarding the current and future potential use of climate forecasts. At the same time, emphasis should be laid on participatory research in an attempt to bridge shortcomings between climate outlooks and applications for yield forecasts. Climate forecasters may, for example, consider issuing lists of feasible analogue seasons which could be used in simulations. Strategically, the adoption of climate outlooks and participation in related research, even though climate outlooks still have limited skills, may prove to be a valuable long-term approach.

A configuration which emphasises diversity was used to simulate crop yields in the South African sugar industry. This allowed not only for the heterogeneous simulations of management × soil × climate combinations, but also compares a wide range of future climate scenarios by substituting future climate with climate data from pre-selected analogue seasons. Although the analogue season approach has been well recognised in various other research studies, a review by Everingham *et al.* (2002b) indicated that nine to ten analogue seasons may be too few and could cause inconsistent results between consecutive forecasts. In this study, selecting a sample of more than ten analogue seasons would, however, become problematic since the full climate record exists of 24 seasons only. Everingham *et al.* (2002b) also argued that crops are often scheduled to be harvested several months into the future, which adds uncertainty when the selection of analogue seasons is only based on a three month lead time.

The system configuration presented above is based on

- A suitably evaluated crop yield model,
- A consistent and well documented rainfall outlook and analogue selection approach and
- A logical spatial and temporal breakdown of representative crop simulations.

It is, however, necessary to evaluate the system, select suitable sub-components and assess the overall reliability and performance of the system. These routines, and the results obtained, are reported in Chapter 8.

8 An Evaluation of the Canesim Model-Based Sugarcane Yield Forecast System and its Subcomponents

8.1 Introduction

The sugarcane yield forecast system based on the Canesim model was fully described in Chapters 4 to 7. For many processes it was acknowledged that certain assumptions made during the development phase of the system required further evaluation. This chapter reports on various evaluations that were performed on the system and provides some recommendations for future research.

Promburom (2001) summarises factors that could influence accuracy when regional sugarcane model-based yield forecasts are conducted. These include inadequate and inappropriate model and aggregation algorithms, incorrect model inputs, inadequate model calibration and large errors in observed yields. Similarly, Parysow *et al.* (2000) lists four sources of uncertainty when using process-based models. These are:

- Single variable uncertainties, such as sampling errors, measurement errors and expert-opinion errors;
- Uncertainties in groups of variables, such as groups of inputs and sub-models;
- Errors associated with characterising input variable attributes, such as stochastic vs. deterministic attributes; and
- Uncontrollable sources of uncertainty (e.g. extreme events such as fire and heavy frosts).

Although it may not be an easy undertaking, Parysow *et al.* (2000) noted that ideally error analyses should be performed to identify and reduce system inaccuracies. Comprehensive error analyses, also called error budgets (Gelb *et al.*, 1974), include the identification and quantification of different sources of errors that propagate through the model to produce an error in global model output (Gertner and Guan, 1991; Parysow *et al.*, 2000). Error budgets are important in prioritising future research aimed at increasing forecast skill and they include attempts to:

- Rank inputs according to their sensitivity in influencing model output;
- Forecast uncertainty in model output as a function of uncertainty in model inputs;

- Partition the error contribution among different model inputs; and
- Provide the foundation for the optimal reduction in error, or cost associated with additional data collection.

Bannayan and Crout (1999), Peterson and Fraser (2001) and Potgieter *et al.* (2003) demonstrate the implementation of multiple simulations, using a Monte Carlo simulation approach to establish output distributions, and then link these with inputs. Variable input can be based on previous seasons (e.g. Potgieter *et al.*, 2003), random sampling (e.g. Bannayan and Crout, 1999) or specialised sampling techniques (e.g. Parysow *et al.*, 2000) and may incorporate attributes of climate, soil, crop and management.

In order to quantify forecast skill, many research studies have compared simulated yields with derivatives of actual regional production (e.g. Hansen and Jones, 1999; Roebeling *et al.*, 1999; Stephens *et al.*, 2000; Promburom *et al.*, 2001; Potgieter *et al.*, 2003). Hansen and Jones (1999) indicated that actual production information is often only available for administrative reporting districts, such as mill supply areas, as opposed to model simulation units, such as Homogeneous Climate Zones (HCZ). Potgieter *et al.* (2003) emphasise the importance of using different evaluation indices to express different types of errors. These include indices such as the Squared Mean Difference (*SMD*) and Absolute Mean Deviation (*AMD*) to reflect bias errors, and the Variance Ratio index (VR) to quantify the error in dispersion (Potgieter *et al.*, 2003). Similarly, Rice and Cochran (1984) derive a composite equation to segregate an overall model error into bias, slope and random components. Hansen and Jones (1999), for example, cite various crop simulation studies where particularly inter-annual variability was over-predicted (Mearns *et al.*, 1992; Rosenberg *et al.*, 1992; Moen *et al.*, 1994; Meinke and Hammer, 1995; Chipanshi *et al.*, 1998; Rosenthal *et al.*, 1998), while in other cases simulated yields were also found to be biased (e.g. Haskett *et al.*, 1995; Russell and van Gardingen, 1997).

The aim of this chapter is to evaluate several components of the Canesim yield forecast system and to provide recommendations for future refinements. These include:

- Establishing the value of additional raingauge data in HCZs,
- Assessing the accuracy and skill of the system using different climate datasets,
- Evaluating the representivity of climate data in different HCZs,
- Assessing the accuracy and skill of the system at different times of the year, and
- Quantifying the value of seasonal rainfall outlooks to enhance yield forecasts.

Figure 8.1 displays a diagrammatic "roadmap" of simulations and analyses performed in this chapter.

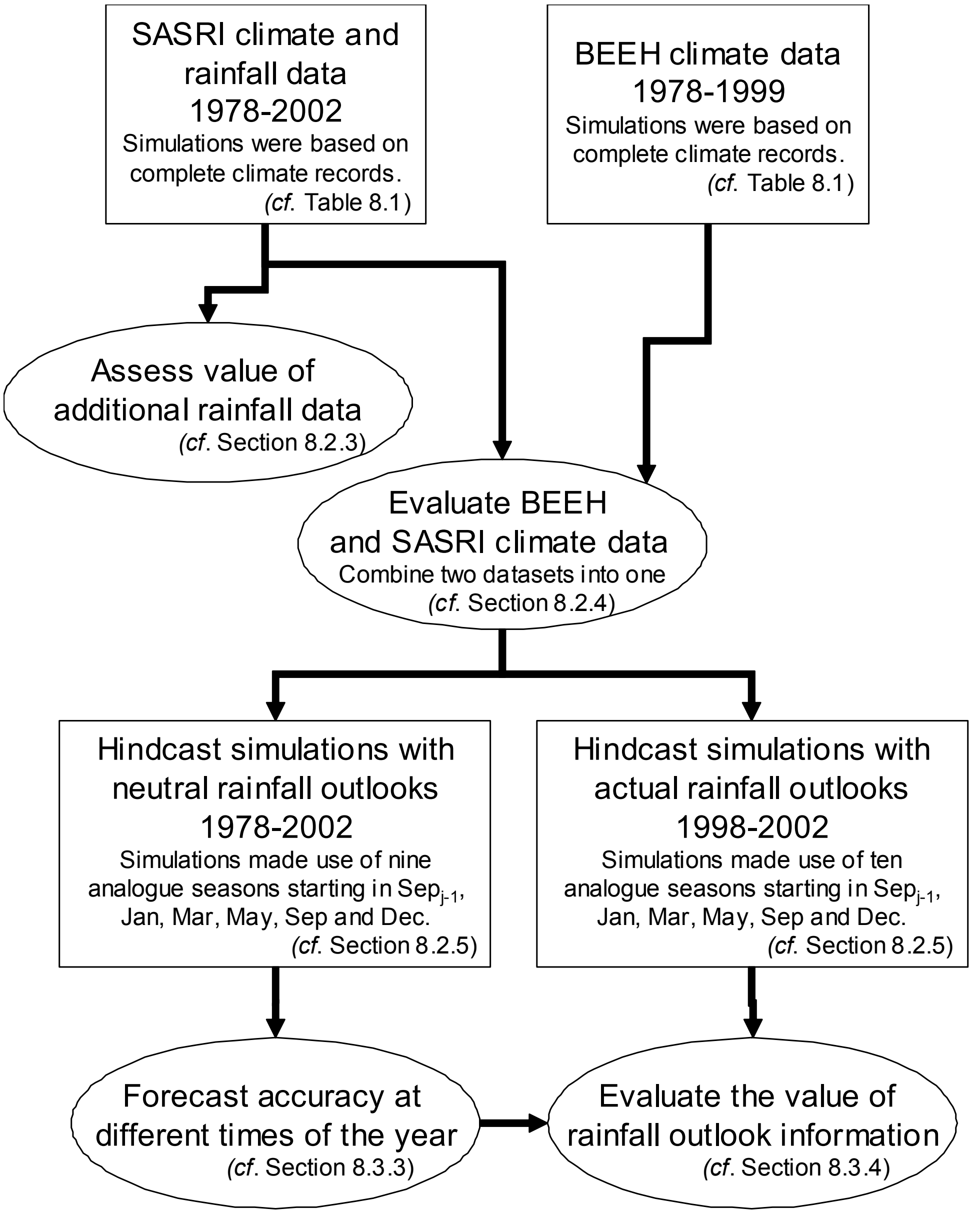

Figure 8.1 A "roadmap" of simulations and analyses performed in this chapter. Simulations are shown in boxes and analyses are shown in oval shapes

8.2 Methods

8.2.1 Industry Production Data Corrections

Hansen and Jones (1999), Stephens *et al.* (2000), Promburom *et al.* (2001) and Everingham *et al.* (2002a) point out that historic production data often contain underlying trends that are not attributed to climate variability. These include:

- Changes to production as a result of spatial expansion and the decommissioning of production areas,
- Changes in agronomic practices, such as reducing harvest age, when attempting to adopt new technologies or to mitigate against pests, diseases and anticipated climate,
- Changes in irrigation systems and strategies,
- Changes to agricultural practices as a result of socio-economic and socio-political preferences and pressures, and
- Changes to physical boundaries of reporting districts as a result of mills that close, merge or move.

Stephens *et al.* (2000) and Everingham *et al.* (2002a) demonstrate the value of historic production data once a certain amount of detrending and corrections have been made. Everingham *et al.* (2002a), for example, removed long-term trends and corrected for changes in production variability over time using a variance stabilising transformation. Stephens *et al.* (2000) overcame several aggregation and representation issues by assuming current production districts and technology for previous years.

Historic figures which report on sugarcane production from 1980 to 2002 for individual mill districts were obtained from the South African Canegrowers Association (SACGA, Mt. Edgecombe, South Africa). Evidence of all the above-mentioned underlying trends was apparent in the data. Various corrections aimed at transforming the data to present conditions and to units comparable with model output were consequently conducted. These included adjusting for mill closures and delivery diversions, converting from total tonnes crushed to mean tonnes per hectare, correcting for expansions in areas under irrigation and correcting for changes in age at harvest.

8.2.1.1 Corrections Resulting from Mill Closures and Sugarcane Diversions

Changes to the industry's milling configuration during the period 1980 to 2002 were evaluated. The Nkwaleni and Empangeni mills, situated in HCZs 20 and 19 (*cf.* Figure 6.1), respectively, were combined with the Felixton mill in 1985. The coastal Illovo mill (in HCZ 41) was relocated to the interior and renamed Eston mill in 1994. The Glendale mill (HCZ 27) closed in 1997 and sugarcane from that area was redirected to the Gledhow mill. The Tongaat mill was renamed Maidstone mill in 1982. In 1987 and 1995 additional sugarcane was delivered to this mill owing to the closure of the Shaka's Kraal (HCZ 29) and Mt. Edgecombe (HCZ 33) mills, respectively. In all the above-mentioned cases, data of total tonnes sugarcane crushed from the different mills were pooled together accordingly. It should be noted, however, that in almost all cases, as well as for the new Komati mill, which opened in 1994, changes to growing areas and, consequently, relative contributions from different HCZs to mills, were likely to have occurred. These changes were omitted because accurate records of new and discontinued old enterprises were unavailable from SACGA.

The total annual tonnes sugarcane crushed at each mill was also corrected for sugarcane that was diverted from other mill supply areas as a result of temporary management arrangements.

8.2.1.2 Conversions from Total Tonnes Crushed to Cane Yield

Sugarcane tonnages as described in Section 8.2.1.1 were converted to yield ($t.ha^{-1}$) by dividing tonnages with each mill's official estimated annual area harvested (data courtesy of the SACGA). It should be noted that estimated areas harvested are often suspect (Wynne, 2001) and areas managed by communal small-scale growers, for example, can undergo rapid and unannounced changes in land use and ownership. No information was available to assess or correct for such inconsistencies.

8.2.1.3 Harvest Age Corrections

Inman-Bamber (1991) notes that growers have been reducing their harvest age since the early 1980s owing to the severe impact of the *Eldana saccharina* stalk borer on older crops. This trend was confirmed in the data through a gradual increase in the percentage of total area harvested per season. Equation 8.1 was used to correct this trend by reducing the yield in a given year J on a pro rata basis for the difference between the mill average crop age in year J (Age_J in months) and the mill average crop age in 2001 (Age_{2001} in months), with

$$Y_J = Act_J \frac{Age_{2001}}{Age_J} \qquad \text{Eq. 8.1}$$

where Y_J and Act_J (in t.ha^{-1}) are the mean corrected and actual yields for the mill, respectively.

The mill's average crop age in 2001 (Age_{2001} in months) was calculated using the weighted average age over the different contributing HCZs (Eq. 8.2).

$$Age_{2001} = \sum_{i=1}^{N} w_i \times ZoneAge_i \qquad \text{Eq. 8.2}$$

where N is the number of HCZs supplying sugarcane to the mill, w_i is the percentage of the 2001 season's mill area harvested in zone i and $ZoneAge_i$ is the average age of harvested sugarcane in zone i in 2001 (*cf.* Table 5.1).

The mill's average crop age in year J (Age_J in months) was calculated using a five year inverse distance weighted average age (Eq. 8.3), such that

$$Age_J = \sum_{j=J-2}^{J+2} \frac{3-|j-J|}{9} Age_j \qquad \text{Eq. 8.3}$$

where

$$Age_j = 12 \times \frac{TotalAreaUnderCane_j}{TotalAreaHarvested_j}. \qquad \text{Eq. 8.4}$$

8.2.1.4 Cane Yield Corrections as a Result of Irrigation Expansion

An additional correction was made for the Felixton and Umfolozi mills, where there were disproportional large expansions of irrigated areas between 1980 and 2002. Irrigated sugarcane normally produces higher yields than rainfed sugarcane and on a mill scale a trend of increasing performance was observed in the corrected yields (Y_J) as a result of the increase in area under irrigation. Data of irrigation area expansions for these mills (courtesy of the SASA Industrial Affairs Division, Mt. Edgecombe, South Africa) were used to make adjustments to the corrected yields. First, a regression between the percentage area under irrigation and Y_J was established. Thereafter, values of Y_J, where J<2002, were adjusted to the same level as Y_{2002} using the above-mentioned regression.

8.2.2 Evaluation Parameters to Assess Forecast Accuracy

Murphy (1993) defined forecast accuracy as the average relationship between individual pairs of forecasted and realised observations. During a yield forecast exercise, Supit (1997) calculated the Relative Root Mean Square Error (*RRMSE* in %, Eq. 8.5) between observed and simulated yields. The *RRMSE* enables forecast accuracy to be compared in similar terms across mills (Supit, 1997).

$$RRMSE = 100 \times \frac{\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(Y_i - Est_i\right)^2}}{\overline{Y}} \qquad \text{Eq. 8.5}$$

where n is the total number of observed and simulated data pairs (23 years in this study), Y_i is a mean corrected observed yield in a given year i (according to Section 8.2.1), Est_i is the simulated yield in year i and $\overline{Y}$ is the mean of corrected observed yields over all years.

Haskett *et al.* (1995), Russell and van Gardingen (1997) and Hansen and Jones (1999) reported that model-based yield forecasts often overestimated actual yields. These overestimations may be attributed to several causes, which have been discussed briefly in Section 2.3. Bezuidenhout and Singels (2001) and Gers *et al.* (2001) identified similar overestimation trends when applying the Canesim model, respectively. A relative root mean square unbiased error (σ_ε in %, Eq. 8.6) was consequently calculated by multiplying model results with a correction factor. This technique calibrates simulated yields against the observed and also reduces the absolute variability of the simulation output (Hansen and Jones, 1999).

$$\sigma_\varepsilon = \frac{\sqrt{\frac{1}{23}\sum_{i=1980}^{2002}\left(Est_i \times \frac{\overline{Y}}{\overline{Est}} - Y_i\right)^2}}{\overline{Y}} \times 100 \qquad \text{Eq. 8.6}$$

where $\overline{Est}$ is the mean of the simulated yields over the period 1980 – 2002.

Forecast skill is a comparison of the quality of a forecast against another reference such as long-term mean, persistence or random guessing (Murphy, 1993; Mason, 2000). For this evaluation, forecast skill (*Skill* in %, Eq. 8.7) was determined from the ratio of σ_ε and the coefficient of variance of observed yields (CV_Y in %, Eq. 8.8). By definition, CV_Y is the forecast accuracy (*RRMSE*) when the the long-term mean yield ($\overline{Y}$) is always assumed for the forecast. Forecast skill can therefore be defined as the percentage of inter-annual production variability captured or explained by the forecast. For reporting purposes forecast skill was categorised into groups of no skill ($Skill < 10\%$), low skill ($10\% \leq Skill < 30\%$), medium skill ($30\% \leq Skill < 60\%$) and high skill ($Skill \geq 60\%$). Values for different mills and the industry were denoted by subscripts (e.g. $Skill_{Pongola}$, $Skill_{Sezela}$ and $Skill_{industry}$)

$$Skill = \left(1 - \frac{\sigma_\varepsilon}{CV_Y}\right) \times 100 \qquad \text{Eq. 8.7}$$

where

$$CV_Y = \frac{1}{\overline{Y}} \sqrt{\frac{1}{23} \sum_{i=1980}^{2002} \left(Y_i - \overline{Y}\right)^2} \times 100 \,. \qquad \text{Eq. 8.8}$$

In addition to the above-mentioned parameters, the frequency when yields were forecasted with the correct sign (*i.e.* higher or lower, when compared with yields of the previous season) was also expressed as a percentage. This was termed the Directional Skill.

8.2.3 Assessment of the Value to Accuracy from Additional Raingauge Data in HCZs within the SASRI Climate Dataset

The simulation configuration described in Chapter 7 was used to simulate nine crops per milling season (April to December) per HCZ from 1980 to 1999 using the SASRI climate dataset. These simulations were carried out for both rainfed and irrigated crops. Additional simulations were carried out when more than one raingauge existed within a HCZ. The simulation configuration was not altered over time and model output therefore reflected a temporal variation solely as a result of variabilities in climate and available water for irrigation. Water restrictions for irrigation were applied according to the procedures described in Section 5.3.3. The simulations were carried out in historic mode, thus using the full climate dataset, and not substituting future scenerios with analogue years. The simulation results were aggregated to a mill scale according to the information provided in Section 5.3.4.

