EMPIRICAL MODELLING AND PREDICTION OF SUGARCANE YIELDS FROM FIELD RECORDS

EA BRÜGGEMANN[ac], JR KLUG[a], PL GREENFIELD[a] AND HM DICKS[b]

[a] *School of Agricultural Sciences and Agribusiness*
[b] *School of Mathematics, Statistics and Information Technology, University of Natal, Private Bag X01, Scottsville, 3209*
[c]*Present address: Rietspruit Farms, PO Box 11560, Dorpspruit, 3206, South Africa*

Abstract

Commercial sugarcane records for 19 seasons from 146 fields were obtained from five estates in the KwaZulu-Natal Midlands. Extensive editing and cleaning of the agronomic records was required. Regression models were developed, and depending on the predictor variables selected, the best model accounted for 55% of the observed yield variation, based on 535 crop cycles. The final models were validated using an independent data set of 47 observations and satisfactory performances by the models were confirmed. The 95% confidence limits of yield predictions for the population mean were within 10% of long-term mean yields and could be useful for resource allocation and harvest planning at the estate level. Reliable yield predictions for individual fields could not be made and within-field resource variations could not be adequately accounted for. Key physical field attributes associated with sugarcane yield were locality, aspect, altitude, soil type and effective rooting depth. Season and rainfall were important climatic variables. Of the factors influenced by management, sugarcane variety, plant-ratoon status, crop cycle, N and K nutrition and the topsoil Ca:Mg ratio were important yield predictors. The relative importance of individual predictors varied with the specific combination of resources for a particular observation.

Keywords: sugarcane, modelling, yield prediction, field records

Introduction

Many growers keep records to facilitate improved farm planning and control. The lowest level of spatial aggregation to which these records generally refer is the commercial sugarcane field and the data are usually available only in an unprocessed form. They are frequently underutilised because of the extreme difficulty experienced in observing and interpreting production trends. In South Africa limited use has been made of commercial field records to guide crop management decisions (Hoekstra, 1976; Landrey *et al*., 1981, Culverwell, 1984), assess farm productivity (Tucker, 1975; Hulbert and Harding, 1980; Hellmann, 1988) and investigate the effect of agronomic practises on sugarcane yields (Hellmann, 1993 and Hellmann *et al*., 1995). The objective of this research was to determine key parameters strongly associated with sugarcane yield in field records, and to derive empirical models for yield prediction. Mondi Forests provided field records and detailed natural resource data for five estates in the Kranskop, Umvoti and Richmond districts. This sugarcane is all grown within the Moist Midlands Mistbelt bioresource group, BRG 5 (Camp, 1999), and the climate is broadly representative of the higher-potential rainfed sugarcane production region of the Midlands.

Method

Records for 984 completed crop cycles (harvest to harvest) from 146 fields (974.3 ha) spanning 19 seasons (1979 – 1997) were provided. Comprehensive records of soil fertility, foliar nutrient concentrations, agronomic inputs and yields were available for 13 seasons. Extensive editing of these records was necessary to provide data suitable for electronic manipulation and statistical analysis (Brüggemann, 1999). Unrepresentative observations were excluded from the investigation using five *a prior* criteria:

i exclusion of all crop cycle types with less than 20 observations;

ii exclusion of all observations without fertilizer input records;

iii exclusion of all observations without topsoil fertility analyses;

iv exclusion of all observations where the duplicate records differed by more than 10%; and

v exclusion of all observations of fields with mixed sugarcane varieties, and varieties with fewer than 20 observations.

The restricted data set of 543 observations was considered suitable for statistical examination. Where a record was not available for an observation, it was coded as a missing value. Linear regression analysis based on the method of least squares was used to establish a functional relation among the variables. The data were analysed using GENSTAT version 5, release 3.2 software. Sugarcane yield (tons cane per hectare (TCH)) was defined as the dependent variable. The predictor variable set comprised 23 qualitative parameters, defined as categorical variables, and 66 continuous variables. Summary statistics for those variables found to be important in the final models are presented in Table 1.

