HOW ACCURATE ARE CANESIM ESTIMATES OF SUGARCANE PRODUCTION?

C N BEZUIDENHOUT* and A SINGELS

South African Sugar Association Experiment Station, Private Bag X02, Mount Edgecombe, 4300, South Africa
**Currently: University of Natal, Pietermaritzburg, P/Bag X01, Scottsville, 3209, South Africa*

Abstract

Forecasting the size of the sugarcane crop is essential to the industry for optimising cane milling and sugar marketing. The SASA Experiment Station provides regular forecasts of mill and industry average cane yields from a climatic perspective. The usefulness of this information depends on its accuracy, and this communication reports briefly on this aspect.

Future yields are calculated by applying the Canesim crop model to historic weather data, likely future weather scenarios, soil characteristics and management practices. Data from approximately 70 weather and rainfall stations are used to simulate the crop for each month of the milling season. Mean yields are calculated for homogenous climate zones, mill areas and the industry as a whole, and are expressed as a percentage of the yield from the previous season. The information is disseminated to the industry bi-monthly from February to August.

Estimation accuracy was determined by calculating the relative mean absolute error between estimated and actual tonnage. Actual yields were adjusted for long term trends in harvest age, the extent of the area that receives irrigation, and any water restrictions. This was compared with the inter-annual coefficient of variation in actual yields to assess the potential value of the forecasts.

The mean absolute error (MAE) for the industry estimate was 6.6% compared with an inter-annual variation in actual yield of 15.4%. This suggests that the industry cane crop estimate should be of value to decision-makers. On average, the MAE at the mill level was 12%, and ranged from 26% for Komatipoort to 8% for Amatikulu. These results indicate that, although further improvements to estimates are needed for irrigated mill areas, the estimates for most dryland areas are sufficiently accurate to be of value to mill management.

Keywords: sugarcane, modelling, Canesim, crop estimate, crop forecast, yield

Introduction

Forecasts of the size of the sugarcane crop are essential for optimising the milling and marketing functions of the industry. The SASA Experiment Station (SASEX) provides regular forecasts on a mill and industry level, of the average cane yield from a climatic perspective. Mean yields are calculated for homogenous climate zones, mill areas and the industry as a whole, and are expressed as a percentage of the yield from the previous season. This information is disseminated to the industry bi-monthly from February to August. The usefulness of the forecasts depend on their accuracy, and this communication reports briefly on this aspect.

Methods

Cane yields have been forecast operationally by SASEX since the 2000/01 milling season (Bezuidenhout and Singels, 2001). Future yields are calculated by applying the Canesim crop model to historic weather data, likely future weather scenarios, soil water holding capacity (TAM) and management data. Information from approximately 70 weather and rainfall stations is used to simulate the crop for each month of the milling season. The same data configuration and forecasting system used in the February 2003 forecasts of the 2003/04 yields (Bezuidenhout and Singels, 2003) was used for a hindcast of yields for the period 1980/81 to 2002/03. (A hindcast is defined as a yield estimation based entirely on historic weather data.) Some of the inputs used for the hindcast are summarised in Table 1.

Estimation accuracy was determined by calculating the mean error between estimated and actual yields on a mill supply area and industry level. Actual yields were adjusted for long term trends such as reduced harvest age, increased area under irrigation and changes in mill cane supply. Adjusted actual yields represent the history of cane yields that would have been achieved under today's agronomic and milling practices. An estimation error could arise from any one or more of the following sources: (1) simulation error, (2) incomplete or unrepresentative weather, soil and management data, and (3) incorrect actual yield data. The relative mean absolute error (RMAE) was calculated according to equation 1. It should be noted that estimated yields (Est) were corrected for any bias.

$$RMAE = \frac{\frac{1}{n}\sum_{y=1980}^{2002}\left|Est_y \times \frac{\overline{Act}}{\overline{Est}} - Act_y\right|}{\overline{Act}} \times 100 \qquad (1)$$

where Est_y and Act_y are the estimated and actual yields (t/ha) in season y respectively, and $\overline{Est}$ and $\overline{Act}$ are the averages (t/ha; 1980 to 2002) for *Est* and *Act* respectively.

The RMAE was compared with the coefficient of inter-annual variation in actual yields (CV%) to assess the potential value of estimates. The value (%) was quantified using equation 2:

$$\text{Value Index} = (CV - RMAE)/CV \times 100 \qquad (2)$$

Estimates were regarded as useful when the value index exceeded 30%.

Operational forecasts (as opposed to hindcasts) of mill and industry production by SASEX and Mill Group Boards were also evaluated, after comparing the forecasts issued in March with the final tonnage. An average relative absolute error (equation 3) over the 2001/02 and 2002/03 seasons was calculated for each mill and the industry.

$$Err = \frac{\left|\frac{Est_y^{Mar}}{Est_{(y-1)}} - \frac{Act_y}{Act_{(y-1)}}\right|}{\frac{Act_y}{Act_{(y-1)}}} \times 100 \qquad (3)$$

where Est_y^{Mar} is the forecast yield (t/ha) issued in March of year y.