Evaluations could only be performed at mill supply scales since no actual production information was available for individual HCZs (*cf.* Section 8.2.1). Two HCZs containing more than one raingauge were identified. These were the Heatonville (Zone 18) and Upper North Coast (Zone 29) zones. These zones contained four and six additional raingauges, respectively, and delivered sugarcane to the Felixton and Gledhow mills, respectively. In both cases, contributions from the respective HCZs to the respective mills were substantial (between 38% and 66% of the total mill demand). It was assumed that multiple raingauges in a zone represented equal proportions of the zone.

In order to assess the value of the additional raingauge data in the two HCZs, the following simulations were performed:

- Data from all raingauges (default);
- Data from all raingauges that did not contain a complete data record from 1978 to 2002 were excluded;
- Data from only two raingauges with comlpete records were included;
- Data from one raingauge with a complete record were included;
- All additional raingauges were removed and only the main climate station data were used.

The forecast skills (Eq. 8.7) for both mills were calculated under each of the above-listed scenarios and results were compared to each other. These results are discussed in Section 8.3.1.

8.2.4 Assessment of the Value of Climate Data

Two sets of model runs, based on the SASRI and the BEEH climate datasets, were carried out from 1980 to 1999 in a similar fashion to those described in Section 8.2.3. For the SASRI climate dataset, these simulations included additional raingauge data. Values for σ_ε (Eq. 8.6) and *Skill* (Eq. 8.7) for both the BEEH and SASRI climate datasets were established for all individual mills and for the sugar industry as a whole. These reflected the overall accuracy of the model when using either one or the other of the two datasets. This information was used later (*cf.* Sections 8.3.2 & 8.3.3) to compile a single composite climate dataset from both resources.

It could be expected that either the SASRI or the BEEH climate dataset would be more representative for a particular HCZ, in which case a higher associated *Skill* value, both at a mill and an industry scale, could be expected. The change in *Skill* could therefore be used to evaluate the suitability of a particular climate dataset in a HCZ, which may provide valuable pointers towards increasing or reducing the current SASRI climate station network. It could therefore be assumed that, if the BEEH climate dataset produced better results than the SASRI dataset, there may be a need to review the location and number of SASRI climate stations within that zone. The increase in $Skill_{industry}$ was hence used to quantify the suitability of a climate dataset. This was done after the relative weighting resulted from different areas under cane within different HCZs was removed (*cf.* Eq. 8.9).

$$\Delta Skill_{industry(i)} = \frac{\lfloor Skill_{industry,BEEH(i)} \rfloor - \lfloor Skill_{industry,SASRI(i)} \rfloor}{Area_i} \times 10^4 \qquad \text{Eq. 8.9}$$

where $Area_i$ (ha) is the total area of cane harvested per annum in the i-th HCZ and $Skill_{industry,BEEH(i)}$ and $Skill_{industry,SASRI(i)}$ are the industry scale forecast skills under the condition that yields in the i-th HCZ were simulated using the BEEH or SASRI climate dataset, respectively. $\Delta Skill_{industry(i)}$ (in %) is therefore the change in $Skill_{industry}$ per 10 000 ha as a result of the change in climate datasets in the i-th HCZ. A positive value indicates better results obtained by the BEEH climate dataset, while a negative value supports the use of the SASRI climate dataset. The results are discussed in Section 8.3.2.

8.2.5 Model Forecast Accuracy in an Operational Context and Quantification of the Value of Seasonal Rainfall Outlook Information

A composite climate dataset based on the best results in the previous section (*cf.* Section 8.3.2) was compiled for the period 1978 to 2002. This dataset comprised of combined BEEH and SASRI climate data for the period 1978 to 1999 and SASRI climate data from 2000 to 2002. The composite dataset was used to re-simulate the yield forecasts from 1980 to 2002 in a hindcast mode. During this undertaking it was assumed that climate data terminated at a certain time of the year (as explained in Section 7.3). Climate data were terminated on 1 September prior to the opening of the milling season, 1 January, 1 March, 1 May, 1 September and 31 December. In each case, with the exception of 31 December, nine analogue seasons were used to complete the simulations. The nine analogue seasons were based on a neutral rainfall outlook with a three-month lead time, resulting in the selection of three below-normal (*i.e.* dry) scenarios, three scenarios with near-normal rainfall and three above-normal (*i.e.* wet) scenarios. The simulations were carried out for each HCZ from 1980 to 2002 for sugarcane crops harvested in each month of the milling season (April – December). Results were aggregated and evaluated at a mill and industry scale. Evaluations included the calculation of σ_{ε} , *Skill* and the directional skill. These evaluations reflected the accuracy of the Canesim sugarcane yield forecast system at different times of the year if it were to be used in an operational context under the scenario where future climate data were still unavailable. Results are discussed in Section 8.3.3.

In addition to the above-mentioned simulations, the seasons from 1998 to 2002 were also simulated after using the actual SAWS seasonal rainfall outlook that was issued at that specific time (information courtesy of the South African Weather Service, Pretoria, South Africa). Compared to the simulations described in the previous paragraph, these simulations were therefore based on different selected analogue seasons. The difference in forecast skill between simulations based on a neutral seasonal rainfall outlook and those based on the actual seasonal rainfall outlook was used to quantify the value of rainfall outlook information. Results are discussed in Section 8.3.4.

8.3 Results and Discussion

8.3.1 The Value of Additional Raingauge Information

Figure 8.2 displays the change in *Skill* at the Felixton and Gledhow mills resulting from the exclusion of additional raingauge information in the Heatonville and Upper North Coast HCZs, respectively. Yield forecasts were marginally better when more raingauges were used in the Heatonville HCZ. For the Upper North Coast HCZ, yield forecasts were most accurate when only two additional raingauges were used.

Forecast skill was generally not affected severely at Felixton, while the skill deteriorated significantly at Gledhow when fewer than two additional raingauges were used in the respective HCZ. This could be an indication that the main climate station used in the Upper North Coast HCZ is unsuitably located to represent the larger area. It could also, however, be an indication that higher rainfall variability in the zone exists and the zone may therefore require more rainfall stations.

From these results it is clear that additional raingauges generally enhanced forecast accuracies. The results, however, are limited and not consistent as to how many additional gauges are required to optimally enhance model-based yield forecasts. It could be expected that certain areas, such as predominantly rainfed areas, may benefit more when data from additional raingauges are used. Further research, preferably at a sub-mill scale, will be required to provide better guidelines regarding the spatial distribution of raingauges.

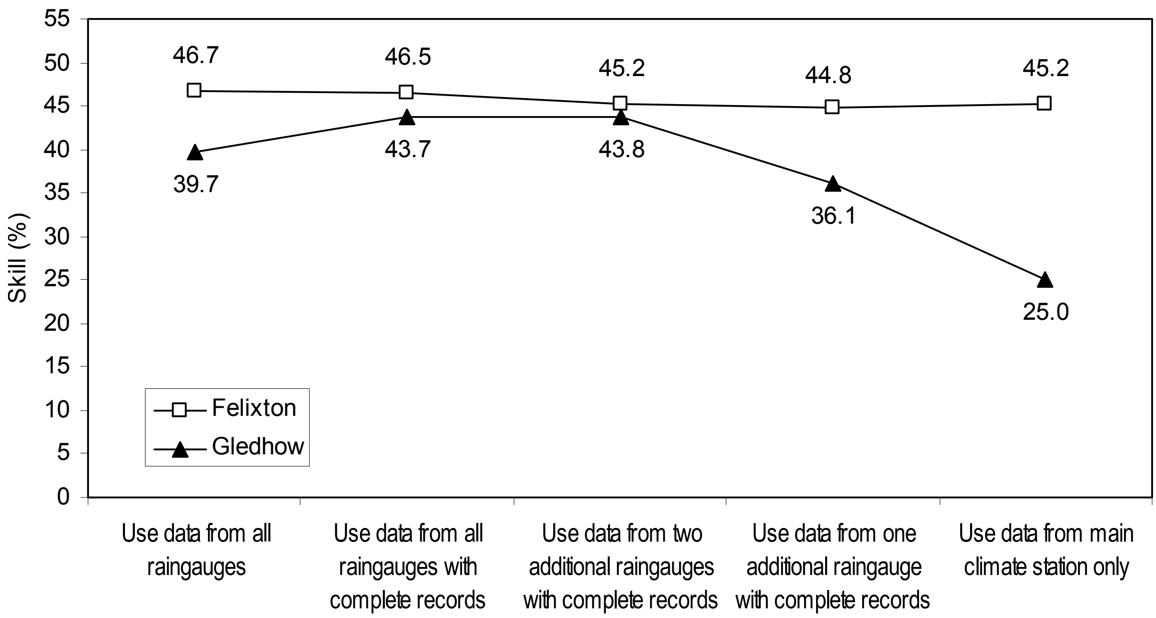

Figure 8.2 Changes in forecast skills for yields at the Felixton and Gledhow sugar mills based on the number of raingauges used in one of each mill's associated homogeneous climate zones

8.3.2 The Value of Climate Data

Table 8.1 gives values for σ_ε and *Skill* for the Canesim yield forecast system at mill and industry scales when the SASRI and BEEH climate datasets were used to simulate yields. These results are based on a complete historical climate dataset from

1978 to 1999, with no substitution using analogue seasons. At seven of the 15 mills, forecasts using the SASRI climate dataset were significantly superior to the BEEH climate dataset (difference in *Skill* > 10%). Suspect evaporation estimates in the SASRI climate data for the Komati HCZ (Komati mill) were found. Forecasts at the Pongola and Umfolozi mills displayed no skill. The reason for low accuracies at these two mills is unknown, but it should be noted that both mill areas are largely irrigated and that little information of actual irrigation restrictions are available. Further research, however, is needed to explain these inaccuracies. The forecast skill on an industry scale was medium and differed by 0.1% between different climate datasets.

Table 8.1 Evaluation statistics of modelled sugarcane yield forecasts based on climate data collated by the South African Sugarcane Research Institute (SASRI) and the School of Bioresources Engineering and Environmental Hydrology (BEEH, University of KwaZulu-Natal), respectively. The bias corrected Relative Root Mean Square Error (σ_ε) and forecast skill (*Skill*) are supplied for mills and the industry as a whole

No.	Mill	Mean of actual yield (t/ha)	Standard deviation of actual yield (t/ha)	SASRI climate dataset σ_ε (%)	SASRI climate dataset *Skill* (%)	BEEH climate dataset σ_ε (%)	BEEH climate dataset *Skill* (%)
1	Komati	78.2	17.44	suspicious data		18.13	-4.0
2	Malelane	82.6	15.68	13.58	13.4	11.47	26.8
3	Pongola	75.7	8.17	8.95	-9.5	12.56	-53.8
4	Umfolozi	66.5	9.35	8.77	6.3	8.45	9.6
5	Entumeni	41.7	6.56	5.68	13.4	6.73	-2.6
6	Felixton	59.7	12.38	6.60	46.7	6.59	46.8
7	Amatikulu	45.8	9.09	3.84	57.8	4.18	54.0
8	Darnall	48.8	9.19	4.12	55.2	5.91	35.7
9	Gledhow	49.2	8.41	5.07	39.7	6.12	27.2
10	Union Co-op	71.8	11.91	6.81	42.8	9.30	21.9
11	Noodsberg	67.7	14.23	6.89	51.6	9.16	35.6
12	Maidstone	49.2	8.73	5.36	38.6	5.91	32.3
13	Eston	59.5	9.54	6.48	32.0	7.45	21.9
14	Sezela	55.6	13.46	5.21	61.3	5.20	61.4
15	Umzimkulu	69.8	13.58	6.74	50.4	4.99	63.2
Mill mean		**61.5**	**11.18**	**7.94**	**30.4**	**8.14**	**25.1**
Industry		**56.5**	**8.93**	**3.69**	**58.6**	**3.70**	**58.5**

Figure 8.3 depicts the difference in $Skill_{industry}$ per 10 000 ha sugarcane based on an evaluation of yields generated by the SASRI and BEEH climate datasets, respectively. Data from SASRI for the Komati HCZ (Zone 1) were suspect, as stated previously, and the BEEH climate dataset was used for this zone, without any comparison between datasets being made. No SASRI climate data existed for Zones 24 (Muden), 37 (Tala Valley) and 40 (Umkomaas) and long-term mean production (as noted in Section 7.3) was used to represent these zones. Zones 37 and 40 were subsequently

represented more skilfully by the BEEH climate dataset. However, simulations based on the BEEH climate dataset were less skilful than assuming the long-term mean yield for the Muden HCZ. All harvested yields for Muden were consequently set to 108 $t.ha^{-1}$.

Climate data from the BEEH climate dataset generally produced better results in the Mpumalanga region (Zones 1 – 8, Figure 8.3). The likely reason for this is that the SASRI climate station configuration in Mpumalanga underwent several changes during the period of 1980 to 1999. These included the termination of manual stations with long historical data records and the installation of several new AWSs at new sites. In contrast, a reliable climate station was operated by SASRI in the Pongola HCZ (Zone 10). Data from this station were also used in Zone 11 and for certain time periods in Zone 9 (*cf.* Table 5.3). In Zululand, the North Coast and the Midlands (Zones 12 – 40) results using the SASRI climate dataset were superior to those produced by the BEEH climate dataset. In these regions, SASRI was managing well established climate stations in Zones 28, 29, 30, 31 and 33. Evaluations of yields using SASRI climate data in these specific zones were all superior to yields produced by the BEEH dataset. In the Midlands, Zones 26, 34, 37 and 40 were better represented by the BEEH climate dataset. None of these zones contained reliable SASRI climate stations and, with the exception of Zone 26, these zones all represented predominantly irrigated sugarcane that was cultivated in the warm climates of deep river valleys. On the South Coast (Zones 41 – 48), most zones were better represented by the BEEH climate dataset. Zone 44 contains a well established SASRI climate station, *viz.* Sezela (*cf.* Table 6.1). This was the only zone that contained a well established SASRI climate station in which the BEEH climate dataset produced superior results. Although the Sezela climate station may have a long and consistent data record, this station is situated on the coast and may be unrepresentative of the zone as a whole.

It should be noted that the results shown in Figure 8.3 are based on the difference in skill between yields generated by the SASRI and BEEH climate datasets. A specific climate dataset may, therefore, be favoured because of its good climate representivity in a particular zone, or because of the other dataset's poor data representivity. It is nevertheless observed that, with the exception of Zone 44, all HCZs containing well managed and continuous SASRI climate stations (*viz.* Zones 10, 21, 31, 28, 30, 29 and 33) were simulated more skilfully when compared to yields produced by the more generic BEEH climate dataset. It is imperative for forecast modelling to have a well-managed climate station network. These results emphasise the importance to maintain climate stations with sound data quality control over long periods of time and also suggest that there may exist room for expanding the current SASRI climate station network.

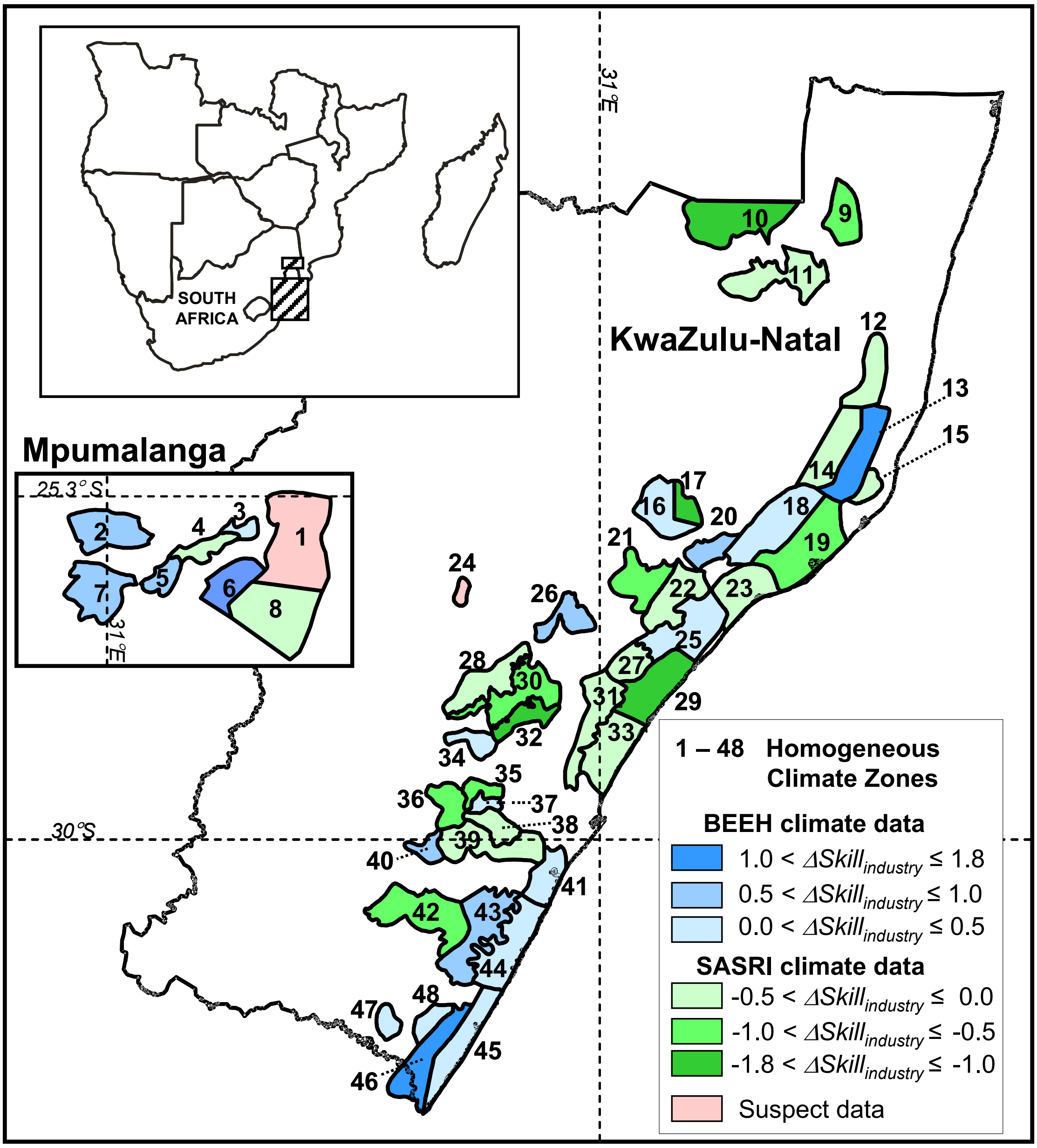

Figure 8.3 Climate data origins used to represent homogeneous climate zones in the South African sugar producing areas. Climate data sources are BEEH, *i.e.* the School of Bioresources Engineering and Environmental Hydrology at the University of KwaZulu-Natal, and SASRI, *i.e.* the South African Sugarcane Research Institute. Colour shadings indicate the relative difference in forecast skills between the two climate data sources

8.3.3 Accuracy of Operational Forecasts

Information contained in Table 8.2a, b and c reflects the 23 year mean accuracy of operational Canesim yield forecasts issued at different times of the year. The subscript "*y-1*" denotes a forecast issued during September in the year prior to the particular milling season. Accuracies are expressed in terms of σ_ε, *Skill* and directional skills, respectively. All results were obtained under a neutral rainfall outlook using a combined SASRI / BEEH climate dataset for the period 1978 to 1999, and appending the remainder of the period 2000 to 2002 using the SASRI climate dataset.