Care was exercised not to consider aliased categorical variables simultaneously. The most parsimonious yet relevant models were developed using a manual forward selection procedure. The use of automated selection procedures, e.g. stepwise and all subsets regression, was precluded by the large number of categorical variables that needed to be screened. Careful evaluation of candidate models was required since, depending on which predictors were included in the terms statement, they were derived from a data set of varying size. Models were built with the aim of achieving a high coefficient of multiple determination (R^2) and low standard error of estimate (SE), matching predicted to measured yields as closely as possible, using the smallest number of easily measured, readily comprehendible, independent variables. The 10% level of significance was used to include or reject variables as predictors. Where the differences between

Table 1. Summary statistics for variables found to be important in the final models.

Dependent variable	Range (Min. - Max.)	Mean	Median	Lower quartile	Upper quartile	Standard deviation	CV %	Number of observations
Tons cane per hectare (TCH)	29 - 169	87.3	88	67	105	25.2	28.9	543
Continuous variables								
N fertilizer (kg ha^{-1})	43 - 184	114.2	112	98	131	23.1	20.2	543
Rain total for cycle (mm)	925 - 3573	2070	2039	1631	2449	533.8	25.8	543
Rain for 12 months’ growth (mm)	436 - 2001	1037	961	821	1255	333.5	32.2	543
Altitude (m a.s.l.)	916 - 1121	1025	1032	970	1071	56.1	5.5	543
Total K (kg ha^{-1}) (fertilizer application + topsoil analysis)	104 - 1316	425	402	304	479	195.7	46.1	543
Effective rooting depth ERD (cm)	52 - 151	125.1	129	114	137	19.0	15.2	543
Topsoil Ca:Mg (using kg ha^{-1} nutrient content) (Topsoil Ca:Mg using charge equivalents)	1.4 - 34.9 (0.9 - 21.2)	5.3 (3.2)	4.0 (2.4)	3.1 (1.9)	6.0 (3.6)	4.0 (2.4)	75.2 (75.2)	543 (543)

Categorical variables	Levels	Number of observations
Season (Regression coefficient differences for some seasons were not statistically significantly different from each other and seasons were grouped where appropriate.)	**1** (1993, 1996) **2** (1990, 1994, 1995, 1997) **3** (1991, 1992) **4** (1985, 1988) **5** (1986, 1987, 1989)	92 192 82 77 100
Soil type	**1** (Humic yellow and red / yellows) **2** (Humic red) **3** (Orthic red)	248 246 49
Variety	**1** (NCo376 and NCo293) **2** (N16 and N12)	338 205
Plant / ratoon crop	**1** (Plant, Ratoons 1 and 2) **2** (Ratoons 3, 4 and 5) **3** (Ratoons 6, 7, 8 and 9)	236 233 74
Aspect	North and West South and East	242 301
District	Umvoti Kranskop Richmond	102 229 212
Crop cycle	Winter harvest Summer harvest	378 165

levels of categorical variables included in the model were statistically non-significant, these were grouped to simplify the models. A regression model was developed to account for as much of the observed yield variation in TCH as possible, screening all available independent variables. This model is referred to as the “yield model”. A second model was developed, also aiming to account for as much of the observed variation in TCH as possible, but screening only those independent variables known at least six months in advance of the planned harvest date. This model is referred to as the “prediction model”. Analyses of residuals were used to check the models for bias and curvilinear trends. In all cases the assumptions of normality and homoscedasticity were upheld and therefore no transformation of TCH was required. In the examination of observations with poor model fit, eight observations with either standardized residuals >3 in absolute value or leverage >0.1 were investigated and confirmed to be at variance with previous field history and the general production trends. They were therefore excluded from the database as abnormal observations and the final models were based on 535 observations. Multicollinearity was assessed and variance inflation was found to be within acceptable limits.

The R^2 and SE statistics refer to model fit alone and do not indicate how models perform when applied to new data sets, even from the same resource base. For this reason the models were validated using independent data from 47 fields harvested during 1998. The models were run in GENSTAT using the fixed conditions for each predictor and the predictive performance evaluated by comparing model output with the actual 1998 yields.