Results and Discussion

The comparison of actual and hindcast yields is summarised in Table 1. Time series of actual and hindcast yields for the industry and for the Felixton mill are illustrated in Figure 1.

Reasonable to good agreement existed between estimated and actual yields for most dryland mills and for the industry. Estimates of industry average yields were on average within 6.6% of the true value, whereas the average error on a mill level was 12%. The error for individual mills ranged from 26% for Komati to 8% for Amatikulu. Estimates for mills with substantial areas under irrigation were not accurate enough to provide value, nor were estimates for the Eston and Entumeni mills. Estimates for all other mills were found to be useful, as measured by the value index of more than 30%.

Table 1. Summary of inputs used for the hindcast, results from the hindcast (error (RMAE), coefficient of variation of actual yields, index of estimation value) and a comparison between the error of Mill Group Board (MGB) and SASEX operational forecasts in March.

Mill	No of rainfall gauges	No. of met sites	Average crop age (months)	Average TAM (mm)	Hindcast RMAE (%)	CV actual yields (%)	Value index (%)	MGB March error (%)	SASEX March error (%)
Amatikulu Mill	4.73	1.00	14.1	82.7	7.8	18.9	58.9	11.1	11.9
Darnall Mill	8.55	4.00	13.9	86.2	8.5	18.0	52.9	15.0	3.4
Entumeni Mill	1.00	1.00	16.0	92.0	12.7	14.7	13.1	11.0	4.0
Eston Mill	3.41	6.00	18.1	111.9	11.9	15.3	22.3	7.5	5.5
Felixton Mill	6.00	3.00	12.7	83.6	11.0	19.6	43.8	17.2	11.0
Gledhow Mill	20.14	4.00	13.9	88.4	10.7	16.5	34.9	9.9	1.4
Komati Mill	0.00	2.00	12.0	50.0	25.8	19.7	31.2	8.5	6.8
Maidstone Mill	12.59	2.00	13.9	92.2	11.4	17.0	32.7	11.4	3.1
Malelane Mill	0.00	3.00	12.0	50.0	15.7	17.6	11.2	3.6	5.9
Noodsberg Mill	3.50	4.00	22.0	98.1	10.8	20.0	46.1	9.7	20.1
Pongola Mill	0.00	2.00	12.0	50.0	15.3	10.2	50.1	2.6	0.8
Sezela Mill	9.73	2.00	15.5	68.1	11.4	22.7	49.8	9.0	4.9
Umfolozi Mill	1.27	4.00	12.0	81.1	13.4	15.2	11.7	11.8	4.4
Umzimkulu Mill	5.24	1.00	19.8	78.4	10.2	18.4	44.6	10.7	6.7
Union Coop Mill	4.50	6.00	21.8	98.9	9.5	16.3	42.1	9.9	20.3
Mill totals, means	**80.65**	**45.00**	**14.8**	**80.3**	**11.96**	**17.88**	**33.1**	**9.9**	**7.3**
Industry means					**6.6**	**15.4**	**57.5**	**7.0**	**2.1**

The poor performance of the model in irrigated areas could be due to a weak simulation of historical water restrictions, as well as incorrect TAM values. It will be difficult to reduce the error below the coefficient of variation of well-irrigated crops.

Complete sets of observed weather data were used for these hindcasts. Operational forecasts early in the season will have an additional uncertainty associated with rainfall outlooks (see e.g. Everingham *et al*., 2002). The review of operational forecasts made in March each year for the 2001/02 and 2002/03 seasons shows better than expected accuracy at mill and industry level (Table 1). The average error of 2.1% for the March industry forecast is remarkable, while the average MGB error of 7.3% is comparable with the 5% of Lumsden *et al*. (2000) for March forecasts of Eston mill average yield. SASEX forecasts for the Midlands North mills were disappointingly poor and need further investigation. SASEX forecasts in March 2001 and March 2002 were better than the conventional MGB forecasts for all mills except Amatikulu, Malelane, Union Co-op and Noodsberg.