Figure 8.4 illustrates six identical time series reflecting average annual historical yields that were achieved over the entire sugar industry in consecutive years between 1980 and 2002 (solid lines). Figure 8.4 also shows different time series of industry mean yields simulated by the Canesim system (dotted lines). These yields were simulated after it was assumed that available climate data terminated at a particular time, *viz.* September$_{y-1}$, January, March, May, September and December. Similar graphs for the 15 mill supply areas are displayed in Appendix D.

Forecast skills of simulated yields, as may be expected, generally increase as time progresses. At individual mill level (*cf.* Table 8.2b), skills increased on average from 11.6% in September prior to the milling season to 33.3% in December. Forecasts issued in the September prior to the milling season were generally more accurate at mills with longer cropping cycles (e.g. Noodsberg and Umzimkulu mills), as a result of the relatively smaller proportions of the crop that were still outstanding at that point in time. With the exception of the Komati mill, which had a shorter historic record (*cf.* Appendix D), the highest forecast skills were achieved at the Umzimkulu, Noodsberg, Amatikulu and Sezela mills (*cf.* Table 8.2c). These mill supply areas cover a diverse range of agronomic and climatic conditions. The fact that the system could manage to perform reasonably well under a wide range of conditions emphasises the proficiency of a model-based yield forecast system. However, it also prompts further investigations into the lower skills obtained at other mills in areas with similar climates.

Table 8.2 A summary of sugarcane yield forecast accuracies at mill and industry scales. Forecast were issued at six different times of the milling season and their accuracies are expressed as (a) bias corrected Relative Root Mean Square Errors (σ_ε), (b) forecast skills and (c) directional skills

(a)

No.	Mill	σ_ε (%) 1 Sep$_{y-1}$	1 Jan	1 Mar	1 May	1 Sep	31 Dec	**Mean**
1	Komati	17.2	7.2	3.9	6.0	14.7	18.7	**11.3**
2	Malelane	16.7	13.6	12.7	12.6	12.8	13.1	**13.6**
3	Pongola	8.8	8.0	6.3	7.6	13.8	11.3	**9.3**
4	Umfolozi	14.6	12.5	11.3	11.1	12.6	12.8	**12.5**
5	Entumeni	12.9	12.5	11.5	11.3	12.6	12.7	**12.3**
6	Felixton	18.9	17.3	13.8	11.9	10.5	10.7	**13.8**
7	Amatikulu	17.7	15.8	11.0	8.6	7.5	7.3	**11.3**
8	Darnall	16.5	14.9	11.3	8.2	8.6	8.5	**11.3**
9	Gledhow	14.7	15.2	12.5	11.1	10.9	10.7	**12.5**
10	Union Co-op	12.8	10.7	9.8	9.7	9.0	9.1	**10.2**
11	Noodsberg	15.6	13.1	11.0	10.9	9.8	10.6	**11.8**
12	Maidstone	15.3	15.5	11.9	10.8	11.6	11.4	**12.7**
13	Eston	13.3	11.9	12.1	11.5	11.9	11.7	**12.1**
14	Sezela	20.4	17.1	13.6	11.6	10.8	10.9	**14.0**
15	Umzimkulu	14.5	12.3	9.9	8.2	8.3	8.2	**10.2**
Mean		**15.3**	**13.2**	**10.8**	**10.1**	**11.0**	**11.2**	
Industry		**13.7**	**11.8**	**8.8**	**7.1**	**6.6**	**6.5**	

(b)

No.	Mill	Forecast Skill (%) 1 Sep$_{y-1}$	1 Jan	1 Mar	1 May	1 Sep	31 Dec	**Mean**
1	Komati	12.4	63.3	80.3	69.4	25.1	5.0	**42.6**
2	Malelane	5.5	23.0	28.0	28.6	27.4	26.0	**23.1**
3	Pongola	14.2	22.1	38.2	25.5	-35.3	-10.9	**9.0**
4	Umfolozi	3.7	18.0	25.8	27.0	16.9	15.6	**17.8**
5	Entumeni	11.7	14.6	21.6	23.0	14.2	13.1	**16.4**
6	Felixton	3.6	11.8	29.8	39.6	46.6	45.7	**29.5**
7	Amatikulu	6.1	16.3	41.9	54.6	60.2	61.5	**40.1**
8	Darnall	8.6	17.2	37.0	54.6	52.2	52.8	**37.1**
9	Gledhow	10.8	8.0	24.4	32.4	33.9	34.9	**24.1**
10	Union Co-op	21.7	34.6	40.3	40.6	45.1	44.0	**37.7**
11	Noodsberg	21.7	34.2	44.8	45.7	51.0	47.1	**40.7**
12	Maidstone	9.8	8.9	29.7	36.4	31.5	32.7	**24.8**
13	Eston	12.9	22.1	20.9	25.2	22.1	23.9	**21.2**
14	Sezela	10.0	24.8	40.1	49.1	52.4	52.1	**38.1**
15	Umzimkulu	21.3	32.9	46.1	55.5	54.7	55.4	**44.3**
Mean		**11.6**	**23.4**	**36.6**	**40.5**	**33.2**	**33.3**	
Industry		**11.3**	**23.2**	**43.2**	**54.0**	**57.3**	**57.9**	

(c)

No.	Mill	Directional Skill (%) 1 Sep$_{y-1}$	1 Jan	1 Mar	1 May	1 Sep	31 Dec	**Mean**
1	Komati	75.0	87.5	87.5	75.0	75.0	75.0	**79.2**
2	Malelane	63.6	63.6	59.1	59.1	50.0	40.9	**56.1**
3	Pongola	63.6	54.5	59.1	54.5	59.1	54.5	**57.6**
4	Umfolozi	54.5	59.1	77.3	81.8	81.8	81.8	**72.7**
5	Entumeni	72.7	59.1	72.7	68.2	77.3	77.3	**71.2**
6	Felixton	54.5	63.6	86.4	90.9	95.5	95.5	**81.1**
7	Amatikulu	77.3	59.1	81.8	90.9	95.5	95.5	**83.3**
8	Darnall	68.2	77.3	86.4	86.4	86.4	90.9	**82.6**
9	Gledhow	72.7	81.8	90.9	90.9	90.9	86.4	**85.6**
10	Union Co-op	63.6	72.7	81.8	77.3	81.8	81.8	**76.5**
11	Noodsberg	68.2	81.8	86.4	81.8	77.3	72.7	**78.0**
12	Maidstone	72.7	68.2	81.8	81.8	81.8	72.7	**76.5**
13	Eston	77.3	63.6	72.7	77.3	77.3	77.3	**74.2**
14	Sezela	63.6	77.3	86.4	81.8	86.4	90.9	**81.1**
15	Umzimkulu	63.6	68.2	72.7	72.7	77.3	72.7	**71.2**
Mean		**67.4**	**69.2**	**78.9**	**78.0**	**79.5**	**77.7**	
Industry		**72.7**	**59.1**	**77.3**	**77.3**	**81.8**	**81.8**	

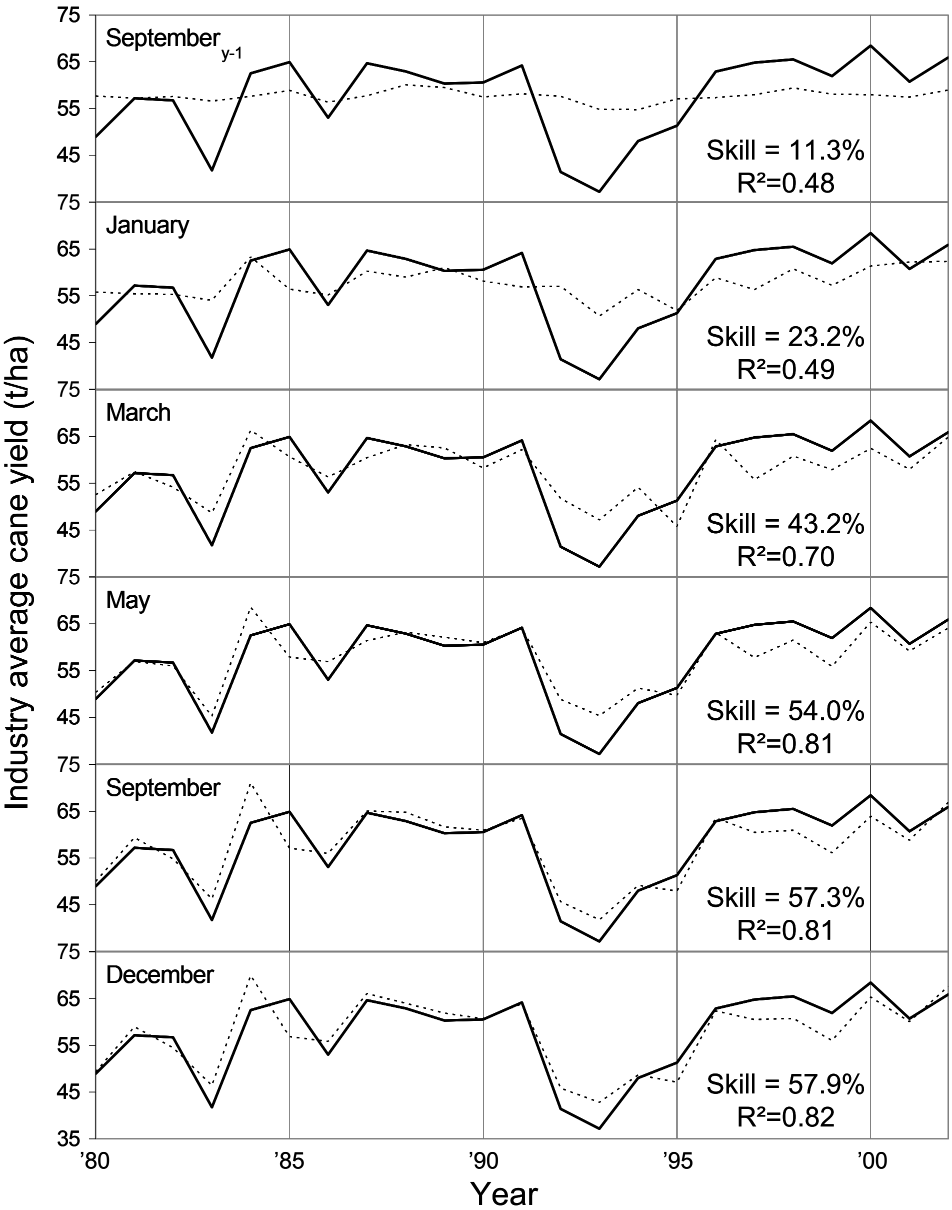

Figure 8.4 Time series of mean actual yield (solid lines) obtained by the South African sugar industry between 1980 and 2002. Dotted lines denote the Canesim system forecasted yields at different times of the season, starting in the September prior to the opening of the milling season (September$_{y-1}$)

It should be noted that skills at the Komati, Pongola, Umfolozi and Entumeni mills deteriorated significantly towards the end of the milling season. In all these instances the deterioration in accuracies commenced between May and September. Closer investigation revealed an over-prediction of inter-annual yield variability when longer periods of actual climate data (as opposed to analogue data) were available. The trend is likely to be caused by an over-simplification of input parameters. For example, if a HCZ is represented by only one simulated crop when more diverse conditions apply in reality, then fluctuations in simulated yields can be expected to exceed fluctuations in the actual production. This problem may be addressed by accepting a more diverse range of agronomic inputs, which will allow for certain portions of the crops in the HCZ to stress, while other portions experience better growing conditions. In the simulations conducted for this study, diverse conditions were simulated when climate data terminated prematurely and different analogue future scenarios had to be assumed. It is suspected that for some HCZs these simulations could indirectly have addressed the lack of diversity in agronomic inputs, since some analogue seasons will allow crops to stress, while others will allow better growing conditions. These suspicions are confirmed for some mills by the increase in forecast errors as the proportion of the crop simulated by analogue data decreases. Negative forecast skills for the Pongola mill in September and December depict that these forecasts were worse than simply assuming the mill's long-term mean as a forecast. The fact that accuracy at these mills deteriorated over the winter season suggests that simulations of winter growing conditions could be oversimplified. These observations point to the need for further research into the diversification of soil, harvest cycle, water use and irrigation strategy input variables within HCZs.

Even though forecast skills (Table 8.2b) were often less than 30%, the system frequently forecasted the *direction* of future yields correctly, compared to the previous season (*cf*. Table 8.2c). For example, the directional skill for the entire industry in the September prior to the milling season was 73%, while the equivalent forecast skill was only 11%. Generally, the directional skill was higher for coastal mills dependent on rainfed sugarcane (e.g. Gledhow, Amatikulu, Darnall, Sezela) rather than for inland mills and mills which were supplied by large areas of irrigated sugarcane.

8.3.4 The Value of Seasonal Rainfall Outlook Information

The average changes in forecast skill over the period 1998 to 2002 are shown in Table 8.3 and Figures 8.5a-e for HCZs, mills and the entire sugar industry. These are the results after analogue years were based on actual SAWS seasonal rainfall outlook information, as opposed to an assumed neutral outlook. Appendix E contains time series graphs for the industry as a whole and for individual mills. Each of these displays mean actual yields and two forecasts, one based on a neutral outlook and another based on the actual SAWS seasonal rainfall outlook.

Seasonal rainfall outlook information was the most valuable in January, where industry scale skills were improved by an average of 7.9% over the period 1998 to 2002. The value of rainfall outlook information deteriorated from January towards the drier winter months, e.g. May. Also, rainfall outlook information issued in September (early spring) often had a negative effect on forecast skills. Generally, the Eston, Amatikulu and Gledhow mills reflected promising improvements in forecast skills when using the seasonal rainfall outlook (*cf.* Table 8.3). Also, forecasts based on the seasonal rainfall outlook issued for the Eston and the Umzimkulu mills were consistently more accurate than those based on a neutral outlook (*cf.* Table 8.3).

The selection of analogue seasons in this study was based on three month accumulated rainfall outlook information. Several other seasonal climate outlooks are also available. These range from monthly accumulated rainfall outlooks, zero to three month mean temperature outlooks and three to six month mean temperature and rainfall outlooks. In addition, several crop response driving parameters, such as the number of rainy days, the number of cold, hot, wet and dry spells and the number of days with temperature and rainfall events exceeding certain threshold amounts may be more valuable than forecasts of means. A close partnership between climate forecasters and the sugar industry (and other sectors) may be necessary to ascertain and make advances on potential synergies.

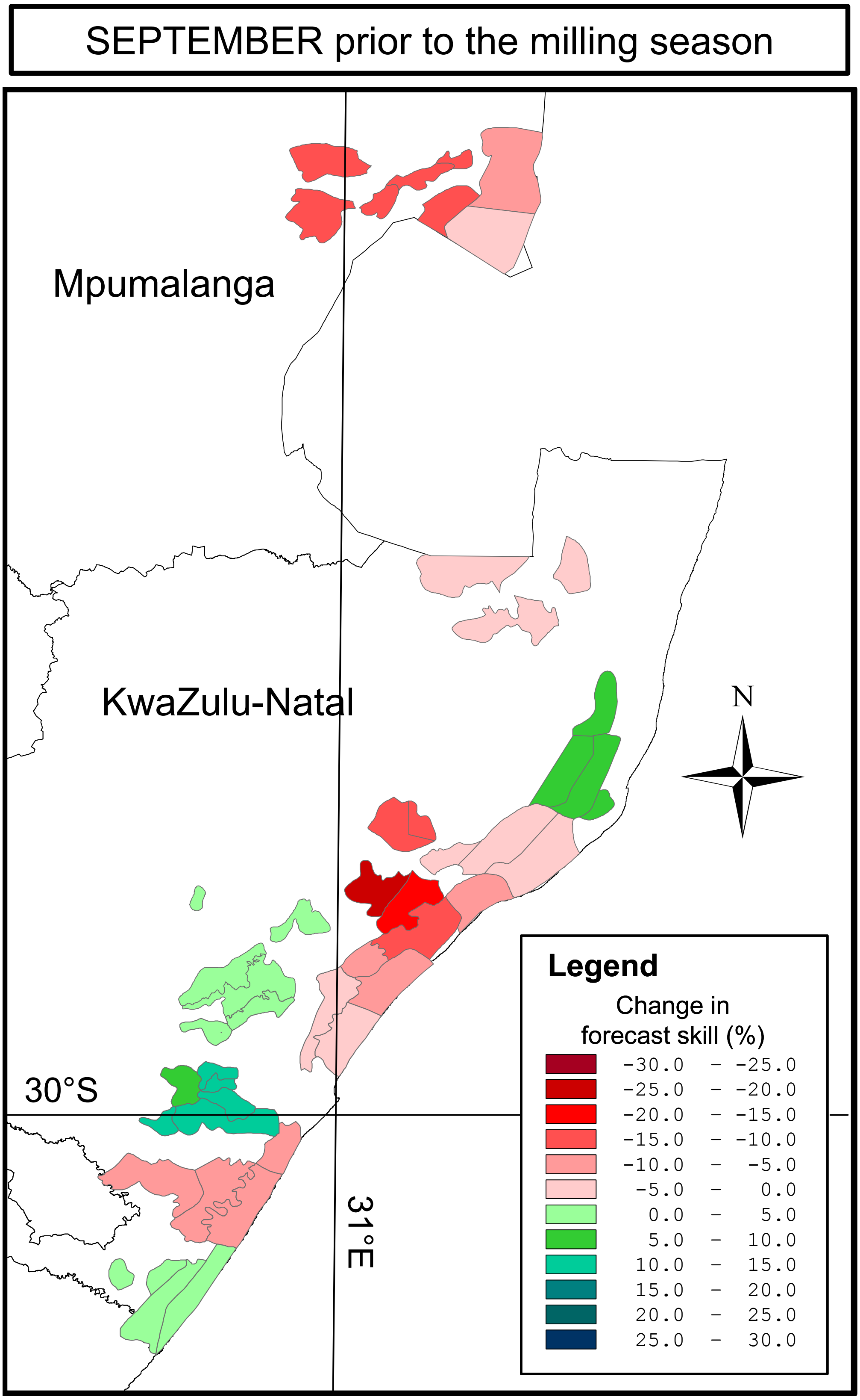

Figure 8.5a Spatial changes in forecast skill as a result of utilising seasonal rainfall outlook information as opposed to assumptions of neutral outlooks for forecasts issued in the September prior to the milling season

