Application of seasonal climate outlooks to forecast sugarcane production in South Africa

C. N. Bezuidenhout*, R. E. Schulze

School of Bioresources Engineering & Environmental Hydrology, University of KwaZulu-Natal, Private Bag X01, Scottsville 3209, South Africa

ABSTRACT: Sugarcane production is an economically important activity in South Africa. Climate variability causes high fluctuations in annual production, and stakeholders in the industry are often compelled to make crop management, harvesting, transport, milling and marketing decisions under high levels of uncertainty. In this study, a previously derived model-based yield forecasting system is used to produce hindcasts of sugarcane production in South Africa. The derivation of climate outlooks by the South African Weather Service (SAWS) and a technique to translate these into yield forecasts are explained. Model generated forecasts are consistently more accurate than forecasts based on climatology and, depending on the time of the year, capture between 11 and 58% of the natural variation in national sugarcane production. Climate outlook information generally enhances production forecasts, suggesting that these outlooks are sufficiently accurate for decision making and that the translation techniques perform satisfactorily. Different regions, however, display different accuracy responses to the climate outlook and several suggestions for future research are made, such as incorporating different climate forecasts and creating consortiums between climate forecasters, crop modellers and decision makers.

KEY WORDS: Seasonal climate outlook · Production forecasting · Sugarcane · Climate forecast application

1. INTRODUCTION

South African sugarcane production areas are situated on the eastern shores of the country. This area has a large rural population and the sugar industry has become an important economic entity with more than 50 000 growers, and supports approximately 1 million people. Climate, especially rainfall, is probably the single most important factor that influences sugarcane production in South Africa. Inter-annual rainfall variability in the sugarcane growing belt is high, coefficients of variation (CV) ranging from 20 to 35% (Schulze 1997). The region is typically subject to relatively frequent severe and wide-spread droughts (e.g. 1983, 1992, 2003), occasional tropical cyclones (e.g. 1984, 2000) and less frequent mid-latitude cut-off low pressure systems which produce excessive rainfalls (e.g. 1987). As a result, the variation in annual national sugarcane production has been high (CV: 17 %) and industry stakeholders, such as growers, millers and marketers are often compelled to make crop management, harvesting, transport, milling and marketing decisions under high levels of uncertainty (Bezuidenhout 2005).

Model-based forecasting of sugar production has been researched by various authors:

• Promburom et al. (2001) derived a forecasting system for Thailand
• Everingham et al. (2002a,b) investigated the enhancement that forecasting provides to the Australian sugar value chain
• Potgieter et al. (2003) used a model to forecast sugarcane yields at field scale in Australia.

In southern Africa, Smith (1992), Jury (1998), Lumsden et al. (1999) and McGlinchey (1999) derived methodologies for forecasting sugarcane. Bezuidenhout & Singels (2001) also developed a crop model-based sugarcane production forecasting system for South Africa. This system has become operational and

*Email: bezuidenhoutc@ukzn.ac.za  © Inter-Research 2006 · www.int-res.com
was evaluated using historical climate and production data over a period of 20 yr. Forecasts were able to capture 57 and 38% of the inter-annual variation in national and mill scale production, respectively (Bezuidenhout & Singels 2003, Singels & Bezuidenhout 2005). Although a large amount of variability is seemingly still unaccounted for, Bezuidenhout (2005) reports that forecasts with this level of accuracy may, as a result of improved decision making, benefit the industry by approximately US $4.2 million yr\(^{-1}\). Internationally still unaccounted for, Bezuidenhout (2005) suggests that forecasts with this level of accuracy may, as a result of improved decision making, benefit the industry by approximately US $4.2 million yr\(^{-1}\). Importantly, still unaccounted for, Bezuidenhout (2005) also states that forecasts with this level of accuracy may, as a result of improved decision making, benefit the industry by approximately US $4.2 million yr\(^{-1}\).

Although a large amount of variability is seemingly still unaccounted for, Bezuidenhout & Singels (2003, Singels & Bezuidenhout 2005), as well as Hansen (2002), elaborate on the usefulness of seasonal climate forecasts, especially in an agricultural context. The aim of this study was to quantify to what extent the SAWS 3-TRO enhanced the accuracy of model-based sugarcane production forecasts in South Africa. This was done by (1) reporting on the methodology used to incorporate information from the 3-TRO into the production forecast, (2) demonstrating the accuracy of the production forecasting system under assumptions of neutral 3-TRO information and (3) quantifying the enhancement in production forecast accuracy by using actual 3-TRO information over a 5 yr period.