Results and discussion

The models

The regression coefficients and related statistics for the yield model are presented in Table 2 and those for the prediction model in Table 3. A term used to describe a variable is printed in *italics* when it refers to a variable in statistical analysis.

Two important considerations should be borne in mind when interpreting the models. Firstly, the predictor variables are not necessarily the cause of sugarcane yield variations, but are related only in a strictly statistical sense. It is therefore not possible to explain why a certain yield was achieved. Secondly,

Table 2. Regression coefficients for the yield model.

Categorical variables	Levels	Regression coefficient differences	Standard deviation	t-value 513 d.f.#	Probability
Constant	Reference categories †	-0.7	30.9	-0.02	0.982
Season §	**1** (1993, 1996)	-	(Reference category)		
	2 (1990, 1994, 1995, 1997)	13.77	2.16	6.36	<0.001
	3 (1991, 1992)	21.42	2.60	8.23	<0.001
	4 (1985, 1988)	29.41	3.49	8.42	<0.001
	5 (1986, 1987, 1989)	33.44	2.94	11.37	<0.001
Soil type	**1** (Humic yellow and red / yellows)	-	(Reference category)		
	2 (Humic red)	5.70	2.14	2.67	0.008
	3 (Orthic red)	13.40	3.05	4.39	<0.001
Variety	**1 (**NCo376 and NCo293)	-	(Reference category)		
	2 (N16 and N12)	10.41	2.46	4.24	<0.001
Plant / Ratoon crop	**1** (Plant, Ratoons 1 and 2)	-	(Reference category)		
	2 (Ratoons 3, 4 and 5)	-5.90	2.00	-2.95	0.003
	3 (Ratoons 6, 7, 8 and 9)	-12.75	2.87	-4.45	<0.001
District	Umvoti	-	(Reference category)		
	Kranskop	50.7	14.9	3.39	<0.001
	Richmond	67.1	13.5	4.97	<0.001
Crop cycle	Winter harvest	-	(Reference category)		
	Summer harvest	4.90	1.70	2.89	0.004
Aspect	North and West	-	(Reference category)		
	South and East	14.65	6.03	2.43	0.015
Continuous variables		Regression coefficient	Standard deviation	t-value 513 d.f.#	Probability
N fertilizer (kg ha^{-1})		0.1985	0.0391	5.08	<0.001
Rain total for cycle (mm)		0.0139	0.0027	6.13	<0.001
Rain × North and West		-	(Reference category)		
Rain × South and East		-0.0059	0.0028	-2.12	0.035
Altitude (m a.s.l.)		-0.0492	0.0248	-1.99	0.048
Total K (kg ha^{-1})		0.0147	0.0040	3.70	<0.001
Effective rooting depth (ERD) (cm)		0.3410	0.0833	4.09	<0.001
ERD × Umvoti		-	(Reference category)		
ERD × Kranskop		-0.2840	0.1770	-2.42	0.016
ERD × Richmond		-0.3963	0.1000	-3.96	<0.001
$R^2 = 0.55$	SE = 16.5 TCH				

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

\# Degrees of freedom.

§ Various seasons exhibited similar growth performances and the model was simplified by grouping these.

Specimen calculation:

Yield (TCH) = Constant + Season + Soil type + Variety + Plant / Ratoon crop + District + Crop cycle + Aspect + (0.1985 x N fertilizer) + (0.0139 x Rain total for cycle) + (-0.0492 x Altitude) + (0.0147 x Total K) + (0.341 x ERD)

NOTE: The regression coefficient differences of the categorical variables modify the magnitude of the Constant.

The interaction terms of the continuous variables modify the magnitude of the regression coefficients.

it is difficult to determine the importance of the variables independently of one another in multiple regression because they are seldom truly independent. It follows that interpretations of the biological meaning of parameters included in a model are tenuous until the trends and relations in question can be tested in controlled experiments, although the model output (predicted yield) *per se* is usually useful to managers for planning and resource allocation.