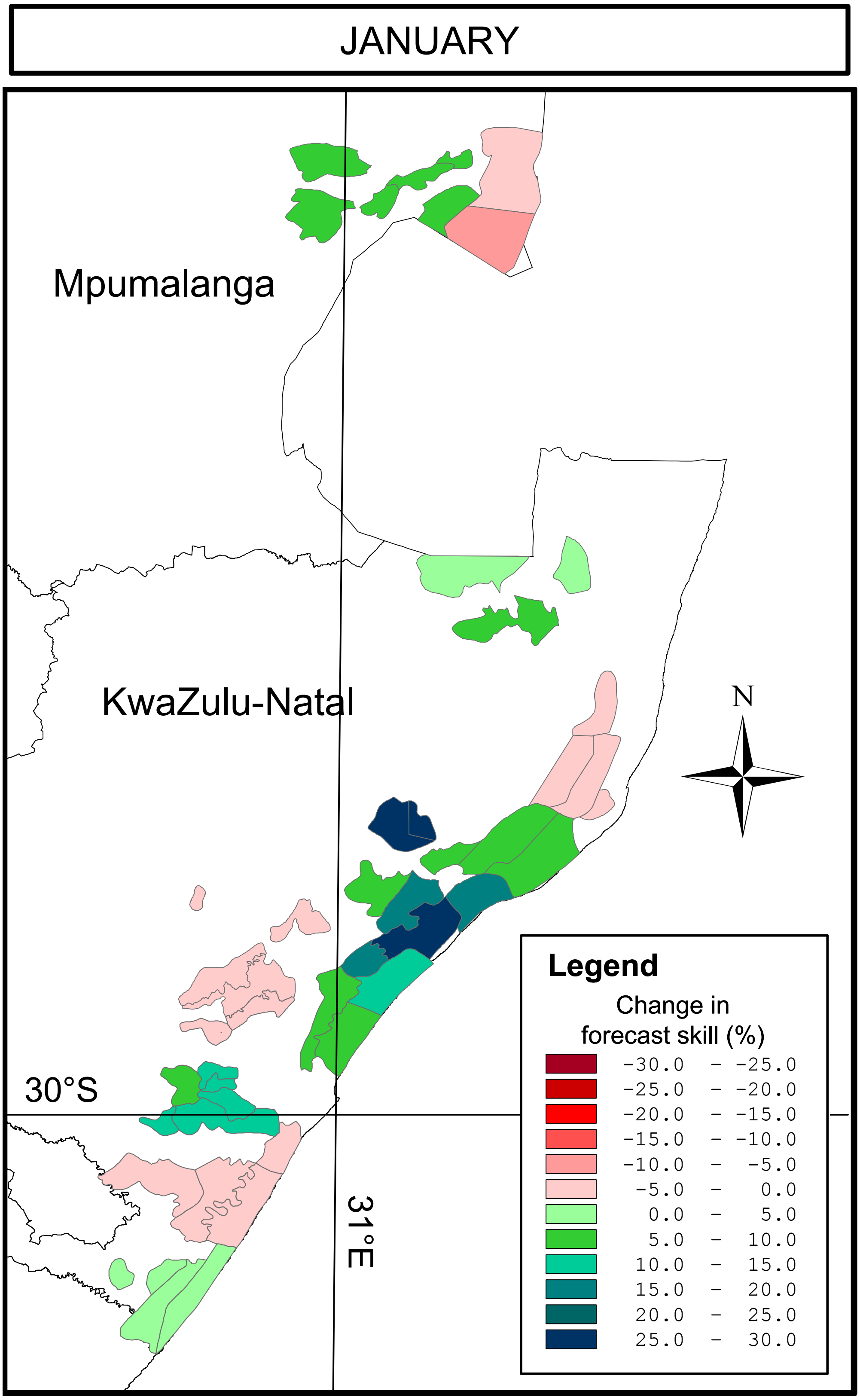

Figure 8.5b Spatial changes in forecast skill as a result of utilising seasonal rainfall outlook information as opposed to assumptions of neutral outlooks for forecasts issued in January

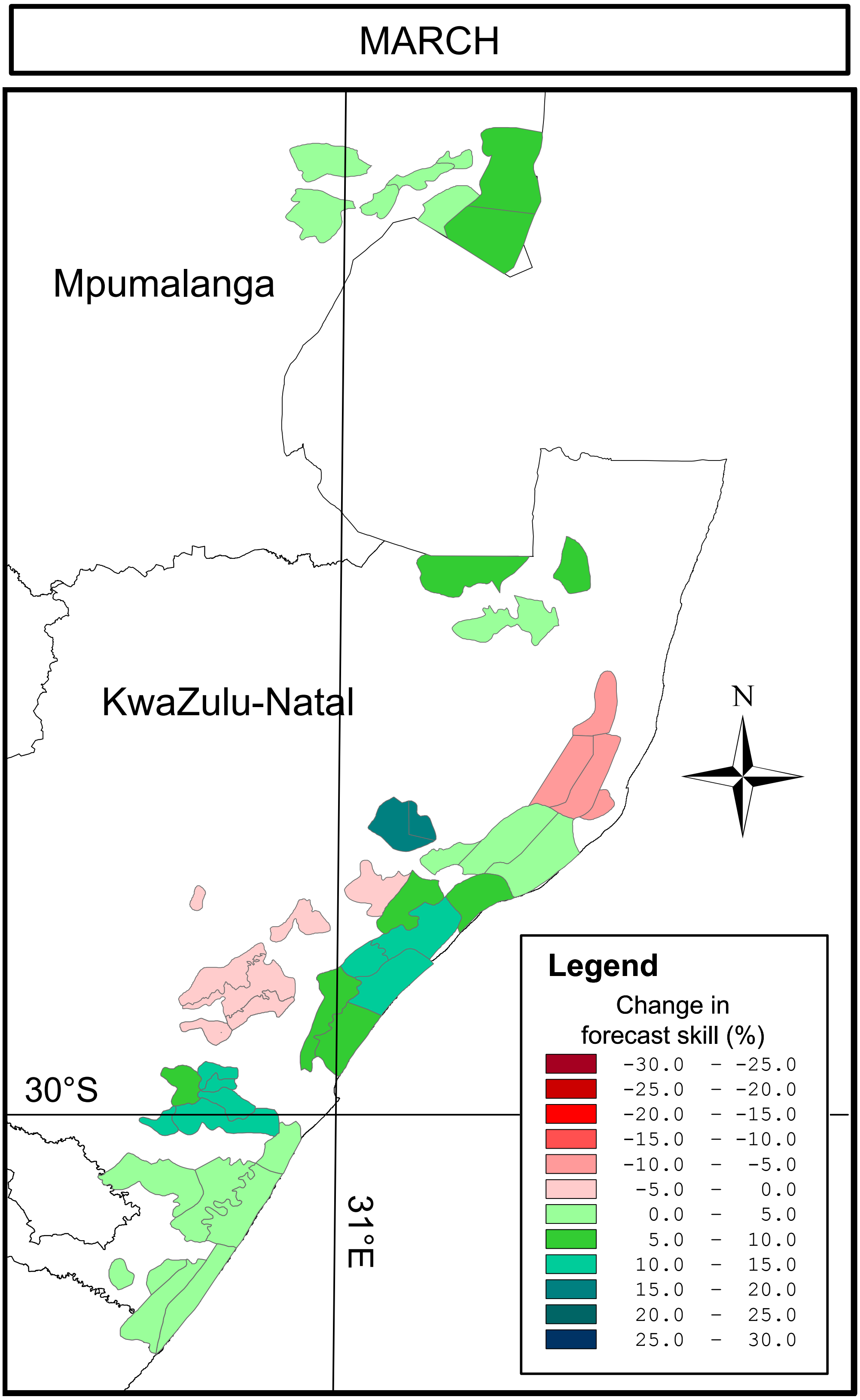

Figure 8.5c Spatial changes in forecast skill as a result of utilising seasonal rainfall outlook information as opposed to assumptions of neutral outlooks for forecasts issued in March

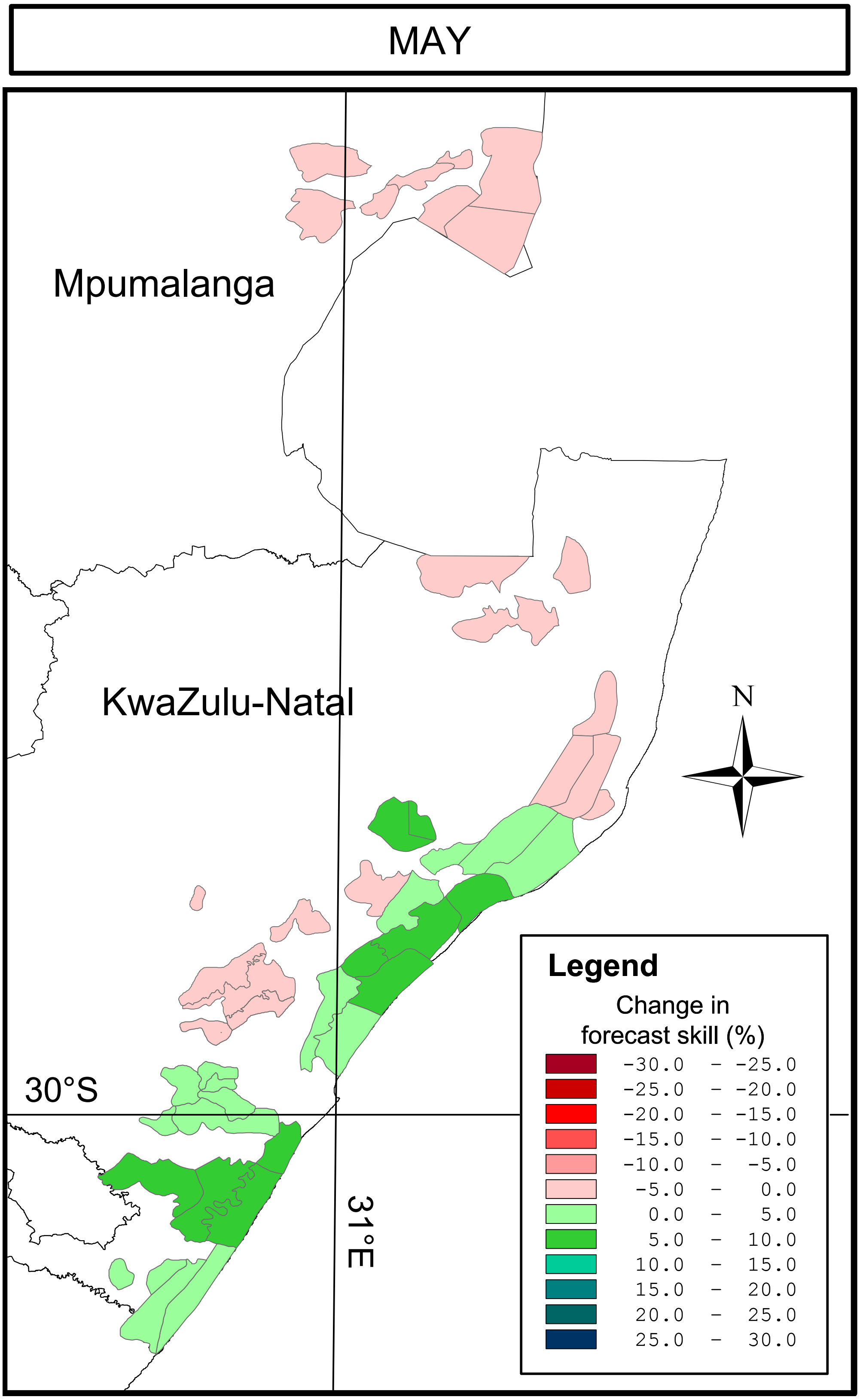

Figure 8.5d Spatial changes in forecast skill as a result of utilising seasonal rainfall outlook information as opposed to assumptions of neutral outlooks for forecasts issued in May

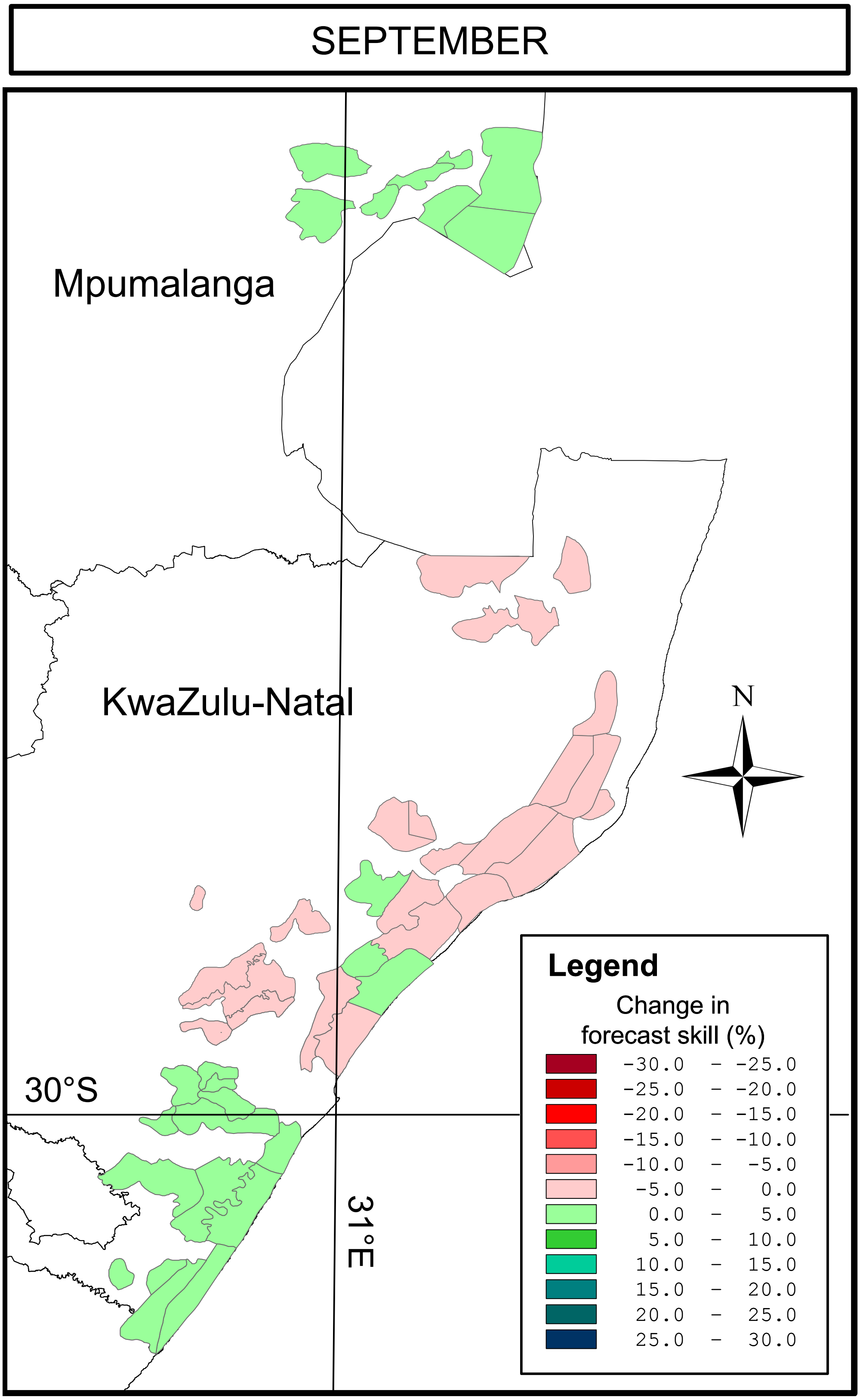

Figure 8.5e Spatial changes in forecast skill as a result of utilising seasonal rainfall outlook information as opposed to assumptions of neutral outlooks for forecasts issued in September

Table 8.3 The change in forecast skill between yield forecasts based on a neutral seasonal rainfall outlook and those based on actual seasonal rainfall outlooks. Positive values depict an improvement in forecast skill when actual outlook information was used. Underlined values are statistically significant (*P*<0.05) according to a binomial distribution test

		Change in forecast skill (%)					
No.	Mill	1 Sep_{y-1}	1 Jan	1 Mar	1 May	1 Sep	**Mean**
1	Komati	<u>-3.49</u>	-7.86	6.64	-4.09	0.03	**-1.75**
2	Malelane	<u>-11.5</u>	6.47	0.81	-1.51	0.05	**-1.13**
3	Pongola	-2.96	1.85	6.16	-1.20	<u>-3.20</u>	**0.13**
4	Umfolozi	9.29	-1.32	-6.04	-1.66	-0.43	**-0.03**
5	Entumeni	-20.6	7.65	-0.14	-2.38	0.06	**-3.09**
6	Felixton	-2.07	<u>8.35</u>	0.81	0.93	-0.17	**1.57**
7	Amatikulu	-11.8	27.74	15.04	9.36	-0.59	**7.95**
8	Darnall	-12.0	15.67	<u>13.24</u>	<u>6.54</u>	0.84	**4.86**
9	Gledhow	-3.15	17.82	<u>12.79</u>	6.45	0.69	**6.92**
10	Union Co-op	1.54	-1.03	-2.15	-0.43	-1.32	**-0.68**
11	Noodsberg	0.92	-1.28	-2.32	-0.37	-1.12	**-0.83**
12	Maidstone	-5.08	4.19	5.66	2.59	-0.45	**1.38**
13	Eston	10.33	12.38	14.83	3.41	3.91	**8.97**
14	Sezela	-9.50	-0.57	0.92	6.42	0.57	**-0.43**
15	Umzimkulu	0.53	2.10	0.25	2.06	0.26	**1.04**
Mean		**-3.97**	**6.14**	**4.43**	**1.74**	**-0.06**	
Industry		**-0.93**	**11.57**	**10.11**	**8.20**	**-0.47**	

8.4 Discussion and Conclusions

Depending on the time of forecast, the Canesim model-based sugarcane yield forecast system managed to capture between 11% and 58% of the natural seasonal variability in mean annual yields at an industry scale. The system also showed a significant capability to forecast whether yields in the forthcoming season could be expected to be higher or lower than in the previous season. Medium to high forecast skills were achieved at several mills located in different agro-climatic regions in South Africa. This emphasises the potential of the model-based system and prompts further research into mill areas, where forecast accuracies are currently low. In most cases, yield forecasts were poor for mills where substantial areas of cane were under irrigation. Evidence exists that several of these mills could be simulated more accurately if model inputs were to be diversified.

The evaluations performed in this chapter indicate favourable results and compare well with previous forecasting assessments. On an industry scale, R^2 values of actual *vs.* forecasted yields ranged from 0.48 to 0.82, depending on the time of forecast (*cf.* Figure 8.4). These compare favourably with results from Jury (1998), who achieved an R^2 value of 0.69 after using a statistical ocean-atmospheric response model. A reevaluation of the Eston mill results from Lumsden *et al.* (1999) revealed a σ_{ε} value of 19.9%, compared to 11.7% achieved in this study. Promburom *et al.* (2001) reported an accuracy of 4.8% when forecasting sugarcane production in Thailand. Their result is unfortunately not comparable with the evaluation parameters used in this study and generally the natural seasonal variability in sugarcane production in Thailand was found to be lower than in South Africa.

An evaluation in the number of raingauges per HCZ has shown that, under rainfed conditions, additional raingauges may increase simulation accuracy. Generally, it could be expected that a carefully selected number of representative additional raingauges with complete records will increase overall simulation accuracy. Further research is, however, required to determine guidelines towards the number and positioning of these additional raingauges within homogeneous climate zones.

Accurate and representative climate data play an important role in the accuracy of forecasts. An evaluation showed significant differences in accuracy between model simulations based on two different sets of climate data. Generally, climate data managed by SASRI produced more accurate forecasts for the Pongola, Entumeni, Darnall, Gledhow, Eston, Union Co-op and Noodsberg mills. In contrast, the Malelane and Umzimkulu mills were better represented by more generically derived climate data originating from temperature and rainfall databases housed at BEEH. A more detailed analysis also indicated that climate data originating from well managed climate stations with long data records often produced superior results to those from more generically derived data. This was evident even though established climate stations may often not be optimally located in order to represent a wider region.