2. METHODS

2.1. Seasonal climate forecasting

The SAWS is the country’s official meteorological service. Each month its scientists produce a 3 mo rainfall and temperature outlook using a multi-tiered forecast system consisting of a dynamic modelling process, combined with a statistical approach and a consensus discussion (Landman et al. 2001, Johnston et al. 2004). The input to the consensus discussion is derived from both in-house and outside sources and include (Johnston et al. 2004):

- A prediction of near-global sea-surface temperatures using Canonical Correspondence Analysis (CCA, Barnett & Preisendorfer 1987)
- 3 mo running forecasts of rainfall and surface temperature anomalies calculated for ensemble means derived from the Centre for Ocean Land Atmosphere (COLA) T30 General Circulation Model (GCM)
- Statistical forecasts of precipitation and temperature using CCA (Landman & Mason 1999)
- Statistical GCM downscaled information for specific rainfall regions (Von Storch & Navarra 1995, Landman et al. 2001)
- Model outputs supplied by the International Research Institute (IRI) and European Centre for Medium Term Weather Forecasts (ECMWF).

2.2. Model configuration

The Canesim sugarcane yield model is a daily time step, point-based simulation tool driven predominantly by soil water availability. The model’s water balance, yield calculation and canopy development are de-
scribed by Singels et al. (1998, 1999) and Singels & Donaldson (2000), respectively. These publications also report on the accuracy of various aspects of the model. The model has been verified at mill supply scale by Gers et al. (2001), who reported excellent agreement between simulated and observed regional yields ($R^2 = 0.87$). These results suggest that the process representations in the model are sound and that it is sufficiently accurate at a regional mill scale.

The technique for forecasting production with the model has been reported by Bezuidenhout & Singels (2001, 2003), Bezuidenhout (2005) and Singels & Bezuidenhout (2005). The model simulates representative sugarcane crops in a specific region using climate data recorded at sites which are considered representative for that region (Bezuidenhout 2005). In the case of operational forecasts, climate data will terminate before the completion of a simulation and 9 to 10 analogue seasons from the historical climate record will be selected as substitutes to forecast probable future crop responses along different trajectories (cf. Fig. 1). The selection of these analogue seasons is based on the 3-TRO. For example, 2 analogues with relatively wet conditions during the selected lead time, 5 with near-normal and 3 with dry conditions will be selected if the 3-TRO has a 20, 50, 30% probability distribution for above-normal, near-normal and below-normal rainfall, respectively. Analogue seasons were selected from historic records between 1980 and 2002 and a season was classified as dry, near-normal or wet based on the total rainfall that occurred over the 3 month lead-time period (highlighted in Fig. 1).

The technique described above is limited by the facts that:

- It assumes that historical climate will re-occur, and thus may lack the ability to be representative under climate change scenarios
- As a result of limited data, it only uses 9 to 10 analogue seasons, hence allowing any extreme events during those analogue seasons to carry relatively high weights
- The selection of analogues is based on the 3-TRO (shaded area in Fig. 1), and subsequent climate after this period within the analogue seasons is not evaluated
- The probabilistic nature of the forecast is lost when means are calculated over different analogue season outcomes.

Some of the above-mentioned limitations may in future be addressed by using data series that were generated by downscaled GCMs. Many of these GCM and downscaling models (e.g. Hewitson & Crane 1996), however, are still considered crude with respect to generating high resolution spatial data.

### 2.3. Simulations

Hindcasts of sugarcane production between 1980 and 2002 were simulated using the above-mentioned model and methodology. The term *hindcast* infers that a production forecast was emulated at some time in the past assuming no climate data had extended beyond that point in time. Thereafter hindcasts were compared to the production eventually achieved. Different hindcasts were simulated for which it was assumed that climate data terminated either on 1 September prior to the opening of the milling season, or on 1 January, 1 March, 1 May, 1 September or 31 December of the actual milling season. In each case, with the exception of 31 December, 9 analogue seasons were used to complete the simulations. The 9 analogue seasons were based on a neutral 3-TRO (33, 33, 33%), resulting in the selection of 3 below-normal (i.e. dry year) scenarios, 3 scenarios with near-normal rainfall and 3 above-normal (i.e. wet year) scenarios. The selection of 3 seasons per category, as opposed to using all historic seasons, was more consistent with the operational system and also excludes extreme events. Results were evaluated at both individual mill and entire industry scales.

