IDENTIFYING CLIMATE VARIABLES HAVING THE GREATEST INFLUENCE ON SUGARCANE YIELDS IN THE TULLY MILL AREA

By

DM SKOCAJ[1,2], YL EVERINGHAM[2]

[1]*Sugar Research Australia, Tully,*
[2]*School of Engineering and Physical Sciences, James Cook University, Townsville*
dskocaj@sugarresearch.com.au

KEYWORDS: Climate, Rainfall, Cane Yield, Variability, Tully.

Abstract

LARGE FLUCTUATIONS in cane yield from one season to the next are problematic for all sectors of the sugar industry. The Wet Tropics region is characterised by high rainfall, excessive soil wetness, low solar radiation and vulnerability to extreme climatic variability. Although many different factors influence productivity, annual fluctuations in cane yield at the farm level in this region are believed to be strongly associated with changes in climatic conditions. To investigate this further, a stepwise linear regression model used atmospheric variables at different times of the growing season to explain Tully mill detrended cane yield data for eight different time blocks. These time blocks ranged from 10 to 80 years. The regression models explained between 32.2 and 94.1% of the variation in detrended cane yields for the Tully mill area. Rainfall, most commonly around spring and summer, was always the first variable entered into the models making it an important predictor. However, the other variables selected for late entry changed over time. Improved yield forecasts coupled with greater knowledge of the influence of climatic conditions on cane yields could be used for a range of management decisions across all sectors of the industry.

Introduction

Large fluctuations in cane yield from one season to the next influences the profitability of all sectors of the sugar industry. The greatest fluctuations often occur in rainfed environments, such as the Wet Tropics, where water supply cannot be controlled. The Wet Tropics sugarcane production area is characterised by high rainfall, excessive soil wetness, low solar radiation and vulnerability to extreme inter-annual climate variability. This provides a difficult management environment, can reduce yield potential and often results in extreme year-to-year cane yield variability. For example the Tully mill area average cane yield of 47 t cane /ha in 2011 was 47.8% lower than 2010 and the lowest since 1948 (Anon., 2012).

Previous research has highlighted the effect of some non-varietal factors on cane yield variability (e.g. Smith, 1991; Leslie and Wilson, 1996; Hurney and Bown, 2000; Lawes *et al.*, 2001, 2002). These factors can be broadly classified as being related to management (time of ratooning, fallow vs plough-out replant, crop cycle duration, cultivation, nutrition and weed, pest and disease control) and location (climate, soil type, topography). Management and location are largely responsible for productivity differences between farms and districts, but these differences tend to remain consistent over time (consistently above or below mill average cane yields). However, at the farm level, where the grower tends not to change management practices dramatically from one year to the next, changes in climatic conditions are believed to be strongly associated with annual fluctuations in cane yield.

It is widely accepted that weather conditions influence cane growth, but specific knowledge relating key atmospheric variables with final cane yield is limited. Research conducted by Smith (1991) on the effect of rainfall variation on cane yield showed that rainfall was responsible for between 34 and 61% (33 and 76%) of the variation in plant (ratoon) cane yields over a 20-year period (1969 to 1988) for three mill areas in far north Queensland.

A review of productivity trends in the Wet Tropics over a 35-year period identified excessive wetness, especially early in the growing season, and low solar radiation adversely effected sugarcane productivity (Leslie and Wilson, 1996; Wilson and Leslie, 1997). Analyses of Tully block productivity data for the period 1988 to 1999 showed that the year of harvest and the month when the crop was ratooned accounted for 20.9% and 11.4% of the variation in cane yield respectively (Lawes *et al.,* 2001).

Subsequent investigations identified crops ratooned from October to December had significantly lower yields the following harvest than those ratooned between July and September (Lawes *et al.,* 2002). However, analysis of block productivity data alone was unable to identify the possible causal factors associated with the year and time of ratooning effect (Lawes *et al.,* 2002).

A different modelling approach was taken by Everingham *et al.* (2002) who discovered a link between the Southern Oscillation Index (SOI) and cane yields. They found deeply negative SOI values during October–November favoured above average cane yields for the Mulgrave and Tully mill areas, and could therefore be used to predict cane yields. Conversely, positive SOI values during October–November favoured below average cane yields.

Knowledge of the key atmospheric variables (rainfall, solar radiation, temperature) and time of year influencing cane yields may help refine yield forecasting techniques and improve decision making capabilities throughout the sugar industry.

At the grower level this may include the fine-tuning of planting systems, nutrition and herbicide programs. Therefore the aim of this paper is to 1) identify which atmospheric variables and time of year have the greatest influence on Tully mill cane yields and 2) investigate if these atmospheric variables remain important irrespective of the historical time period analysed.