Seasonal rainfall outlook information issued over the period 1998 to 2002 generally improved forecast accuracies. Rainfall outlooks issued in January increased the forecast skill for the industry by 11.6%. At the same time, forecast skills on a mill scale were, on average, increased by 6.1%. Rainfall outlooks became less valuable towards the drier winter period and were generally inferior to a neutral rainfall outlook assumption in September. It should be noted, however, that these results were only based on five years of information and that several advances in climate forecasting technology have been phased in over these years (*pers. comm.* Willem

Landman, SAWS, Pretoria, South Africa). The results do, however, indicate significant potential enhancements in yield forecasting capabilities and should encourage collaborative research between the sugar industry and seasonal climate forecasters.

Several additional issues arise once a model-based yield forecast system is implemented on an industry scale. Channels of information transfer, the communication of risk and communicating warnings and certain signals of concern need to be addressed. The Canesim model-based yield forecast system was used to generate official forecasts between 2000 and 2003. The following chapter will assess the accuracy and several other issues concerning these forecasts.

9 A Review of Historic Yield Forecasts and Related Information Transfer in the South African Sugar Industry

9.1 Introduction

In the previous chapter several evaluations of the Canesim sugarcane yield forecast system and some of its subcomponents were carried out. It was concluded that the system may be potentially valuable to decision making among stakeholders in the South African sugar industry. Several operational yield forecasts were issued between 2001 and 2003. This chapter reviews some aspects of these forecasts and compares the results with conventional Mill Group Board (MGB) forecasts.

Stern and Easterling (1999) noted that the effectiveness of forecast information depends on the systems that distribute the information, the channels used for distribution, the recipients' understanding and judgement concerning the information and the presentation thereof.

Mill Group Boards have been issuing forecasts of anticipated mill production (tons cane.an^{-1}) for their respective mills since the 1998 milling season. These forecasts are issued on a monthly basis starting at the end of March prior to the opening of the milling season and continuing until the milling season has closed (around December). The nature and some concerns of the forecasting process were discussed in Section 1.2.2. It was noted that MGBs were unable to quantify and attach a level of certainty to their forecasts, which according to several previous studies have been an important reason why forecasts often have failed (Thornton and Wilkes, 1998; Stern and Easterling, 1999; Hammer, 2000b; Hammer, 2000c). Also, no additional information from other forecasts was available to evaluate or confirm official MGB forecasts.

Strengths of the MGB forecasts are as follows:

- Monthly updates are normally carefully reviewed after feedback of actual production at the mill has been received. This technique is supported in the literature (Arkin and Dugas, 1981; Duchon, 1986; Bannayan and Crout, 1999) and results in accuracies that will asymptotically approach 100% towards the end of the milling season. Also, near real-time production information is used as opposed to climate data, the availability of which may lag up to six weeks. This concurs with the findings of Horie *et al.* (1992), who emphasised the importance for forecasts to utilise the most recently available actual data before switching over to probabilistic assumptions of future scenarios.

- Forecasts by MGBs are also viewed as the official source of anticipated production. Widespread information adoption hence exists among industry leaders, government and international marketers. The consolidated MGB forecast, for example, regularly appears as the official production forecast on the Internet home page of the South African Sugar Association (*cf.* www.sugar.org.za).

The aim of this chapter is to evaluate the operational Canesim sugarcane yield forecasts in light of the existing and more widely used MGB forecasts. Specific objectives are to (1) verify and compare the accuracies of MGB forecasts with Canesim-based forecasts, (2) review information transfer methodologies and (3) investigate the potential role of the Canesim model-based yield forecasts as a source of additional information in the South African sugar industry.

9.2 Methods

9.2.1 Operational Model-Based Forecasts of Sugarcane Production

Canesim model-based yield forecasts were issued at approximately two month intervals for the 2001, 2002 and 2003 milling seasons. In all these cases, simulations were carried out for both the current and the previous season. Simulations for the previous season were performed similarly to those described in Sections 8.2.3 and 8.2.4, where results were based exclusively on existing climate record data. Simulations for the current milling season were carried out using the SAWS rainfall outlook to select and simulate multiple future scenarios (similar to Section 8.2.5). Climate data originated solely from the SASRI climate station network (*cf.* Section 6.2.1) and various data quality checks were performed prior to the simulations.

The simulated yields for both seasons were subsequently aggregated to mill and industry scales (following approaches outlined in Section 5.3.4) and the forecast for the current season (EST_i) was expressed relatively to the previous season (EST_{i-1}). This relative value was used to project the known actual production figures from the previous season onto the current season. All forecasts were scrutinised by performing random spot checks and cross checks with previous simulations. Problems were often identified and were usually associated with missing or incorrect climate data. Once the forecast was validated, a report was sent out via email to milling companies, grower representatives, MGBs, extension officers, marketers and other industry leaders.

9.2.2 An Evaluation of Mill Group Board and Model-Based Forecasts

All MGB and model-based forecasts for mills and the industry were compared with actual mill production achieved at the end of the season (Y in $t.an^{-1}$). Accuracies were expressed in terms of a *RRMSE* (%, Eq. 9.1) and a *Skill* (%, Eq. 9.2).

$$RRMSE = \frac{\sqrt{\frac{1}{n}\sum_{y=1998}^{2003}\sum_{i=1}^{I}\left(Y_y - Est_{y,i}\right)^2}}{\overline{Y}} \qquad \text{Eq. 9.1}$$

where Y_y ($t.an^{-1}$) is actual annual production achieved at the end of season y, $Est_{y,i}$ ($t.an^{-1}$) is a forecast of production issued for season y some time before the end of the season, I is the number of forecasts that were issued during a specific time window (e.g. September to December), n is the total number of forecasts issued over the five year period within the specified time window and $\overline{Y}$ is the mean actual production over the five year period.

$$Skill = \left(1 - \frac{RRMSE}{CV_Y}\right) \times 100 \qquad \text{Eq. 9.2}$$

where CV_Y (%) is the coefficient of variance in annual production (Eq. 8.8).

Separate *RRMSE* and *Skill* values were calculated for MGB forecasts issued in different months in the year (March - December). In all these cases the value for I (Eq. 9.1) was always 1. Canesim-based yield forecasts were not issued every month and, based on their date of issue, forecasts were consequently grouped into four time windows relative to the milling season. Those included the periods September to December in the previous year, January to April, May to August and September to December. Values of *RRMSE* and *Skill* were calculated for each of these. Mill names were not disclosed as requested by the South African Sugar Association Industrial Affairs Division.

9.3 Results of Operational Forecast Accuracies

Tables 9.1 and 9.3 reflect *RRMSE* values for MGB and Canesim model-based forecasts at mill and industry scales, respectively. Tables 9.2 and 9.4 reflect the equivalent *Skill* values for the same forecasts. Mills were grouped into four sugar producing regions in South Africa, *viz.* Northern Irrigated and the KwaZulu-Natal North Coast, -South Coast and -Midlands regions.

Figure 9.1 displays time series over the period 1998 to 2003 of MGB and Canesim forecasts at an industry scale. These time series reflect how forecasts changed over time as each season progressed. It is evident that the consolidated MGB forecast for the industry (thick solid lines) significantly over-estimated the size of the crop in 1998, 1999 and 2001. In two of these cases (1999, 2001) MGBs seem to have assumed the next crop to be similar to the previous year's crop.

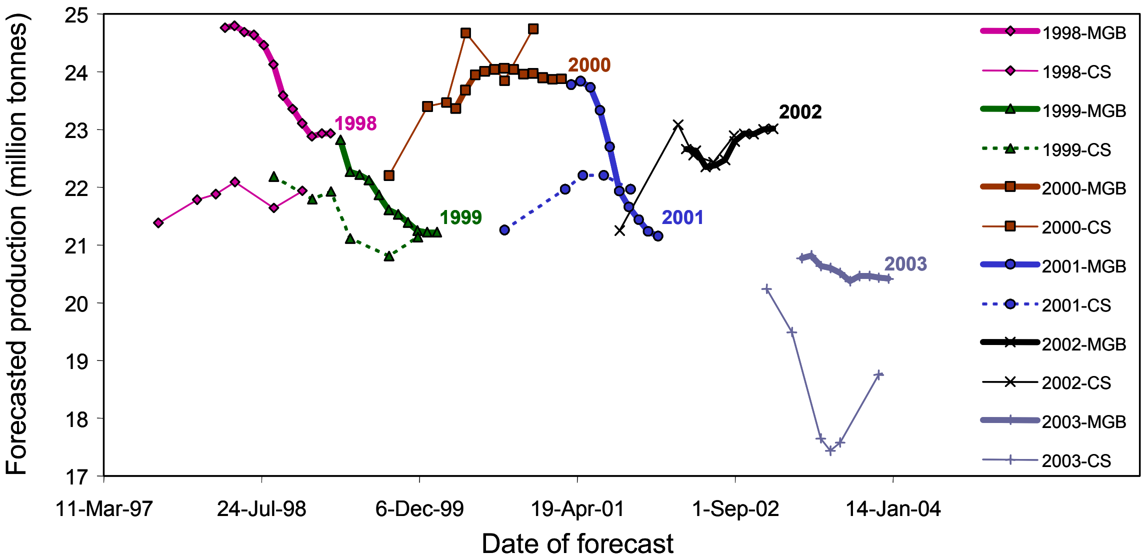

Figure 9.1 Time series of Mill Group Board forecasts (MGB, thick lines) and Canesim model-based forecasts (CS, thin and dotted lines) of total industry-scale sugarcane production in South Africa. All forecasts are plotted against the date when they were issued. The last forecast by MGBs for a season (indicated by the year's digit) may be assumed correct

Tables 9.1 and 9.2 confirm the asymptotical increase in accuracy of MGB forecasts towards the end of the season, which may be attributed to adjustments made by the MGBs to production information that had been fed back before the next forecast was issued. The consolidated MGB forecast for the industry displays no skill until July (*cf.* Table 9.2). Large differences in mean forecast skills of the MGB forecasts were observed between different mills over the 1998 – 2003 period (*cf.* Table 9.2). These ranged from no skill (7.8% at Mill C) to high skill (74.3% at Mill M). Further research may be required to determine why certain MGBs generally had the ability to forecast their production more accurately than others. Various factors could be expected to influence MGB forecast accuracies. These may include the following:

- Characteristics of typical crops grown within the mill supply area (e.g. age at harvest and cultivars);
- Homogeneity of soils, climate and topography within the mill supply area;
- The level of experience and skill of individual MGB members;
- The tools and resources that are used during the forecasting process (e.g. utilising the SAWS climate outlook);
- The total size of the crop and spatial expanse of the mill area;
- The spread of socio-economic and cultural diversity among growers and stakeholders; and
- Other institutional policies and procedures, such as communication, penalty systems and incentives.

Table 9.1 Relative Root Mean Square Error values (%) for Mill Group Board forecasts and for the consolidated industry forecast at different times of the milling season. All values were calculated over the period 1998 to 2003

Region	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Mean
Northern Irrigated											
Mill A	8.51	7.88	6.54	6.29	5.50	4.25	3.61	2.40	0.83	0.18	**4.60**
Mill B	4.10	4.83	4.47	4.76	4.37	4.19	3.06	2.33	1.15	1.10	**3.44**
Mill C	10.06	10.38	9.42	9.09	8.13	6.12	5.38	3.57	1.46	0.28	**6.39**
Mill D	12.84	13.81	8.36	8.71	8.37	6.44	3.41	1.71	0.64	0.23	**6.45**
KwaZulu-Natal Midlands											
Mill E	11.88	9.88	10.54	9.57	8.09	5.55	3.07	3.19	1.85	0.01	**6.36**
Mill F	5.96	6.37	6.63	5.45	4.25	2.90	1.47	1.32	1.27	0.96	**3.66**
Mill G	9.33	8.61	7.96	6.98	6.14	4.11	3.08	1.76	0.91	0.13	**4.90**
Mill H	11.46	10.80	9.63	8.15	6.57	4.53	2.79	1.21	0.49	0.22	**5.58**
KwaZulu-Natal North Coast											
Mill I	10.01	8.13	6.69	6.19	5.79	4.50	1.68	1.37	0.25	0.01	**4.46**
Mill J	11.62	9.55	9.49	9.71	8.01	4.80	1.54	1.00	0.17	0.03	**5.59**
Mill K	15.29	10.48	11.77	9.86	7.51	5.00	2.48	1.34	0.51	0.00	**6.42**
Mill L	8.88	9.25	9.26	9.34	8.45	5.34	3.44	1.97	1.58	0.32	**5.78**
Mill M	13.97	7.87	7.26	6.96	4.76	3.61	1.33	1.06	0.65	0.02	**4.75**
KwaZulu-Natal South Coast											
Mill N	6.89	6.72	6.35	6.35	6.29	5.25	2.24	2.18	2.00	1.24	**4.55**
Mill O	9.94	8.23	6.76	6.55	6.36	5.66	2.64	1.69	2.28	0.54	**5.07**
Mean	**10.05**	**8.85**	**8.07**	**7.60**	**6.57**	**4.82**	**2.75**	**1.87**	**1.07**	**0.35**	
National	**6.74**	**6.43**	**6.18**	**5.52**	**4.32**	**2.78**	**1.67**	**1.02**	**0.40**	**0.10**	**3.52**

Table 9.2 Forecast skills (%) of Mill Group Board and consolidated industry forecasts at different times of the milling season. All values were calculated over the period 1998 to 2003

Region	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Mean
Northern Irrigated											
Mill A	28.02	33.33	44.63	46.74	53.43	64.00	69.44	79.71	92.98	98.50	**61.08**
Mill B	34.05	22.36	28.15	23.42	29.72	32.62	50.85	62.48	81.48	82.35	**44.75**
Mill C	-45.13	-49.82	-35.96	-31.15	-17.33	11.69	22.41	48.50	78.92	95.92	**7.81**
Mill D	-43.29	-54.04	6.72	2.77	6.63	28.11	61.92	80.91	92.84	97.42	**28.00**
KwaZulu-Natal Midlands											
Mill E	-29.03	-7.27	-14.44	-4.00	12.12	39.71	66.69	65.35	79.88	99.86	**30.89**
Mill F	47.95	44.37	42.11	52.44	62.93	74.66	87.16	88.50	88.87	91.63	**68.06**
Mill G	37.08	41.91	46.35	52.95	58.62	72.27	79.24	88.16	93.84	99.11	**66.95**
Mill H	10.62	15.82	24.94	36.49	48.79	64.64	78.23	90.58	96.15	98.25	**56.45**
KwaZulu-Natal North Coast											
Mill I	35.81	47.85	57.08	60.31	62.88	71.13	89.23	91.20	98.36	99.93	**71.38**
Mill J	11.35	27.13	27.57	25.92	38.91	63.41	88.23	92.34	98.71	99.75	**57.33**
Mill K	-37.30	5.90	-5.69	11.46	32.59	55.15	77.71	87.95	95.41	99.99	**42.32**
Mill L	46.43	44.20	44.13	43.66	48.99	67.76	79.27	88.13	90.45	98.07	**65.11**
Mill M	24.39	57.42	60.70	62.36	74.25	80.44	92.82	94.28	96.46	99.90	**74.30**
KwaZulu-Natal South Coast											
Mill N	33.28	34.92	38.52	38.54	39.06	49.14	78.31	78.90	80.66	87.98	**55.93**
Mill O	3.38	20.02	34.34	36.33	38.20	44.99	74.38	83.59	77.81	94.80	**50.78**
Mean	**10.51**	**18.94**	**26.61**	**30.55**	**39.32**	**54.65**	**73.06**	**81.37**	**89.52**	**96.23**	
National	**-9.98**	**-4.97**	**-0.92**	**9.89**	**29.49**	**54.55**	**72.73**	**83.30**	**93.43**	**98.40**	**42.59**

Canesim model-based forecasts of production at mill scales were generally inferior to MGB forecasts, with forecast skills ranging between -114% for Mill E (KwaZulu-Natal Midlands) and 45% for Mill J (KwaZulu-Natal North Coast, *cf.* Table 9.4). It should be noted that forecast skills at mills situated in the KwaZulu-Natal Midlands were considerably lower than those of other regions. These results stand in contrast with those is Chapter 8 (*cf.* Table 8.2b), where it was argued that better operational forecast accuracies existed at mills where crops were harvested at older ages, *viz.* Union Co-op and Noodsberg. It should also be noted that during the January to April period model-based yield forecasts outperformed MGB forecasts at six of the 15 mills (*cf.* Table 9.4). Four of these mills originated from the Northern Irrigated region, while the other two were on the KwaZulu-Natal North Coast. Most of these results suggest that operational model-based forecasts of annual production were more reliable in areas where crops were harvested at a younger age. This could be owing to the fact that forecasted annual production was always expressed relative to the annual production of the previous season. This approach may be acceptable in regions such as the Northern Irrigated region, where most crops are harvested annually. However, in areas, such as the KwaZulu-Natal Midlands, where crops are often harvested at ages up to two years, this approach may become unsuccessful since it relates production in one year with areas not harvested during the previous year.

Table 9.3 Relative Root Mean Square Error (*RRMSE* in %) values for Canesim model-based forecasts of production over the period 1998 to 2003 at mill and industry scales. Values in brackets reflect mean *RRMSE* values of Mill Group Boards over the same time (derived from Table 9.1)

Region	Sep - Dec*	Jan - Apr		May - Aug		Sep - Dec		Mean
Northern Irrigated								
Mill A	12.36	10.25	(8.19)	20.14	(5.65)	21.44	(1.75)	**16.05**
Mill B	8.83	9.35	(4.46)	15.65	(4.45)	16.15	(1.91)	**12.49**
Mill C	9.66	7.16	(10.22)	9.04	(8.19)	13.77	(2.67)	**9.91**
Mill D	9.40	10.38	(13.33)	11.84	(7.97)	12.54	(1.50)	**11.04**
KwaZulu-Natal Midlands								
Mill E	14.02	10.74	(10.88)	10.85	(8.44)	9.58	(2.03)	**11.30**
Mill F	9.42	8.43	(6.17)	11.42	(4.81)	13.27	(1.26)	**10.63**
Mill G	22.96	20.16	(8.97)	14.20	(6.30)	19.17	(1.47)	**19.12**
Mill H	19.43	17.81	(11.13)	15.83	(7.22)	17.92	(1.18)	**17.75**
KwaZulu-Natal North Coast								
Mill I	7.29	9.28	(9.07)	7.04	(5.79)	8.49	(0.83)	**8.02**
Mill J	6.94	5.05	(10.58)	12.47	(8.00)	9.65	(0.69)	**8.53**
Mill K	14.37	11.24	(12.89)	12.07	(8.53)	9.97	(1.08)	**11.91**
Mill L	11.67	10.36	(9.06)	13.43	(8.10)	16.20	(1.83)	**12.91**
Mill M	14.48	11.05	(10.92)	13.90	(5.65)	11.64	(0.76)	**12.77**
KwaZulu-Natal South Coast								
Mill N	12.27	10.01	(6.81)	14.19	(6.06)	14.03	(1.92)	**12.62**
Mill O	16.12	12.00	(9.09)	7.61	(6.33)	11.70	(1.79)	**11.86**
Mean	**12.61**	**10.88**	**(9.45)**	**12.64**	**(6.77)**	**13.70**	**(1.51)**	
National	**6.15**	**3.31**	**(6.58)**	**7.64**	**(4.70)**	**4.24**	**(0.80)**	**5.34**

* Depicts forecasts issued during the year prior to the respective milling season

At an industry scale, the Canesim model-based forecasts of annual production were satisfactory during the early season. Model-based forecasts of industry production were of medium skill during the September to December period of the previous year and significantly outperformed MGB forecasts in the early season (63% *vs.* -7%). These forecasts may be valuable to stakeholders, such as international marketers, who require an early estimate of the entire industry's annual production. Later forecasts were, however, often less accurate than the consolidated MGB forecast and sometimes included erratic shifts between consecutive forecasts (*cf.* Figure 9.1). These shifts may be ascribed to irregularities in available climate data and also to changes in selected analogue seasons used to reflect the current seasonal rainfall outlook. It should also be noted that, while consecutive MGB forecasts asymptotically approached actual production, model-based forecasts remained inaccurate throughout the course of the milling season. This is as a result of the fact that production information to date was generally fed back to MGBs before the next forecast was issued. The same information is being made available to the modellers who perform model-based forecasts. However, no procedure currently exists to incorporate this information into the forecast.