In addition to the above-mentioned simulations, hindcasts of the 1998 to 2002 seasons were also simulated.
using the actual SAWS 3-TRO issued at that specific time. These simulations were therefore based on different selected analogue seasons. The difference in production accuracies between hindcasts based on an assumed neutral 3-TRO and those based on the actual 3-TRO was used to quantify the value of the SAWS 3-TRO.

2.4. Evaluation parameters

Because model-based yield simulations often overestimate actual yields (Haskett et al. 1995, Russell & van Gardingen 1997, Hansen & Jones 2000), hindcast accuracy was expressed as the bias-corrected relative root mean square error \( S_i \) [%] Eq. 1. This parameter not only removes model bias, but also allows comparison of \( S_i \) values originating from different mills (cf. Supit 1997) through expressing the error relatively to mean annual production over 23 yr \((Y, \text{t ha}^{-1})\). Annual regional production data \((Y, \text{t ha}^{-1})\) were detrended by Bezuidenhout & Singels (2006) for changes in management practices (especially as a result of pests and diseases), changes in milling configurations and expansion. \( S_i \) is hence expressed as:

\[
S_i = \left[ \frac{1}{n} \sum_{i=1}^{n} \frac{(\text{Est}_i \times \frac{Y}{\text{Est}_i} - Y_i)^2}{Y} \right] \times 100 \tag{1}
\]

where \( n \) is the total number of observed and simulated data pairs (23 yr), \( \text{Est}_i \) (t ha\(^{-1}\)) is the simulated production in Year \( i \), and \( Y \) and \( \text{Est} \) are the mean observed and simulated production over all years, respectively.

Table 1 reports the mill average and national \( S_i \), hindcast skills \( \text{Skill} \) and Directional skill

Hindcast skill \( \text{Skill} \) [%]; Eq. 2) was determined as the ratio between \( S_i \) and the CV of observed yields \( (CV_Y) \) according to Murphy (1993) and Mason (2000). The CV (SD divided by the mean) is, by definition, the relative root mean square error of forecasts assuming the long-term mean. Eq. (2), therefore, reflects the improvement in forecast accuracy above a forecast that will simply assume the long-term mean. \( \text{Skill} \) can alternatively be defined as the amount of the long-term SD in inter-annual production that has been captured or explained by the hindcast.

\[
\text{Skill} = \left(1 - \frac{S_i}{CV_Y}\right) \times 100 \tag{2}
\]

In addition to the above-mentioned parameters, the coefficient of determination \( (R^2) \) and the frequency with which yields were hindcast with the correct sign (i.e. higher or lower, when compared with yields of the previous season), termed the Directional skill, were also calculated.

3. RESULTS AND DISCUSSION

Table 1 reports the mill average and national \( S_i \), \( \text{Skill} \) and Directional skill values based on the 1980 to 2002 simulations while assuming a neutral 3-TRO at all times. Fig. 2 illustrates 6 time series of mean national annual yield achieved and hindcasted between 1980 and 2002. These illustrate hindcast accuracies at particular times, viz. in September\(_{n-1}\), January, March, May, September and December. From these results it may be seen that:

- All \( \text{Skill} \) values were positive; therefore model-based hindcasts were always superior to simply assuming the long-term mean
- Hindcast accuracies were, as expected, higher at national scale than at mill scale
- As time progressed through the milling season (allowing for the inclusion of more actual climate data) hindcast accuracies increased
- Significant increases in hindcast accuracies occurred over the summer months (November to March)
- Directional skill values were high, even while hindcast skills in September prior to the milling season were relatively low.

These evaluations compare well with previous forecasting assessments. On an industry scale, \( R^2 \) values of actual vs. hindcasted yields ranged from 0.48 to 0.82, depending on the time of the hindcast (cf. Fig. 2). Jury (1998) achieved an \( R^2 \) value of 0.69 in December of the preceding year by using a statistical ocean-atmospheric response model. This \( R^2 \) is considerably higher compared to the model-based forecast at the same time, but neutral outlooks were used in the
model and more research is needed to establish whether Jury's (1998) approach is indeed superior. Promburom et al. (2001) reported an accuracy of 4.8% when forecasting sugarcane production in Thailand. Their result is not comparable with the evaluation parameters used in this study and generally the natural seasonal variability in cane production in Thailand was lower than in South Africa.

Early season forecasts (September, January) did not predict yield losses in 1983, 1992 and 1993 (Fig. 2). These were severe drought seasons and it is expected that neither the crop model, as it was configured for the forecast, nor the method used to select analogue seasons are capable of predicting such extreme events.