The yield model accounted for 55% of the observed yield variation (Table 2). For the prediction model (Table 3) *season* and *rain total for cycle* were excluded from model building because they can be determined only after harvesting. Since both these variables were important predictors in the yield model, alternative variables were used to predict yield. It was encouraging that most predictors were common to both models although the statistical relations were different. The predictors (Table 2 & 3) can be grouped as those defining physical field attributes, climatic variables and agronomic parameters.

Table 3. Regression coefficients for the prediction model.

Categorical variables	Levels	Regression coefficient differences	Standard deviation	t-value 520 d.f.#	Probability
Constant	Reference categories †	46.7	31.2	1.49	0.136
Soil type	**1** (Humic yellow and red / yellows)	-	(Reference category)		
	2 (Humic red)	4.43	2.36	1.88	0.061
	3 (Orthic red)	13.39	3.38	3.96	<0.001
Plant / Ratoon crop	**1** (Plant, Ratoons 1 and 2)	-	(Reference category)		
	2 (Ratoons 3, 4 and 5)	-10.12	1.81	-5.59	<0.001
	3 (Ratoons 6, 7, 8 and 9)	-20.71	2.56	-8.10	<0.001
District	Umvoti	-	(Reference category)		
	Kranskop	11.04	3.07	3.60	<0.001
	Richmond	12.50	4.22	2.96	0.003
Crop cycle	Winter harvest	-	(Reference category)		
	Summer harvest	7.17	1.83	3.91	<0.001
Continuous variables		Regression coefficient	Standard deviation	t-value 520 d.f.#	Probability
N fertilizer (kg ha^{-1})		0.2190	0.0414	5.29	<0.001
Rain for 12 months' growth (mm)		0.0756	0.0138	5.49	<0.001
(*Rain for 12 months' growth*)^2 (mm^2)		-0.00002114	0.00000598	-3.54	<0.001
Altitude (m a.s.l.)		-0.0595	0.0269	-2.22	0.027
Total K (kg ha^{-1})		0.0184	0.0044	4.18	<0.001
Topsoil Ca:Mg		-0.7840	0.2170	-3.62	<0.001
Effective rooting depth (ERD) (cm)		0.0903	0.0437	2.07	0.039
R^2 = 0.43	SE = 18.4 TCH				

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

\# Degrees of freedom.

Key physical field attributes associated with sugarcane yield were district (locality), aspect, altitude, soil type and effective rooting depth (ERD). Yields differed significantly between Umvoti and Richmond / Kranskop in both models. The magnitude of the partial regression coefficient differences for *district* in the yield model (Table 2) should be considered in association with *ERD × district* which reduces, on average, the expected yield differences between Umvoti and the other districts. The reasons for these yield differences are not easily explained and any statements based on the models are purely speculative. The role of management differences between districts is particularly difficult to quantify, although the general climate is more favourable for sugarcane production at Richmond and Kranskop than at Umvoti.

Field aspect was an important yield predictor and significant interactions with *rain total for cycle* were identified (Table 2). As expected in the southern hemisphere, northern and western aspects generally produced lower yields than southern and eastern aspects. The yield response to water was significantly greater on the hotter northern and western aspects (*Rain x North and West*) than on the cooler southern and eastern aspects (*Rain x South and East*) (Table 2). It is likely that *aspect* did not contribute significantly to the prediction model because *rain total for cycle* was deliberately excluded as a yield predictor.

Altitude was a significant yield predictor in both models, yields being reduced by between 5 TCH and 6 TCH, on average, for every 100 m increase in altitude (Table 2 & 3). Mean field altitude was used with the rationale that it functions as a proxy for temperature; decreasing sugarcane yields are expected in response to decreasing temperature and therefore the negative *altitude* regression coefficients were anticipated. Climatic data were not available at an appropriate scale to confirm this.