Table 9.4 Forecast skill values for Canesim model-based forecasts of production over the period 1998 to 2003 at mill and industry scales. Values in brackets reflect equivalent values from Mill Group Board forecasts (derived from Table 9.2)

Region	Sep - Dec*	Jan - Apr		May - Aug		Sep - Dec		Mean
Northern Irrigated								
Mill A	29.12	41.25	(30.67)	-15.49	(52.20)	-22.92	(85.16)	**7.99**
Mill B	43.71	40.35	(28.21)	0.22	(28.48)	-3.03	(69.29)	**20.32**
Mill C	-18.22	12.33	(-47.47)	-10.64	(-18.19)	-68.59	(61.44)	**-21.28**
Mill D	-0.54	-10.98	(-48.67)	-26.66	(11.06)	-34.11	(83.27)	**-18.07**
KwaZulu-Natal Midlands								
Mill E	-113.69	-63.66	(-18.15)	-65.42	(8.35)	-45.99	(77.95)	**-72.19**
Mill F	1.29	11.63	(46.16)	-19.66	(58.03)	-39.15	(89.04)	**-11.47**
Mill G	-61.33	-41.65	(39.50)	0.19	(57.55)	-34.73	(90.09)	**-34.38**
Mill H	-63.12	-49.50	(13.22)	-32.93	(43.72)	-50.49	(90.80)	**-49.01**
KwaZulu-Natal North Coast								
Mill I	19.84	-2.06	(41.83)	22.58	(62.85)	6.58	(94.68)	**11.74**
Mill J	24.50	45.05	(19.24)	-35.74	(38.95)	-4.98	(94.76)	**7.21**
Mill K	-16.07	9.18	(-15.70)	2.54	(23.38)	19.50	(90.26)	**3.79**
Mill L	-38.75	-23.21	(45.32)	-59.68	(51.14)	-92.57	(88.98)	**-53.55**
Mill M	-65.90	-26.58	(40.90)	-59.19	(69.44)	-33.38	(95.87)	**-46.26**
KwaZulu-Natal South Coast								
Mill N	8.81	25.66	(34.10)	-5.39	(41.32)	-4.27	(81.46)	**6.20**
Mill O	-18.74	11.66	(11.70)	43.96	(38.46)	13.88	(82.65)	**12.69**
Mean	**-17.94**	**-1.37**	**(14.72)**	**-17.42**	**(37.78)**	**-26.28**	**(85.05)**	
National	**31.09**	**62.90**	**(-7.47)**	**14.46**	**(23.25)**	**52.56**	**(86.97)**	**40.25**

* Depicts forecasts issued during the year prior to the respective milling season

9.4 A Synthesis on Forecast Accuracies and Information Transfer

Bezuidenhout and Singels (2001) highlighted several problems in the operational Canesim yield forecast system, which included:

- A time lag of up to six weeks in climate data availability resulting from slow data mailing, capturing and processing procedures;
- Uncertainties regarding the accuracy of SAWS rainfall outlook information;
- A lack of data with adequate spatial coverage;
- Slow and inappropriate software and data formats; and
- Limited computational and human resources.

In addition to these problems, it should also be noted that owing to an irregular influx of climate data, different combinations of climate stations were often used to execute consecutive yield forecasts in a season. Also, some problems arise when expressing the forecasted yield for the current season relatively to the simulated yield of the previous season:

- First, it should be noted that differences between simulated yields of the two seasons are due solely to differences in climate and irrigation water regimes. These results may be incompatible at mill and industry scales where other factors, such as changes in the areas harvested, different lengths of milling seasons and different management practices could play important roles.
- Secondly, as shown in Section 8.3.2, it should be noted that there is still a certain amount of error in simulated results, even if actual climate data have been used to simulate the entire season. The forecast error will therefore be exacerbated by expressing a forecasted yield in relative terms, using a simulated yield of the previous season as a benchmark.

The timing of the Canesim yield forecasts may also inhibit the uptake of the information. Currently, data preparation and simulations are carried out once climate data have been received from collaborators, which occurs around the 15th of the following month. The final forecast is normally disseminated around the 22nd (five working days later). The SAWS seasonal climate outlook is usually issued on the 25th of each month, which results in a relatively dated rainfall outlook being used for the forecasts. It has, for example, occurred that the Canesim yield forecast was issued a day before the next month's SAWS climate outlook became available.

Stakeholder requirements should also be considered. In Chapter 3 (Fig. 3.3) it was established that stakeholders would need yield forecasts at the beginning of each month. Under the procedure described above, the most recent climate data of the last day of the previous month may be 15 days old and the rainfall outlook that was assumed is likely to be outdated by the time stakeholders use the information.

A considerable amount of growth takes place in the early season and a model-based forecast of this growth, guided by the rainfall outlook, seems feasible. Mill Group

Boards, on the other hand, tend to assess the current status of crops and their forecasts therefore become more accurate when growth is restricted by dry winter conditions and when large proportions of the crop have already been harvested. It is believed that the accuracies of model-based forecasts of annual production at both mill and industry scales could be significantly enhanced if the correct feedback mechanisms of production to date are incorporated into the system.

The author believes that a collaborative research approach between modellers and MGBs should be considered. De Lange and Singels (2003) have demonstrated some potential advantages of such an initiative. Canesim model-based forecasts of annual production and MGB forecasts can not be viewed as independent information sources since model-based forecasts have been disseminated to MGBs for their consideration since 2001. A collaborative attempt to forecast production could include growth models, large GIS-based databases of production and ground scouting information of yields, pest, diseases, sugarcane flowering, lodging and insufficient irrigation and ripening.

9.5 Conclusions

Operational Canesim model-based forecasts of annual production may be valuable to some mills and for the industry as a whole during the early season and during the period that MGB forecasts are unavailable (before March). Although some MGB forecasts seem more accurate than others, these forecasts generally had medium to high skills and outperformed model-based forecasts significantly, especially at a mill scale. Tentative information suggests that model-based forecasts may be more valuable if production were not expressed relative to that of the previous season, especially for mills where crop areas between consecutive seasons may differ substantially.

Several constraints have been identified that prohibit more efficient regimes of information transfer of model-based forecasts. These are related mainly to the timing of information. Forecasts depend on recent climate data, rainfall outlook information and time constraints in executing the simulations. These information sources are not synchronised and also do not currently allow for information transfer at the most optimum time for decision making.

Several synergies exist between MGB and model-based forecasts. Model-based simulations can be used to quantify the impacts of sequences of climate events on crops. Mill Group Boards, on the other hand, have the ability to readily capture and project current infield and production information. Some resources, such as large spatial databases, may currently be under-utilised by both MGBs and modellers and the author believes that a collaborative effort between these parties should yield the most suitable results.

This chapter concludes the results of this thesis. Chapter 10 contains the final conclusions and briefly highlights several areas that were identified for future research and consideration.

10 Conclusions and Recommendations

10.1 Main Conclusions

A crop growth model integrated with remote sensing technologies may provide highly valuable, if not *the* most suitable, forecasts of sugarcane yields and production for South Africa. Crop growth models have the ability to quantitatively translate information of recent climate into sugarcane yield responses, quantify impacts of probabilistic future climate scenarios and quantify risk and uncertainty. Remote sensing technologies, in addition, can be used to supply an economically feasible regional estimate of the current size of the crop and estimate the production area. Crop growth modelling and remote sensing in combination have, therefore, the ability to forecast future scenarios and aggregate information to larger areas. A sugarcane crop growth model, *viz.* Canesim, was selected for yield forecasts owing to its simplicity in algorithms as well as input data requirements and reasonable verification and representivity within the South African sugar producing areas. The adoption of remote sensing technologies in the South African sugar belt has been slow and remote sensing was, for this reason, excluded from the scope of this study.

Industry stakeholders generally envisaged that yield forecasts could significantly enhance decision making in the South African sugar industry. International marketing, national financing and mill operations are the probable areas to show immediate benefits from using yield forecast information. Yield forecasts are usually required at the commencement of each month and certain stakeholders may need forecasts of anticipated industry scale production by as soon as September prior to the commencement of the new milling season in April of the following year. Different stakeholders require updates of forecasts of sugarcane production at a range of different time intervals, spatial resolutions and for window periods within the milling season. Stakeholders may, however, need to acquire the correct decision making skills to optimally utilise the probabilistic information supplied through model-based forecasts. Several of these stakeholder requirements could be addressed by a Decision Support Program, that will not only filter out the most relevant information, but which may also assist with probabilistic decision making techniques.

Sugarcane producing areas in South Africa were subdivided into 48 relatively Homogeneous Climate Zones (HCZs). The uniformity within, and differences between, HCZs were verified and the HCZs could potentially assist researchers to extrapolate various experimental outcomes over wider areas. These may include agronomic recommendations and pest and disease advisories. Climatically, all zones had less than 10% internal variability, but differed by more than 18% from neighbouring zones. For this study, HCZ were used to spatially extrapolate point-based model simulation results.

This study provided a powerful and flexible crop growth simulation system that was used to evaluate several system components and information and data sources. The system's ability to provide forecasts of sugarcane production at climate zone, mill and industry scales was not only demonstrated, but was also operationally implemented in South Africa for several years. The study provided the first system, to the author's knowledge, that produced model-based operational forecasts of sugarcane production at an industry scale for one of the 15 large sugarcane producing countries in the world.

At an industry scale, the system could manage to capture up to 58% of the climatically driven variability in mean annual sugarcane yields. Accuracies compared well with results from previous literature and the system showed a significant capability to forecast whether yields in the approaching season could be expected to be higher or lower than those of the previous season. The reliability of forecasts, both at mill and industry scales, changed with time as the milling season progressed. Forecast accuracies differed widely between different mills and several factors, such as data quality from climate stations and the number of additional raingauges within HCZs, were identified and which may explain some of these inconsistencies. Generally, it was concluded that, should sufficient information and climate data exist, the system has the ability to favourably forecast yields over a wide range of agro-climatic conditions.

Climate forecasts contain potentially valuable information and should be consulted when forecasts of crop production are made. For this study, a methodology was developed that quantitatively translated a three month lead time rainfall outlook into anticipated sugarcane yield responses. Rainfall outlook information for the February to April lead time increased the accuracy of industry scale forecasts of mean annual yields by 11.6% during the 1998 – 2002 period. Weaker responses existed for rainfall outlooks issued after February and outlook information issued in September was generally not of any value within the system. It should, however, be emphasised that valuable progress in the adoption of seasonal climate outlook information could be expected if climate forecasters and crop modellers increased collaboration.

Several new issues, which were often omitted in more theoretical studies in the literature, come to the forefront when attempts are made to provide stakeholders with operational yield forecasts. These include:

- Time constraints, such as fixed times when seasonal climate outlooks are issued and when climate station data are processed;
- Conveying the forecast in logical and usable terms, such as expressing yields relatively to the previous season; and
- Comparing forecasts with existing and more conventional ones already available in the industry.

Operational forecasts of production at mill and industry scales were issued between 2001 and 2003. In addition, forecasts for 1998 to 2000 were emulated in order to compare accuracies with conventional Mill Group Board (MGB) forecasts that were issued between 1998 and 2003. Model-based forecasts outperformed those of MGBs at some mills, mainly those in the Northern irrigated region, during the early season (Jan – Apr). Forecasts aimed at an industry wide scale also significantly outperformed early season MGB forecasts.

Operationally, the system still has some limitations. Forecasts of sugarcane production, for example, are always expressed relative to yields of the previous season, potentially ignoring non-climatic changes between seasons. Also, near real-time production information from mills, which can be used to correct certain errors, are not currently fed back into the system. The author believes that enough synergy exists between modellers and MGBs to encourage a collaborative effort to enhance forecasts of sugarcane yields and production within the sugar belt of South Africa.

10.2 Recommendations for Future Research

The following list of recommendations for future research and refinements to the current Canesim model-based yield forecast system has been compiled. These are not in any order of importance:

GENERAL

- A multidisciplinary research strategy could be expected to enhance research outcomes. It is believed that different skills offered by modellers, statisticians, crop physiologists, GIS experts, experts in management, climatologists and experts in communication can significantly contribute to this field of study.

- Based on Figure 1.1, it is important to identify why the South African sugar industry is vulnerable to yield losses and to make strategic (*i.e.* long-term), tactical (*i.e.* seasonal) and operational (*i.e.* daily and weekly) mitigation plans against such vulnerabilities. Small scale growers, for example, seem more vulnerable to yield losses than commercial growers, but also seem more limited in their capacity to gain from operational forecasts.

- Collaboration between modellers and Mill Group Boards is essential. This should stimulate new ideas and will also streamline information flow. Further research is needed to establish why certain MBGs generally achieve more accurate forecasts than others. Accuracy may simply be dictated by the complexity, size and diversity of the respective mill areas, but may also include other issues, such as tools employed to make the forecast, the level of experience of board members and other organisationally and politically related issues, such as institutional policies and penalty systems. Similar to forecasts made by MGBs, there is also a need for the Canesim model-based system to include near real-time production feedback from mills.

- More research is needed to improve model aggregation errors. Models, for example, can be calibrated to local conditions. This, however, may be suboptimal with a strong bias towards local historic information that may not be relevant to new practices, cultivars and production areas. Alternatively, remote sensing technologies may be employed. These can be used to assess the crop's status, calculate areas under cane, determine pest and disease infestations and highlight areas under water and temperature stress.

- It has been shown that expressing forecasted yields relative to the previous season may introduce additional errors. Further refinements to this approach are needed and these may include local model calibration, using statistical indices and accounting for non-climate related changes in year-to-year production.

- The Canesim yield model and input variables require further development and expansion to account for different cultivars, sugarcane flowering and impacts of pests (e.g. Horton *et al.*, 2002) and diseases. In contrast to the CANEGRO model (*cf.* Singels and Bezuidenhout, 2002), the Canesim model does not currently simulate the dynamics of sucrose accumulation. The model also lacks information on fibre and reduced sucrose contents. In Chapter 7 it was concluded that the simulation of irrigated crops may need further refinement. These should include scenarios with water restrictions. In addition, the initial soil moisture content at the start of each crop, which currently does not vary between years, may also need to be varied based on conditions prior to the crop's initialisation. Some results from this study suggest that the description of heterogeneity among crops may still be limited. Further research is needed to establish the optimal number of input variables needed to describe the diversity in soils, management and climate. It was shown, for example, that the number of raingauges providing data used to represent heterogeneity within climate zones may play an important role.

- The Canesim model needs more extensive verification. Verifications of simulated cane yields have not been published and more confidence in the model is needed for different locations in the sugar belt, such as in the KwaZulu-Natal midlands, and during seasons with different climate regimes.

CLIMATE

- A strong emphasis needs to be laid on closer collaboration with climate forecasters. Climatologists provide expertise in General Circulation Modelling and their subsequent downscaling to reflect regional climate patterns. Closer collaboration may also result in the provision of customised and correctly timed seasonal rainfall outlook information. Climate forecasters are likely to soon be able to forecast changes in the frequencies of rainy days, heavy rainstorms and duration of dry spells. Such forecasts, as opposed to those of only three month rainfall totals, may be more valuable to crop yield modellers, since they provide a better explanation of probable changes in available soil water regimes. The evaluations on the value of rainfall outlook information in this study (*cf.* Section 8.3.4) were only based on five years of data and further research is needed to confirm the value of seasonal rainfall outlooks within the Canesim model-based yield forecast system.

- There is also a need to refine the selection criteria for analogue seasons. Strong indications exist that not enough analogue seasons are currently simulated. In addition, other outlook information, such as a one-month rather than only three-months rainfall outlooks as well as temperature forecasts, are currently already available for South Africa and need to be incorporated into the system.

- It was shown that the current network of reporting climate stations is likely to be sub-optimal (*cf.* Section 8.3.2). This was confirmed in several areas where climate surrogates derived better results compared to data from nearby stations situated in other climate zones. The climate station network and additional raingauge network need to be assessed and a methodology needs to be developed to establish optimum locations and densities for these stations. At the same time it should also be ensured that new stations and existing stations measure all the required climatic variables, including solar radiation. Climate data communication also needs to be accelerated and near-real time climate data integrity checks should be performed.

- Integrated research strategies are needed to address issues of climate change. Generally, it is believed that higher climatic variability may be expected under climate change scenarios owing to the increased energy levels within the atmosphere. This will increase crop production vulnerabilities and will place a higher demand on the necessity to acturately forecast crop responses in advance.

MANAGEMENT

- There is a need to train decision makers on how to make management decisions under risk and uncertainty. Forecasts should be made understandable without sacrificing important information on risk and uncertainties. The level of confidence associated with each forecast needs to be conveyed. This should not only include an expression of future climate uncertainty, but also the magnitude of input, model and aggregation errors. At the same time, successive forecasts should not differ substantially from one another without an explanation being provided. Decision support tools, as well as specialist assistance are needed to help decision makers with forecast interpretations.

The above-mentioned recommendations, in conjunction with the research conveyed in this study, emphasises the importance of a multi-disciplinary research approach to readily and accurately forecast climate variability with suitable lead times, correctly translate these into yield responses, quantify the associated risks and mitigate against these by using capable decision makers to implement alternative plans.

11 References

Addiscott, T. M. 1993. Simulation modelling and soil behaviour. *Geoderma* 60: 15-40.

Ahmadi, R., Baier, G. J., Shipley, D. J. 2000. Improvement of cane quality at Ubombo sugar. In *Proceedings of South African Sugar Technologists' Association Workshop on Farming for Recoverable Value*, Mt. Edgecombe, South Africa, 13 June 2000.

Allen, R. G., Pereira, L. S., Raes, D., Smith, M. 1998. *Crop evapotranspiration - Guidelines for computing crop water requirements*. FAO Irrigation and Drainage Paper 56. 301p, FAO, Rome, Italy.

Alley, W. M. 1984. The Palmer Drought Severity Index: limitations and assumptions. *Journal of Climate and Applied Meteorology* 23: 1100-1109.

Anon. 1996. *Gloabal Information and Early Warning System (GIEWS) - Guidelines for crop and food supply assessment missions*. Food and Agriculture Organization of the United Nations (FAO), Rome, Italy.

Anon. 1997. *The precision-farming guide for agriculturists*. Ed. J. R. Kuhar. ISBN 0-86691-245-2, John Deere Publishing. Moline, USA, 117p.