Fig. 3 illustrates the improvements denoted in hindcasts as a result of including information originating from the SAWS 3-TRO. On a national scale the 3-TRO in September of the previous year generally did not enhance hindcasts. However, hindcasts of production based on the SAWS 3-TRO issued from January to May of the milling season were generally more accurate compared to neutral 3-TRO assumptions. Improvements in Skill were highest in the midsummer month of January (11.57%) and deteriorated towards the winter. A forecast skill improvement of >10% at a national scale should be regarded as significant, although the true value of such an accuracy improvement resides with the benefits generated through enhanced decision making. This remains a complex management issue and falls outside the scope of this study.

There seems to be no clear spatial pattern as to where hindcast accuracies were consistently enhanced or reduced by the SAWS 3-TRO. Hindcast accuracies in the northern coastal production areas in KwaZulu-Natal were most markedly influenced, with Skills initially being reduced in the September prior to the milling season and significantly enhanced thereafter.

The selection of analogue seasons in this study was based on 3 mo accumulated rainfall outlook information. Several other seasonal climate outlooks are also available from the SAWS, as well as from other forecasting groups, such as ECMWF and IRI. These range from monthly accumulated rainfall outlooks, 0 to 3 mo mean temperature outlooks and 3 to 6 mo 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 simply totals and means. A close partnership between climate forecasters and the sugar industry (and other agriculture sectors) may be necessary to develop synergies.

Significant scope exists for decision makers, climate forecasters and crop modellers to work together more closely. Forecasters, for example, may be able to customise information for specific agricultural sectors, such as generating lists of analogue seasons, in order for this information to be readily incorporated into yield simulations. Further research is needed in climate forecast downscaling to, e.g., sugar mill supply areas, as well as in determining the effect of micro climatic conditions at different locations. There is also significant scope to incorporate climate forecasts other than those with a 3 mo lead time and with only rainfall, such as by incorporating forecasts of temperature. In addition, quantifying an operational forecast uncertainty at different scales remains a difficult task as a result of limited data on the spatial variability in climate.

4. CONCLUSIONS

Depending on the time of forecast, the Canesim model-based sugarcane yield forecast system was able to capture between 11 and 58% of the natural seasonal
Fig. 3. Spatial changes in forecast skill as a result of utilising seasonal rainfall outlook information as opposed to assumptions of neutral outlooks. Overall skill enhancement at a national scale in top right of each panel.
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.

Seasonal rainfall outlook information issued over the period 1998 to 2002 generally improved forecast accuracies. Rainfall outlooks issued in the January prior to a milling season (April to December) increased the forecast skill for the industry by 11.6%. Rainfall outlooks became less valuable after May as the drier winter period started to set in. Actual rainfall outlooks in September were generally no more valuable than neutral rainfall outlook assumptions. It should be noted, however, that these results were based on only 5 years of information, and that several advances in climate forecasting technology have been phased in over these years (W. Landman, SAWS, pers. comm). The results do, however, indicate significant potential enhancements in yield forecasting capabilities and should encourage collaborative research between the South African sugar industry and seasonal climate forecasters.

Future research should aim at bringing climate forecasters, crop modellers and decision makers into closer collaboration, and to incorporate other readily available climate forecasts into the model-based sugarcane production forecasting framework.

Acknowledgements. Research funding for this project was made available by the University of KwaZulu-Natal and the South African Sugarcane Research Institute. The South African Sugarcane Research Institute, in particular Dr. A. Singels, is thanked for advice and for making data and the Canesim sugarcane crop yield model available. A history of seasonal climate forecasts, and 3-TRO information, were made available by the SAWS, Pretoria, South Africa. Other climate data used in this study were obtained through kind made available by the University of KwaZulu-Natal and the South African Sugar Commission.

LITERATURE CITED


Mason SJ (2000) Definition of technical terms in forecast verification and examples of forecast verification scores. IRI-CW/01/1, International Research Institute for Climate Prediction, Palisades, NY


Schulze RE (1997) South African atlas of agrohydrology and climate data used in this study were obtained through kind made available by the SAWS, Pretoria, South Africa. Other seasonal climate forecasts, and 3-TRO information, were made available by the South African Sugarcane Research Institute.

Editorial responsibility: Bryson C. Bates, Canberra, Australia
Submitted: May 20, 2005; Accepted: February 16, 2006
Proofs received from author(s): March 16, 2006