Difficulty was experienced defining pure field soil type classes because field boundaries seldom correspond with soil changes and the statistical yield responses to *soil type* may have been distorted by the relatively impure groupings. Since yields were recorded at a field scale, it was impossible to refine the scale of the investigation to investigate pure soil groupings. It is encouraging that in spite of these data limitations, significant yield differences were found between most of the *soil type* levels. Yellow subsoils were associated with lower yields than red subsoils and, on average, superior yields were achieved on orthic topsoils compared with humic topsoils. The reasons for these yield relations remain poorly understood and possible explanations lie beyond the scope of this study.

Yields increased in response to increasing rooting depth and the (associated) increase in water holding capacity. The yield model and prediction model both used ERD as a yield predictor. A strong interaction between *district* and *ERD* was established for the yield model with *ERD* being statistically important only at Umvoti.

When accounting for observed differences in sugarcane yields, physical field attributes were useful only for between-field comparisons and yield differences of up to 100 TCH for successive crops from the same field had to be explained using the climate and agronomic variables. The climate variables *rain total for cycle*, *rain for 12 months' growth* and *season*, were important yield predictors. As expected, rain was positively related to yield (Table 2 & 3). For the yield model (Table 2) this relation was linear over the range of 925 mm to 3573 mm rainfall (Table 1). A significant non-linear relation between yield and *rain for 12 months' growth* was identified for the prediction model (Table 2), indicating that yield responses to high rainfall were proportionally smaller than to low rainfall during the first year's growth. This relation may be partially attributed to the period of incomplete canopy when crop water demand is low and radiation is usually growth-limiting in the Midlands.

The role of *season* in the yield model is not clear. The 13 harvest seasons could be grouped using five classes, all significantly different from each other (Table 2). Partial regression coefficients suggest yield differences of up to 33 TCH between the season classes and it would be useful to understand the underlying relations defined by this variable. Under rainfed conditions agriculturalists expect considerable yield variation between seasons – traditionally ascribed mainly to differences in annual rainfall. Since *rain total for cycle* is included in the yield model, the *season* yield differences are in addition to this effect. It is not possible to explain the dynamics of *season* because of limitations of the regression techniques and the field data. Nevertheless, yields during the 1990's were found to be consistently lower than during the 1980's.

Of the agronomic factors influenced by management, sugarcane variety, plant-ratoon status, crop cycle, N and K nutrition and the topsoil Ca:Mg ratio were important yield predictors. A highly significant yield benefit from using N–varieties in preference to the older NCo–varieties was identified in the field records. The partial regression coefficient differences show that N12 and N16 yield approximately 10 TCH more than NCo376 and NCo293. Although the magnitude of these partial coefficients should be used with the necessary discretion, it is encouraging that field-scale records showed significant benefits from adopting new technology. *Variety* did not contribute significantly to the prediction model because of peculiarities of the data set, and was dropped as a predictor.

Plant / ratoon status was a highly significant yield predictor, confirming that higher yields are generally obtained from younger ratoons than from older ones. The crop groupings should be of particular interest to management for guiding the replant decision. Based on the yield trends investigated for the 535 observations from 13 seasons, there were no significant differences within the crop groups. Plant crops, and first and second ratoons yielded equally well on average. After the second ratoon a significant yield decline was found in the field records and a further significant yield reduction was identified beyond the fifth ratoon (Tables 2 & 3).

Crops harvested in summer produced a significantly higher cane yield, on average, than crops harvested in winter, irrespective of the season in which the cycles started. This relation is contrary to the findings of Inman-Bamber (1994) who reported superior radiation use efficiencies and faster rates of canopy formation for summer ratoons compared with winter ratoons, concluding that the crop starting season was more important than the harvest season. It is only possible to speculate about the reasons for this apparent contradiction because of the limitations of this empirical investigation. Seasonal variations in sugarcane moisture content and extraneous matter may account for a large portion of the yield differences indicated by the regression coefficients.