Anon. 1999a. *Climate research for Africa*. 16/1999, ICPO Publication Series No. 29. 83p, World Climate Research Programme, Southampton, UK.

Anon., 1999b. Sugar first with across industry approach. In *CLIMAG - Newsletter of the Climate Variability in Agriculture R&D program (CVAP)*, July 1999, Issue 2, ISSN: 1441-7987, Indooroopilly, Australia.

Antony, G., Everingham, Y. L., Smith, M. 2002. Financial benefits of using climate forecasting: A case study. In *Proceedings of Australian Society of Sugar Cane Technologists*, 153-159. Cairns, Australia, 29 April - 2 May 2002.

Arkin, G. F., Dugas, W. A. 1981. Making weather and climate dependent crop management decisions. In *Proceedings of the Workshop on Computational Techniques of Meteorological Data Applications for Problems in Agriculture and Forestry.* Ed. A. Weiss. 223-237, Anaheim, Canada, 30-31 Marchpp, American Meteorological Society, Boston, USA.

Arkin, P. A., Meisner, B. N. 1987. The relationship between large-scale convective rainfall and cold cloud over the Western Hemisphere during 1982 - 84. *Monthly Weather Review* 115: 51-74.

Ba, M. B., Nicholson, S. E. 2001. Satellite-derived surface radiation budget over the African continent. Part II: Climatologies of the various components. *Journal of Climate* 14: 60-76.

Banitz, E. 2001. Evaluation of short-term weather forecasts in South Africa. *Water SA* 27 (4): 489-498.

Bannayan, M., Crout, N. M. J. 1999. A stochastic modelling approach for real-time forecasting of winter wheat yield. *Field Crops Research* 62 (1): 85-95.

Bates, B. C., Charles, S. P., Huges, J. P. 2000. Stochastic down-scaling of general circulation model simulations. In *Applications of Seasonal Climate Forecasting in Agriculture and Natural Ecosystems - The Australian Experience.* Eds. G.L. Hammer, N. Nicholls, C. Mitchell: ISBN 0-7923-6270-5, Kluwer Academic Publishers, Dordrecht, Netherlands.

Bezuidenhout, C. N. 2000. *A model review and proposed mechanistic tiller model for the CANEGRO sugarcane crop model*. M Tech Engineering thesis. Centre for Engineering Research, Technikon Natal, Durban, South Africa.

Bezuidenhout, C. N., Gers, C. 2002. Homogeneous Climate Zones for the South African Sugar Industry: Preliminary boundaries. In *Proceedings of the 76th South African Sugar Technologists' Association Congress*, 601-605. Durban, South Africa.

Bezuidenhout, C. N., Singels, A. 2001. The use of simulation crop modelling to forecast sugarcane yield. In *Proceedings of the SASTA Workshop on Burn/Harvest to Crush Delays & Crop Estimating*, South African Sugar Technologists' Association, 20-29. Mt. Edgecombe, South Africa, 8 November 2001.

Bland, W. L., Clayton, M. K. 1994. Spatial structure of solar radiation in Wisconsin. *Agricultural and Forest Meteorology* 69: 75-84.

Boughton, W. C. 1981. Rainfall variability in rainfall-runoff modelling. *Civil Engineering Transcript, Australia* 23: 68-73.

Boyer, D. G., Feldhake, C. M. 1994. Identification of thermally homogeneous subunits in a steep Appallachian pasture. *Journal of Applied Meteorology* 33: 1200-1209.

Bristow, K. L., Campbell, G. S. 1984. On the relationship between incoming solar radiation and daily maximum and minimum temperature. *Agricultural and Forest Meteorology* 31: 159-166.

Calder, I. R. 1992. Hydrologic effects of land-use change. In *Handbook of Hydrology.* Ed. D. R. Maidment: McGraw-Hill, Inc., New York, USA.

Camp, K. 1999. *Bioresource classification of KwaZulu-Natal*. M Sc dissertation, University of Natal, Pietermaritzburg, South Africa.

Cane, M. A. 1999. Current capabilities in long-term weather forecasting for agricultural purposes. In *Proceedings of the International Workshop on Climate Change in Agriculture*, 24p. Geneva, Switzerland, 27-29 September 1999.

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

Cheeroo-Nayamuth, F. B., Bezuidenhout, C. N., Kiker, G. A., Nayamuth, A. R. H. 2003. Validation of CANEGRO-DSSAT v3.5 for contrasting sugarcane varieties in Mauritius. In *Proceedings of the 77th South African Sugar Technologists' Association Congress*, 601-604. Durban, South Africa.

Cheeroo-Nayamuth, F. C., Robertson, M. J., Wegener, M. K., Nayamuth, A. R. H. 2000. Using a simulation model to assess potential and attainable sugar cane yield in Mauritius. *Field Crops Research* 66: 225-243.

Chetty, K., Smithers, J. C., Schulze, R. E. 2003. Towards a continuous simulation modelling approach for design flood estimation in South Africa. In *Proceedings of the 2nd IWRM Symposium*, 16p. Stellenbosch, South Africa.

Chipanshi, A. C., Ripley, E. A., Lawford, R. G. 1998. Large-scale simulation of wheat yields in a semi-arid environment using a crop-growth model. *Agricultural Systems* 59: 1-10.

Chmielewski, F. M., Potts, J. M. 1995. The relationship between crop yields from an experiment in southern England and long-term climate variations. *Agricultural and Forest Meteorology* 73: 43-66.

Comrie, A. C., Glenn, E. C. 1998. Pricipal components-based regionalization of precipitation regimes across the southwest United States and northern Mexico, with an application to monsoon precipitation variability. *Climate Research* 10: 201-215.

Day, W. 2001. Modelling crop physiology for integrated decision making. *Annals of Applied Biology* 138: 215-219.

de Jager, J. M., Potgieter, A. B., 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.

de Lange, J., Singels, A. 2003. Using an Internet based crop model to assist with crop estimation for the Umfolozi mill supply area. In *Proceedings of the 77th South African Sugar Technologists' Association Congress*, 592-595. Durban, South Africa.

de Wit, C. T., van Keulen, H. 1987. Modelling production of field crops and its requirements. *Geoderma* 40: 253-265.

Dent, M. C., Schulze, R. E., Angus, G. R. 1989. *Crop water requirements, deficit and water yield for irrigation planning in southern Africa*. WRC Report No. 118/1/88. 183p, Water Research Commission, Pretoria, South Africa.

Downing, T. E., Washington, R. 1997. Seasonal forecasting of African rainfall: Prediction, responses and household food security. In *Proceedings of Climate Variability Prediction, Water Resources and Agricultural Productivity: Food Security in Tropical Sub-Saharan Africa*, 16-48. Cotonou, Benin, 22-25 July 1997.

Duchon, C. E. 1986. Corn yield prediction using climatology. *Journal of Climate Applications and Meteorology* 25 (5): 581-590.

Dyson, L. L., van Heerden, J., Marx, H. G. 2002. *Short term weather forecasting techniques for heavy rainfall*. Report No 1011/1/02. Water Research Commission, Pretoria, South Africa.

Eakin, H. 2000. Smallholder maize production and climatic risk: A case study from Mexico. *Climate change* 45: 19-36.

Everingham, Y. L., Inman-Bamber, N. G., Smith, D. M. 2002a. Seasonal climate forecasts to enhance decision-making across the sugar industry value chain. In *Proceedings of the Australian Society for Sugar Cane Technologists*, 67-74. Cairns, Australia, 29 April - 2 May 2002.

Everingham, Y. L., Muchow, R. C., Stone, R. C., Inman-Bamber, N. G., Singels, A., Bezuidenhout, C. N. 2002b. Enhanced risk management and decision-making capability across the sugarcane industry value chain based on seasonal climate forecasts. *Agricultural Systems* 74: 459-477.

Fischhoff, B. 1994. What forecasts (seem to) mean. *International Journal of Forecasting* 10: 387-403.

Gadekar, M. D. 1998. *Crop production forecasting & its use in production planning & control - A study for groundnut*. 5p, ActionAid India, New Delhi, India.

Gardner, B. R., Blad, B. L., Garrity, D. P., 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.

Gelb, A., Kasper, J., Nash, R., Price, C. 1974. *Applied Optimal Estimation*. ISBN 0-262-57048-3, The MIT Press, Cambridge, USA. 374p.

Georgiev, G. A., Hoogenboom, G. 1999. Near real-time agricultural simulations on the web. *Simulation* 73 (1): 22-28.

Gers, C. 2003. Relating remotely sensed multi-temporal Landsat 7 ETM + imagery to sugarcane characteristics. In *Proceedings of the 77th South African Sugar Technologists' Association Congress*, 313-321. Durban, South Africa.

Gers, C., Schmidt, E. J., Bezuidenhout, C. N. 2001. Validation of the CANESIM sugarcane growth model using GIS within the South African sugar industry. In *Proceedings of the 21st Annual ESRI International User Conference*, Paper 416. San Diego, USA, 9-13 July 2001.

Gertner, G. Z., Guan, B. T. 1991. Using an error budget to evaluate the importance of component models within a large scale simulation model. In *Proceedings of*

Mathematical Modelling of Forest Ecosystems, 62-74. Frankfurt am Main, Germany.

Górski, T., Górska, K., 2003. The effects of scale on crop yield variability. *Agricultural Systems* 78: 425-434.

Gungula, D. T., Kling, J. G., Togun, A. O. 2003. CERES-Maize predictions of maize phenology under nitrogen-stressed conditions in Nigeria. *Agronomy Journal* 95: 892-899.

Hammer, G. L. 2000a. Applying seasonal climate forecasts in agricultural and natural ecosystems - a systhesis. In *Applications of Seasonal Climate Forecasting in Agriculture and Natural Ecosystems - The Australian Experience.* Eds. G.L. Hammer, N. Nicholls, C. Mitchell: ISBN 0-7923-6270-5, Kluwer Academic Publishers, Dordrecht, Netherlands.

Hammer, G. L. 2000b. A general systems approach to applying seasonal climate forecasts. In *Applications of Seasonal Climate Forecasting in Agriculture and Natural Ecosystems - The Australian Experience.* Eds. G.L. Hammer, N. Nicholls, C. Mitchell: ISBN 0-7923-6270-5, Kluwer Academic Publishers, Dordrecht, Netherlands.

Hammer, G. L., Hansen, J. W., Phillips, J. G., Mjelde, J. W., Hill, H., Love, A., Potgieter, A. 2001. Advances in the application of climate prediction in agriculture. *Agricultural Systems* 70: 515-553.

Hansen, J. W. 1999. Stochastic daily solar irradiance for biological modeling applications. *Agricultural and Forest Meteorology* 94: 53-63.

Hansen, J. W. 2002. Realizing the potential benefits of climate prediction to agriculture: issues, approaches, challenges. *Agricultural Systems* 74 (3): 309-330.

Hansen, J. W., Hodges, A. W., Jones, J. W. 1998. ENSO influences on agriculture in the Southeastern United States. *Journal of Climate* 11 (3): 404-411.

Hansen, J. W., Jones, J. W. 1999. Scaling-up crop models for climate prediction applications. In *Proceedings of the START CLIMAG Geneva Workshop*, 52. Geneva, Switzerland, 27-30 September 1999.

Hansen, J. W., Jones, J. W. 2000. Scaling-up crop models for climate variability applications. *Agricultural Systems* 65 (1): 43-72.

Hansen, J. W., Jones, J. W., Irmak, A., Royce, F. 2001. *El Niño-Southern Oscillation impacts on crop production in the southeast United States*. Special publication No. 63: 55-76. American Society of Agronomy,

Hargreaves, G. L., Samani, Z. A. 1985. Reference crop evapotranspiration from temperature. *Transactions of the American Society of Agricultural Engineers* 1: 96-99.

Haskett, J. D., Pachepsky, Y. A., Acock, B. 1995. Use of the beta distribution for parameterizing variability of soil properties at the regional level for crop yield estimation. *Agricultural Systems* 48: 73-86.

Hayes, M. J. 2001. *Drought indices - a technical comparison of drought indices*. 11p, National Drought Mitigation Center, Lincoln, USA.

Heuvelink, G. B. M. 1998. Uncertainty analysis in environmental modelling under a change of spatial scale. *Nutrient Cycling in Agroecosystems* 50: 255-264.

Hewitson, B. C., Crane, R. G. 1996. Climate downscaling: techniques and application. *Climate Research* 7: 85-95.

Hildebrandt, Q. L. 1998. Maximising profits from sugar through optimising the length of the milling season. In *Proceedings of the Annual General Meeting of the South African Sugar Agronomists Association*, 1-6. Mt. Edgecombe, South Africa.

Hodges, T., Botner, D., Sakamoto, C., Hays Haug., J. 1987. Using the CERES-Maize model to estimate production for the U.S. cornbelt. *Agricultural and Forest Meteorology* 40: 293-303.

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

Horton, P. M., Hearne, J. W., Apaloo, J., Conlong, D. E., Way, M. J. 2002. Investigating strategies for minimising damage caused by the sugarcane pest Eldana saccharina. *Agricultural Systems* 74 (2): 271-286.

Hughes, A. D. 1992. *Sugarcane yield simulation with the ACRU Modelling System*. M Sc dissertation. Department of Agricultural Engineering, University of Natal, Pietermaritzburg, South Africa.

Hunt, L. A., Kuchar, L., Swanton, C. J. 1998. Estimation of solar radiation for use in crop modelling. *Agricultural and Forest Meteorology* 91: 293-300.

Inman-Bamber, N. G. 1991. A growth model for sugar-cane based on a simple carbon balance and the CERES-Maize water balance. *South African Journal of Plant and Soil* 8 (2): 93-99.

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. In *Proceedings of the 69th South African Sugar Technologists' Association Congress*, 55-59. Durban, South Africa.

Inman-Bamber, N. G., Kiker, G. A. 1997. CANEGRO 3.10 - DSSAT, 1998 distribution software, version 3.1. IBSNAT. Honolulu, USA.

Isaacs, Z. 2003. Sugar growers, millers & refiners. *EngineeringNews*, Vol 23(20), p.50-62, May 2003.

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

Jones, C. A., Kiniry, J. R. 1986. *CERES-Maize: A simulation model of maize growth and development*. ISBN 0-89096-269-3, Texas A&M University Press, College Station, USA. 194p.

Jury, M. R. 1998. Statistical analyses and prediction of KwaZulu-Natal climate. *Theoretical & Applied Climatology* 60: 1-10.

Karl, T. R., Knight, R. W. 1985. *Atlas of monthly Palmer Hydrological Drought indices (1931-1983) for the contiguous United States*. Historical Climatology Series 3-7. National Climatic Data Center, Asheville, USA.

Keating, B. A., Robertson, M. J., Muchow, R. C., Huth, N. I. 1999. Modelling sugarcane production systems I. Development and performance of the sugarcane module. *Field Crops Research* 61: 253-271.

Kienzle, S. W., Lorentz, S. A., Schulze, R. E. 1997. *Hydrology and water quality of the Mgeni catchment.* Report TT87/97. 88p, Water Research Commission, Pretoria, South Africa.

King, A. W. 1991. Translating models across scales in the landscape. In *Quantitative Methods in Landscape Ecology: The Analysis and Interpretation of Landscape Heterogeneity.* Eds. M. G. Turner, R. H. Gardner: 479-517pp, Springer-Verlag, New York, USA.

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

Kuhnel, I. 1994. Relationship between the southern oscillation index and Australian sugarcane yields. *Australian Journal of Agricultural Research* 45: 1557-1568.

Landman, W. A., Mason, S. J. 1999. Operational long-lead prediction of South African rainfall using canonical correlation analysis. *International Journal of Climatology* 19: 1073-1090.

Landman, W. A., Mason, S. J., Tyson, P. D., Tennant, W. J. 2001. Retro-active skill of multi-tiered forecasts of summer rainfall over southern Africa. *International Journal of Climatology* 21: 1-19.

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

Liu, D. L., Scott, B. J. 2001. Estimation of solar radiation in Australia from rainfall and temperature observations. *Agricultural and Forest Meteorology* 106 (1): 41-59.

Lumsden, T. G. 2000. *Development and evaluation of a sugarcane yield forecasting system*. M Sc dissertation. School of Bioresources Engineering and Environmental Hydrology, University of Natal, Pietermaritzburg, South Africa.

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

Lumsden, T. G., Schulze, R. E., Lecler, N. L., Schmidt, E. J. 1999. *An assessment of the potential for sugarcane yield forecasting using seasonal rainfall forecasts and crop yield models*. Internal Report. 94p, South African Sugar Association, Mt. Edgecombe, South Africa.

Luxmoore, R. J., King, A. W., Tharp, M. L. 1991. Approaches to scaling up physiologically based soil-plant models in space and time. *Tree Physiology* 9: 281-292.

Lynch, S. D. 2004. *Development of a raster database of annual, monthly and daily rainfall for Southern Africa.* WRC Report 1156/1/03. 78p, Water Research Commission, Pretoria, South Africa.

Maas, S. J. 1988. Using satellite data to improve model estimates of crop yield. *Agronomy Journal* 80: 655-662.

Mains, L. A. 1996. An experimental examination of subjective forecast combination. *International Journal of Forecasting* 12: 223-233.

Marquis, D. L., Ray, D. E. 1981. Use of stochastic simulation to value improved crop forecast information. In *Proceedings of the American Agricultural Economics Association Meeting*, Atlanta, USA, July 1981.

Maselli, F., Conese, C., Petkov, L., Gilabert, M. A. 1993. Environmental monitoring and crop forecasting in the Sahel through the use of NOAA NDVI data. A case study: Niger 1986-1989. *International Journal of Remote Sensing* 14: 3471-3487.

Mason, S. J. 2000. *Definition of technical terms in forecast verification and examples of forecast verification scores*. IRI-CW/01/1, 6-9. International Research Institute for Climate Prediction, Palisades, USA.

Matthews, R. B., Stephens, W. 2002. *Crop-Soil Models: Applications in Developing Countries*. ISBN: 0851995632, CAB International, Wallingford, UK. 304p.

Matthews, R. B., Stephens, W., Hess, T., Mason, T., Graves, A. 2000. *Applications of crop/soil simulation models in developing countries*. PD82. 173p, Cranfield University Institute of water and environment, Silsoe, UK.

McGlinchey, M. G. 1999. Computer crop model applications: Developments in Swaziland. In *Proceedings of the 73th South African Sugar Technologists' Association Congress*, 35-38. Durban, South Africa.

McGlinchey, M. G., Inman-Bamber, N. G. 1996. Predicting sugarcane water use with the Penman-Monteith equation. In *Proceedings of Evapotranspiration and Irrigation Scheduling*, American Society of Agricultural Engineers, 592-597. San Antonio, USA.

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

Mearns, L. O., Giorgi, F., McDaniel, L., Shields, C. 1995. Analysis of daily variability of precipitation in a nested regional climate model: Comparison with observations and doubled CO_2 results. *Global and Planetary Change* 10: 55-78.

Mearns, L. O., Rosenzweig, C., Goldberg, R. 1992. The effect of changes in interannual climatic variability on CERES-Wheat yields: Sensitivity and $2\times CO_2$ studies. *Agricultural and Forest Meteorology* 62: 159-189.

Meinke, H., Baethgen, W. E., Carberry, P. S., Donatelli, M., Hammer, G. L., Selvaraju, R., Stöckle, C. O. 2001. Increasing profits and reducing risks in crop production using participatory systems simulation approaches. *Agricultural Systems* 70: 493-513.