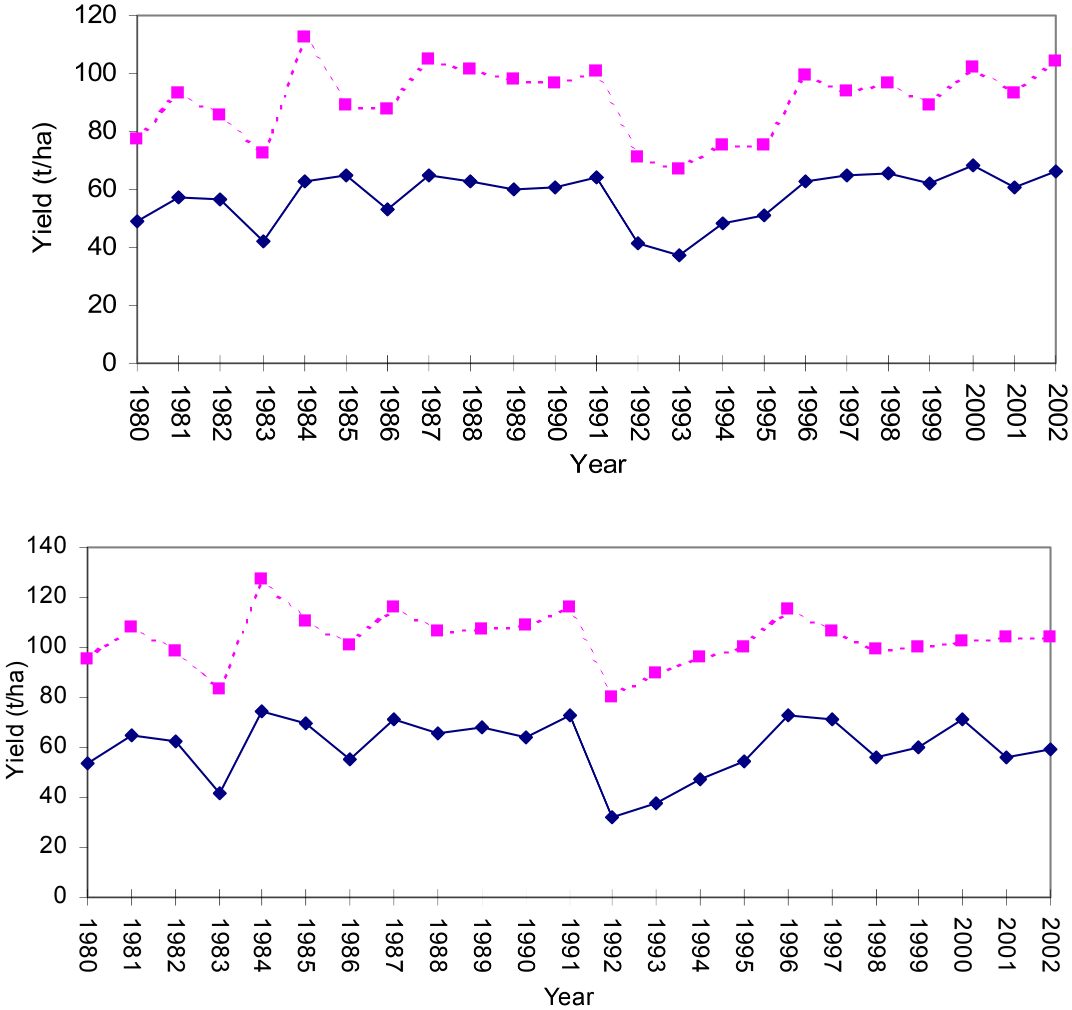

Figure 1. Time series of calculated and actual yields for (a) the industry as a whole and (b) the Felixton mill supply area.

Acknowledgements

The Agricultural Engineering and Biometry Departments at SASEX are acknowledged for capturing and processing the climate data used in this study. SASEX Extension Officers (in conjunction with other industry stakeholders) provided valuable agronomic and milling information for their relevant mill areas. The SASA Industrial Affairs division is thanked for collating and supplying MGB estimates, while historic production data were supplied by the SA Canegrowers Association. The authors would also like to thank the Department of Water Affairs, the Impala Water Board and Transvaal Sugar Ltd (TSB) for supplying river flow and water restriction information for the Pongola and Onderberg regions.

REFERENCES

Bezuidenhout CN and Singels A (2001). The use of simulation crop modelling to forecast sugarcane yield. pp 20-29 In: Proceedings of SASTA Workshop on Integration of Appropriate Technology to Provide Practical Solutions to Burn/harvest to Crush Delays and Crop Estimating in the South African Sugar Industry (CD-ROM). 8 November 2001, South African Sugar Association, Flanders Drive, Mount Edgecombe, South Africa.

Bezuidenhout CN and Singels A (2003). Climatic cane yield estimates for the 2003/04 season. Internal Report, February 2003. South African Sugar Association Experiment Station, P/Bag X02, Mount Edgecombe, South Africa.

Everingham YL, Muchow RC, Stone RC, Inman-Bamber NG, Singels A and Bezuidenhout CN (2002). Enhanced risk management and decision-making capability across the sugarcane industry value chain based on seasonal climate forecasts. *Agric Systems* 74(3): 459-477.

Lumsden TG, Schulze RE, Lecler NL and Schmidt EJ (2000). Assessing the potential for improved sugarcane yield forecasting using seasonal rainfall forecasts and crop yield models. *Proc S Afr Sug Technol Ass* 74: 131-139.