Fertilizer N was a highly significant yield predictor in both the yield and prediction models. It was not surprising that this yield response was linear over the range of 43-184 kg N ha^{-1} (Table 1) since these rates are moderate in terms of general agriculture. (Up to 300 kg N ha^{-1} an^{-1} is recommended for rainfed tropical pastures in BRG 5 (Manson *et al*., 1993)). The partial regression coefficient for N fertilizer is also particularly large, suggesting a 0.2 TCH increase for every 1 kg N ha^{-1} applied. In real terms the magnitude of this responses is unlikely and the danger of interpreting partial regression coefficients independently of the other variables must be borne in mind.

Total K was defined as the sum of the soil analysis test and fertilizer application since inputs are generally large for nutrients deficient in the soil and small for those present in sufficiency. *Total K* was a highly significant yield predictor in both models. Sugarcane is known to have a high K requirement relative to other crops and luxury uptake is common (Van Dillewijn, 1952). The range of total K for the observations used to develop the models was 1212 kg K ha^{-1} (Table 1) and although leaf analyses generally showed no K deficiencies there may be merit in improving the management of this nutrient.

The significance of the topsoil Ca:Mg ratio in the prediction model suggests some evidence of a nutrient imbalance, possibly created by heavy gypsum applications to some fields with the consequent migration of Mg^{2+} (and K^+) with the SO_4^{2-} down the soil profile. Since *topsoil Ca:Mg* was included as a yield predictor only when *season* was dropped from the prediction model, it is not possible to determine to what extent the ratio merely acts as a proxy for *season*, and to what extent it is directly related to cane yields. Some information about the source variables is always lost when using ratios. The negative yield response to a widening topsoil Ca:Mg ratio has not been documented in SASEX research although negative yield responses to heavy lime applications, and no yield response to gypsum, have been reported (Meyer *et al*., 1991; Schroeder *et al*., 1993; 1995). It is possible that at least a portion of the negative regression coefficient (Table 3) can be attributed to the peculiarities of the data set. This is because *season* showed that yields during the 1990's were generally lower than during the 1980's, while the Ca:Mg ratio during the 1990's tended to be wider than during the 1980's. However, *topsoil Ca* on its own contributed negatively to the regression, although not significantly, while *topsoil Mg* was associated with a significant positive yield response. Since the yield response to *topsoil Mg* was statistically weaker that for *topsoil Ca:Mg*, the ratio was used in the model. It follows that there is evidence of a yield response to increasing Mg nutrition, especially in relation to topsoil Ca and while it is likely that the importance of the relation is overstated, it may provide useful crop management information and warrants further research.

Validation

The yield model fitted the independent data well since 75% of the predicted sugarcane yields were within 20% of the observed yields; 47% of the predicted sugarcane yields were within 10% of the observed yields (Figure 1a). The distribution of observed versus predicted TCH yields (Figure 1b) revealed no systematic pattern, the majority of observations lying within the 95% confidence interval for individual predictions. This confirmed that yield determining factors that were important during the 1998 season had been included in the yield model.

The broad confidence interval for predictions at a field level (95% mean confidence limit ±33.6 TCH) reduces their usefulness for management. Yield predictions for the population mean are, however, considerably more reliable (95% mean confidence limit ±8.4 TCH) and can be used to categorise expected responses of regions / fields with similar characteristics to the independent variables.

The prediction model was developed specifically to predict future sugarcane yields and therefore all predictors are known at least six months in advance of the scheduled harvest date. Since the prediction model R^2 is lower and the SE is larger than for the yield model, poorer yield predictions were expected from this model. The validation results are shown in Figure 2a and 2b.

Although the model under-predicts the sugarcane yields that were achieved during the 1998 season (Figure 2b), 66% of predicted sugarcane yields lie within 20% of the observed yields; 36% of predicted sugarcane yields lie within 10% of the observed yields. The prediction model predicts sugarcane yields independent of season and since the 1998 harvest was particularly large, the underestimation of actual yields should have been anticipated. Similarly, the model should be expected to over-predict yields in poor seasons. It is possible to calibrate the model as information becomes available during the harvest season and improve predictive accuracies.