Meinke, H., Hammer, G. L. 1995. Climatic risk to peanut production: A simulation study for Northern Australia. *Australian Journal of Experimental Agriculture* 35: 777-780.

Meinke, H., Hammer, G. L. 1997. Forecasting regional crop production using SOI phases: An example for the Australian peanut industry. *Australian Journal of Agricultural Research* 48: 789-793.

Meinke, H., Hochman, Z. 2000. seasonal climate forecasts to manage dryland crops in Northern Australia - experiences from the 1997/98 seasons. In *Applications of Seasonal Climate Forecasting in Agriculture and Natural Ecosystems - The Australian Experience.* Eds. G.L. Hammer, N. Nicholls, C. Mitchell: ISBN 0-7923-6270-5, Kluwer Academic Publishers, Dordrecht, Netherlands.

Meyer, J. H. 1984. An integrated system for soil identification in the South African sugar industry. In *Proceedings of the 58th South African Sugar Technologists' Association Congress*, 184-191. Durban, South Africa.

Meyer, J. H., Wood, R. A. 1990. Soils of the Eastern Transvaal sugar industry and some aspects of their management. In *Proceedings of the 64th South African Sugar Technologists' Association Congress*, 11-16. Durban, South Africa.

Meyer, S. J., Hubbard, K. G., Wilhite, D. A. 1991. The relationship of climatic indices and variables to corn (maize) yields: A principal components analysis. *Agricultural and Forest Meteorology* 55: 59-84.

Midgley, D. C., Pitman, W. V., Middleton, B. J. 1994. *The Surface Water Resources of South Africa*. Report Numbers 298/1.1/94 to 298/6.1/94 (text) and 298/1.2/94 to 298/6.2/94 (maps). Water Research Commission, Pretoria, South Africa.

Moen, T. N., Kaiser, H. M., Riha, S. J. 1994. Regional yield estimation using a crop simulation model: concepts, methods and validation. *Agricultural Systems* 46: 79-92.

Moreno, C. A., Betancourt, E. C., Gonzalez, J. C. 2001. Determining of an index for estimating the effect of climate on the sugar cane crop. In *Proceedings of the Congress of the 24th International Sugar Cane Technologists Association*, 101-106. Brisbane, Australia.

Müller, F. 1992. Hierarchical approaches to ecosystem theory. *Ecological Modelling* 63: 215-242.

Murphy, A. H. 1993. What is a good forecast? An essay on the nature of goodness in weather forecasting. *Weather and Forecasting* 8: 281-293.

Naylor, R., Falcon, W. P., Wada, N., Rochberg, D. 2002. Using El Niño-Southern Oscillation climate data to improve food policy planning in Indonesia. *Bulletin of Indonesian Economic Studies* 38 (1): 75-91.

Nicholls, N. 2000. Opportunities to improve the use of seasonal climate forecasts. In *Applications of Seasonal Climate Forecasting in Agriculture and Natural Ecosystems - The Australian Experience.* Eds. G.L. Hammer, N. Nicholls, C. Mitchell: ISBN 0-7923-6270-5, Kluwer Academic Publishers, Dordrecht, Netherlands.

Nielsen, D. C., Mab, L., Ahujab, L. R., Hoogenboom, G. 2002. Simulating soybean water stress effects with RZWQM and CROPGRO models. *Agronomy Journal* 94: 1234-1243.

Ogoshi, R. M. 1995. *Determination of genetic coefficients from yield experiments for CERES-Maize and SOYGRO crop models*. PhD thesis. Agronomy and Soil Sciences, University of Hawaii, Honolulu, USA.

O'Neill, R. V., Deangelis, D. L. 1986. *A Hierarchical Concept of Ecosystems*. Ed.G.E. Allen. ISBN 0-691-08437-8, Princeton University Press, Princeton, USA. 262p.

Parysow, P., Gertner, G. Z., Westervelt, J. 2000. Efficient approximation for building error budgets for process models. *Ecological Modelling* 135: 111-125.

Perks, L. A. 2001. *Refinement of modelling tools to assess potential agrohydrological impacts of climate change in Southern Africa*. Ph.D. School of Bioresources Engineering and Environmental Hydrology, University of Natal, Pietermaritzburg, South Africa.

Peterson, E. H., Fraser, R. W. 2001. An assessment of the value of seasonal forecasting technology for Western Australian farmers. *Agricultural Systems* 70: 259-274.

Piwowar, J. M., Ledrew, E. F., Dudycha, D. J. 1990. Integration of spatial data in vector and raster formats in a GIS environment. *Journal of GIS* 4 (4): 429-444.

Podesta, G. P., Messina, C. D., Grondona, M. O., Magrin, G. O. 1999. Associations between grain crop yields in Central-eastern Argentina and El Niño-Southern Oscillation. *Journal of Applied Meteorology* 38 (10): 488-498.

Potgieter, A. B., Everingham, Y. L., Hammer, G. L. 2003. On measuring quality of a probabilistic commodity forecast for a system that incorporates seasonal climate forecasts. *International Journal of Climatology* 23: 1195-1210.

Potgieter, A. B., Hammer, G. L., Butler, D. G. 2002. Spatial and temporal patterns in Australian wheat yield and their relationship with ENSO. *Australian Journal of Agricultural Research* 53: 77-89.

Promburom, P., Jintrawet, A., Ekasingh, M. 2001. Estimating sugarcane yields with Oy-Thai interface. In *Proceedings of the International Society of Sugar Cane Technologists*, 81-86. Brisbane, Australia.

Pulwarty, R. S., Redmond, K. T. 1997. Climate and salmon restoration in the Columbia River basin: The role and usability of seasonal forecasts. *Bulletin of the American Meteorological Scociety* 78: 381-397.

Rabbinge, R. 1993. The ecological background of food production - Crop protection and sustainable agriculture. 2-29, In *Ciba Foundation Symposium* 177. John Wiley & Sons, Chichester, UK.

Reybold, W. U., TeSelle, G. W. 1989. Soil geographic data bases. *Journal of Soil and Water Conservation* 44: 28-29.

Rice, J. A., Cochran, P. A. 1984. Independent evaluation of a bioenergetics model for Largemouth bass. *Ecology* 65: 732-739.

Richardson, C. W. 1981. Stochastic simulation of daily precipitation, temperature, and solar radiation. *Water Resources Research* 17: 182-190.

Richardson, C. W. 1985. Weather simulation for crop management models. *Transactions of the American Society of Agricultural Engineers* 28: 1602-1606.

Roebeling, R., Rosema, A., Oroda, A. 1999. *REFEWS: Regional Famine Early Warning System for the IGAD Regional*. BCRS Report NRSP-2/99 - 17, 9-22. Delft, Netherlands.

Rosenberg, N. J., McKenney, M. S., Easterling, W. E., Lemon, K. M. 1992. Validation of EPIC model simulation of crop responses to current climate and CO_2 conditions: Comparisons with census, expert judgement and experimental plot data. *Agricultural and Forest Meteorology* 59: 35-51.

Rosenthal, W. D., Hammer, G. L., Butler, D. G. 1998. Predicting regional grain sorghum production in Australia using spatial data and crop simulation modelling. *Agricultural and Forest Meteorology* 91: 263-274.

Russell, G., van Gardingen, P. R. 1997. Problems with using models to predict regional crop production. In *Scaling-up from Cell to Landscape.* Eds. P.R. van Gardingen, G. M. Foody, P. J. Curran: 273-294pp, University Press, Cambridge, UK.

Schmidt, E. J. 1998. Forecasting the sugarcane crop. *The Link*, Vol 7(4), p.11, Publication of the South African Sugar Association, Mt. Edgecombe, South Africa, September 1998.

Schmidt, E. J., Gers, C., Narcisco, G., Frost, P. 2000. Application of remote sensing technology in the SA Sugar Industry – A review of recent research findings. In *Proceedings of the 74th South African Sugar Technologists' Association Congress*, 192-201. Durban, South Africa.

Schmidt, E. J., Gers, C., Narcisco, G., Frost, P. 2001. Remote sensing in the South African sugar industry. In *Proceedings of the International Society of Sugar Cane Technologists*, 241-245. Brisbane, Australia.

Schulze, R. E. 1983. *Agrohydrology and -Climatology of Natal*. ACRU Report No. 14. 138p, Water Research Commission, Pretoria, South Africa.

Schulze, R. E. 1995. *Hydrology and Agrohydrology: A Text to Accompany the ACRU 3.00 Agrohydrological Modelling System*. WRC Report No. TT 69/95. 552p, Water Research Commision, Pretoria, South Africa.

Schulze, R. E. 1997. *South African Atlas of Agrohydrology and -Climatology*. WRC Report No. TT82/96. 276p, Water Research Commission, Pretoria, South Africa.

Schulze, R. E., Lumsden, T. G., Horan, M. J. C., Maharaj, M. 1999. *Regional simulation analysis of hydrological and yield responses of sugarcane under dryland and irrigated conditions. ACRU*cons Report No. 28. 94p, School of Bioresources Engineering and Environmental Hydrology, University of Natal, Pietermaritzburg, South Africa.

Schulze, R. E. and Maharaj, M. 2004. *Development of a Database of Gridded Daily Temperatures for Southern Africa*. ACRUcons Report No. 41. 81p, School of Bioresources Engineering and Environmental Hydrology, University of Natal, Pietermaritzburg, South Africa.

Schulze, R. E. and Smithers, J. C. 2004. The *ACRU* Modelling System as of 2002: Background, Concepts, Structure, Output, Typical Applications and Operations. In *Modelling as a Tool in Integrated Water Resources Management: Conceptual Issues and Case Study Applications.* Ed. R.E. Schulze: Chapter 2, 47-83pp, Water Research Commission, Pretoria, South Africa.

Shaffer, M. J. 1988. Estimating confidence bands for soil-crop simulation models. *Soil Science Society of America Journal* 52: 1782-1789.

Sharpley, A. N., Williams, J. R. 1990. *EPIC - Erosion/ Productivity Impact Calculator. 1. Model Documentation*. Technical Bulletin No. 1768. 127p, US Department of Agriculture, Washington DC, USA.

Shuttleworth, W. J. 1992. Evaporation. In *Handbook of Hydrology.* Ed. D.R. Maidment: McGraw-Hill, Inc., New York, USA.

Sinclair, T. R., Seligman, N. A. G. 1996. Crop modelling: from infancy to maturity. *Agronomy Journal* 88 (5): 698-703.

Singels, A., Bezuidenhout, C. N. 1999. The relationship between ENSO and rainfall and yield in the South African sugar industry. *South African Journal of Plant and Soil* 16 (2): 96-101.

Singels, A., Bezuidenhout, C. N. 2002. A new method of simulating dry mater partitioning in the CANEGRO sugarcane model. *Field Crops Research* 78: 151-164.

Singels, A., Bezuidenhout, C. N., Schmidt, E. J. 1999a. Evaluating strategies for scheduling supplementary irrigation of sugarcane in South Africa. In *Proceedings of the Australian Society for Sugar Cane Technologists*, 219-226. Brisbane, Australia.

Singels, A., Donaldson, R. A. 2000. A simple model of unstressed sugarcane canopy development. In *Proceedings of the 74th South African Sugar Technologists' Association Congress*, 151-154. Durban, South Africa.

Singels, A., Kennedy, A. J., Bezuidenhout, C. N. 1998. IRRICANE: A simple computerised irrigation scheduling method for sugarcane. In *Proceedings of the 72th South African Sugar Technologists' Association Congress*, 117-122. Durban, South Africa.

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

Smith, D. I., Hutchinson, M. F., McArthur, R. J. 1993. Australian climatic and agricultural drought: payments and policy. *Drought Network News* 5 (3): 11-12.

Smith, J. M. B. 1992. *Sugar yield estimation*. Report No. N/A/94/4. 30p, KwaZulu-Natal Department of Agriculture, Cedara, South Africa.

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

Smithers, J. C., Schulze, R. E. 2000. *Long duration design rainfall estimates for South Africa*. Report No. 811/1/00. Water Research Commission, Pretoria, South Africa.

Söderström, M., Magnusson, B. 1995. Assessment of local agroclimatological conditions - a methodology. *Agricultural and Forest Meteorology* 72: 243-260.

Spitters, C. J. T., Toussaint, H. A. J. M., Goudriaan, J. 1986. Seperating the diffuse and direct component of global radiation and its implications for modelling canopy photosynthesis. Part 1 - Components of incoming radiation. *Agricultural and Forest Meteorology* 38: 217-229.

Steele, D. D., Scherer, T. F., Prunty, L. D., Stegman, E. C. 1996. Correction frequencies for four irrigation scheduling methods for corn. In *Proceedings of* *the International Evapotranspiration and Irrigation Scheduling*, San Antonio, USA, 3-6 November 1996.

Stephens, D. J., Butler, D. G., Hammer, G. L. 2000. Using seasonal climate forecasts in forecasting the Australian wheat crop. In *Applications of Seasonal Climate Forecasting in Agriculture and Natural Ecosystems - The Australian Experience.* Eds. G.L. Hammer, N. Nicholls, C. Mitchell: ISBN 0-7923-6270-5, Kluwer Academic Publishers, Dordrecht, Netherlands.

Stern, P., Easterling, W. E. 1999. *Making climate forecasts matter*. ISBN 0-309-06475-9, National Academic Press, Washington D.C., USA. 175p.

Stern, W., Miyakoda, K. 1995. Feasibility of seasonal forecasts inferred from multiple GCM simulations. *Journal of Climate* 8 (5): 1071-1085.

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

Stone, R. C., Smith, I., McIntosh, P. 2000. Statistical methods for deriving seasonal climate forecasts from CGMs. In *Applications of Seasonal Climate Forecasting in Agriculture and Natural Ecosystems - The Australian Experience.* Eds. G.L. Hammer, N. Nicholls, C. Mitchell: ISBN 0-7923-6270-5, Kluwer Academic Publishers, Dordrecht, Netherlands.

Struzik, P. 2001. *Spatialisation of Solar Radiation - draft report on possibilities and limitations*. COST Report No. 718 - WG1.1. 12p, European Cooperation in the Field of Scientific and Technical Research, Brussels, Belgium.

Supit, I. 1997. Predicting national wheat yields using crop simulation and trend models. *Agricultural and Forest Meteorology* 88: 199-214.

Taylor, V. 2001. *Hydrological modelling applications for water resources management in the Mkomazi Catchment*. M Sc dissertation. School of Bioresources Engineering and Environmental Hydrology, University of Natal, Pietermaritzburg, South Africa.

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

Thompson, G. D. 1978. The production of biomass by sugarcane. In *Proceedings of the 76th South African Sugar Technologists' Association Congress*, Mt. Edgecombe, South Africa.

Thompson, P. D. 1977. How to improve accuracy by combining independent forecasts. *Monthly Weather Review* 105: 228-229.

Thorburn, P. J., van Antwerpen, R., Meyer, J. H., Bezuidenhout, C. N. 2002. The impact of trash management on soil carbon and nitrogen: (I) Modelling long-term experimental results in the South African sugar industry. In *Proceedings of the 76th South African Sugar Technologists' Association Congress*, 260-268. Durban, South Africa.

Thornton, P. K., Bowen, W. T., Ravelo, A. C., Wilkens, P. W., Farmer, G., Brock, J., Brink, J. E. 1997. Estimating millet production for famine early warning: An application of crop simulation modelling using satellite and ground-based data in Burkina Faso. *Agricultural and Forest Meteorology* 83: 95-112.

Thornton, P. K., Wilkes, P. W. 1998. Risk assessment in food security. In *Systems Approaches for Sustainable Agricultural Development: Understanding Options for Agricultural Production.* Eds. G.Y. Tsuji, G. Hoogenboom, P. K. Thornton: Kluwer Academic Press, Dordrecht, Netherlands.

Todini, E. 1999. Using phase-state modelling for inferring forecasting uncertainty in nonlinear stochastic decision schemes. *Journal of Hydroinformatics* 1 (2): 75-82.

van Antwerpen, R. 1998. *Modelling root growth and water uptake of sugarcane cultivar NCo376*. PhD thesis. University of the Orange Free State, Bloemfontein, South Africa.

van Heerden, J., Steyn, P. C. 1999. *Weather radar measurement of rainfall for hydrological and other purposes*. WRC Report No. 693/1/99. Water Research Commision, Pretoria, South Africa.

van Lanen, H. A. J., van Diepen, C. A., Reinds, G. J., de Koning, G. H. J., Bulens, J. D., Bregt, A. K. 1992. Physical land evaluation methods and GIS to explore the crop growth potential and its effects within the European communities. *Agricultural Systems* 39: 307-328.

Walker, N. D. 1990. Links between South African summer rainfall and temperature variability of the Aghulhas and Benguela current systems. *Journal of Geophysical Research* 95: 3297-3319.

Welding, M. C., Havenga, C. M. 1974. The statistical classification of rainfall stations in the Republic of South Africa. *Agrochemophysica* 6: 5-24.

Wilby, R. L., Wigley, T. M. L. 2000. Precipitation predictors for downscaling: observed and general circulation model relationships. *International Journal of Climatology* 20: 641-661.

Wilby, R. L., Wigley, T. M. L., Conway, D., Jones, P. D., Hewitson, B. C., Main, J., Wilks, D. S. 1998. Statistical downscaling of general circulation model output: A comparison of methods. *Water Resources Research* 42: 2995-3008.

Winkler, R. L., Makridakis, S. 1983. The combination of forecasts. *Journal of the Royal Statistical Society* 146: 150-157.

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

Wood, A. W. 1995. Application of crop growth models to the sugar milling and raw sugar marketing. In *Workshop proceedings on the Research and Modelling Approaches to Examine Production Opportunities and Constraints.* Ed. M.J. Robertson, University of Queensland, Brisbane, Australia.

Wörlen, C., Schulz, K., Huwe, B., Eiden, R. 1999. Spatial extrapolation of agrometeorological variables. *Agricultural and Forest Meteorology* 94: 233-242.

Wynne, A. T. 2001. Delivery efficiencies and cane quality in the South African sugar industry: Benchmarking and penalty allocations. In *Proceedings of the 75th South African Sugar Technologists' Association Congress*, 38-42. Durban, South Africa.

Xie, P., Arkin, P. A. 1997. Global precipitation: A 17-year monthly analysis based on gauge observations, satellite estimates and numerical model outputs. *Bulletin of the American Meteorological Scociety* 78: 2539-2558.

Young, F. J., Nammer, R. D., Williams, F. 1998. Evaluating central tendency and variance of soil properties within map units. *Soil Science Society of America Journal* 62: 1640-1646.

Zelenka, A., Perez, R., Seals, R. 1998. Effective accuracy of models converting satellite radiances to hourly surface insolation. In *Proceedings of the Satellite Data Users Conference, EUMETSAT*, 710-713. Paris, France, 25-29 May 1998.

Zhou, M. M., Singels, A., Savage, M. J. 2003. Physiological parameters for modelling differences in canopy development between sugarcane cultivars. In *Proceedings of the 77th South African Sugar Technologists' Association Congress*, 610-621. Durban, South Africa.

Zucchini, W., Adamson, P. T. 1984. *The occurrence and severity of droughts in South Africa*. WRC Report No. 91/1/84. 198p, Water Research Commission, Pretoria, South Africa.