Ideally the prediction model should predict individual field yields reliably. However, the mean 95% confidence limit for individual predictions (±36.8 TCH) is extremely large and therefore the predicted yield for a field is expected to vary by almost as much as the yield range (82 TCH) for 1998. It follows that the model cannot be realistically applied at this scale. The mean 95% confidence limit for the population mean is ±6.1 TCH. These predictions may be particularly useful to management for resource allocation and harvest planning at a general scale because the confidence limit is less than 10% of the median yield (88 TCH) calculated over 13 years (Table 1).

Conclusions

It is important to remember that the models apply only to the particular area studied, and only to the particular site conditions sampled within the study area. Extrapolation should not be made without testing. The quality of a model cannot be sufficiently judged until it has been in large-scale practical use for some time. Since the data used in this study are typical of

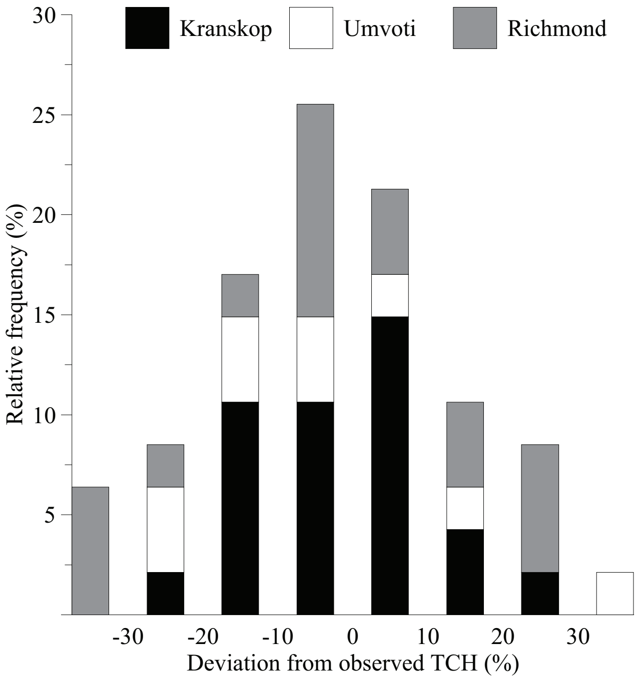

Figure 1a. Graph showing the relative frequency of deviations of predicted sugarcane yields from observed yields (TCH) for the 1998 harvest season using the yield model.

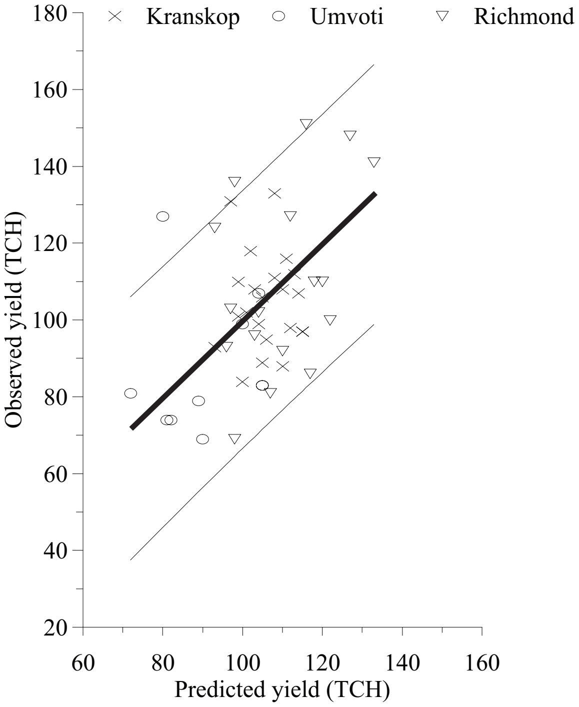

Figure 1b. Graph of observed versus predicted sugarcane yield (TCH) for the 1998 harvest season (heavy solid line) using the yield model. The 95% confidence limits for individual predictions (light solid lines) are shown.

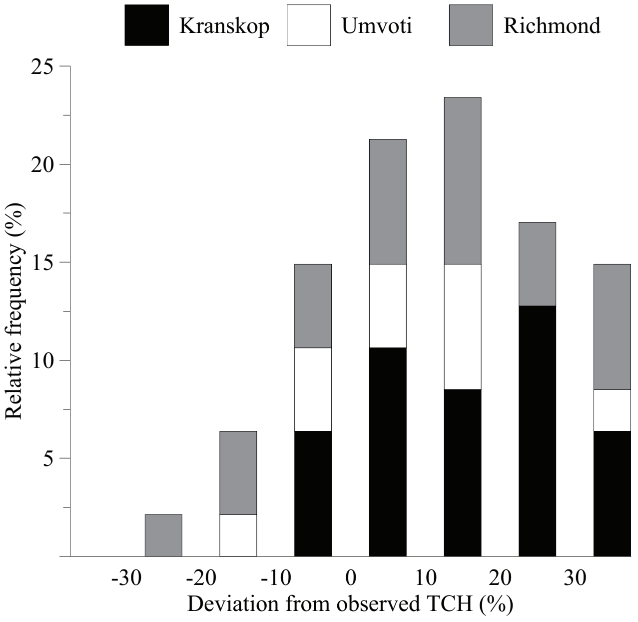

Figure 2a. Graph showing the relative frequency of deviations of predicted sugarcane yields from observed yields (TCH) for the 1998 harvest season using the prediction model.

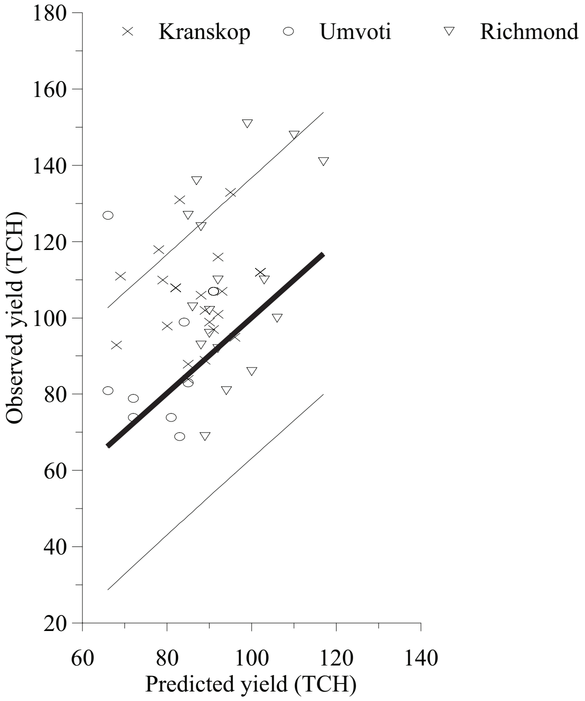

Figure 2b. Graph of observed versus predicted sugarcane yield (TCH) for the 1998 harvest season (heavy solid line) using the prediction model. The 95% confidence interval for individual predictions (light solid lines) is shown.

good commercial records, the availability of large data sets is essential to ensure sufficient appropriate data for modelling. The models identified and quantified the statistical importance of some predictors that have historically been heralded as agronomically important factors defining the yield potential of a site. Other yield relations that were identified as being poorly understood warrant further research.

Acknowledgements

This research was commissioned and funded by Mondi Forests, a division of Mondi Ltd., who provided data recorded by Spencer Holley Agronomic Services and Keith Snyman & Associates. Mondi Forests are thanked for agreeing to the publication of the results. Financial support received from the Sugar Industry Trust Fund for Education and the Foundation for Research and Development is also gratefully acknowledged.

REFERENCES

Brüggemann, EA (1999). Modelling the production potential of land for sugarcane in the KwaZulu-Natal Midlands sugarcane belt. Unpubl. M.Sc (Agric.) thesis. Univ. Natal, Pietermaritzburg, South Africa.

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

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

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

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

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

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

Hulbert, EO and Harding, RL (1980). The computer analysis of farm records from an extension area. *Proc S Afr Sug Technol Ass* 65: 103-108.

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

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

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

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

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

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

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

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