Materials and methods

Data collection and pre-processing techniques

Average cane yields (t cane/ha) for the Tully mill area from 1933 to 2012 (80 years) were obtained from Tully Sugar Limited. Many factors influence cane yields so it was important to remove the influence of technological improvement, while still maintaining year-to-year variability in yields that is largely attributed to climate variation. To do this average cane yields for the Tully mill area were detrended according to the procedure outlined by Everingham *et al.* (2003).

Average daily atmospheric values of minimum temperature, maximum temperature and radiation were obtained from the SILO climate data archive (Jeffrey *et al.,* 2001) using the patched point option for the Tully Sugar Limited meteorological station. The patched point data option was selected as it uses original Bureau of Meteorology observations for a particular meteorological station with missing or suspect data ‘patched’ with interpolated values (which are estimates). Unfortunately minimum and maximum temperature and radiation are not measured for Tully Sugar Limited so interpolated values for these variables were used in the analysis. Total daily rainfall data was obtained from Tully Sugar Limited.

The climate data were aligned with the growing season, which was defined from June to May. Single-, two-, three-, four-, five- and six-monthly rolling and seasonal (summer, autumn, winter and spring) average minimum temperature, maximum temperature and radiation values were then calculated. For rainfall the total single-, two-, three-, four-, five- and six-monthly rolling, seasonal and annual values were calculated from the daily dataset. This provided a total of 245 different variables for inclusion in the analysis.

Lastly, the climate data was related to the Tully mill area detrended cane yield for the following year i.e. climate data from June 1932 to May 1933 was analysed against 1933 cane yields and so on.

Analysis method

A stepwise linear regression model (Norušis, 1997) was used to identify which of the 245 variables (independent variables) best explained detrended cane yields (the dependent variable). In this model the selection of independent variables proceeds by steps. Firstly, in a process termed forward selection, the independent variable resulting in the largest increase in multiple R^2 is added to the model (Norušis, 1997). A variable is only added if the change in R^2 reaches a predetermined significance level.

The significance level was set at 0.05 so it was not too easy for variables to enter the model (Norušis, 1997). Next, backward elimination removes the variable that changes R^2 the least, provided that the change in R^2 meets the observed significance level of 0.1 (Norušis, 1997). The process of forward selection and backward elimination continues until no more variables meet the entry criterion.

The order in which variables are entered into the model is also important. Variables entered into the model earlier can be considered more important in explaining the relationship with detrended cane yields than those entered later.

The analysis was run over different historical periods to investigate if the independent variables and/or sequence of those selected changed over time. The time blocks analysed were 1933–2012 (80 years), 1943–2012 (70 years), 1953–2012 (60 years), 1963–2012 (50 years), 1973–2012 (40 years), 1983–2012 (30 years), 1993–2012 (20 years) and 2003–2012 (10 years). The analysis was completed using IBM® SPSS® Statistics for Windows Version 21 (IBM Corp. Released 2012. Armonk, NY: IBM Corp).

To be conservative an additional stopping rule was applied to identify the minimum number of independent variables to include in the final model according to the procedure outlined by Coakes and Steed (2006).

The ratio of cases to independent variables was selected using the minimum requirement of having at least five times more cases than independent variables. For example, a maximum of 16 variables could be used for the 80-year time block, and 2 variables for the 10-year time block.

The adjusted R-squared (R^2_{adj}) value was used to determine the amount of variability in detrended cane yields explained by the model. It was used instead of the R^2 value as it has been adjusted for the number of variables (predictors) included in the model. R^2 tends to overestimate the strength of the association especially if the model has more than one predictor (independent variable) (Norušis, 1997). The R^2_{adj} value is explained by:

$$R^2_{adj} = 1 - \frac{(1-R^2)(N-1)}{N-k-1}$$

R^2 = sample R-square, k = number of predictors and N = total sample size.

The estimate of the residual mean square (S^2) also called the estimate of the error variance was used to investigate the spread of values about the regression models (Norušis, 1997).

The larger the S^2 the more the values are spread out (Norušis, 1997). An S^2 of zero indicates that all values are identical. The S^2 was calculated using the following formula:

$$S^2 = \frac{RSS}{\text{Residual df}}$$

RSS = residual sum of squares and Residual df = residual degrees of freedom.

The beta coefficient was used to indicate the impact of each climate variable selected in the final models on detrended cane yields as the climate variables were measured in different units (rainfall is reported in mm and temperature in ^{o}C).

The sign of the beta coefficient indicates if the variable had a positive or negative impact on detrended cane yield. Beta coefficients are the same as partial regression coefficients when all independent variables have been computed in standardised form (Norušis, 1997).

Results

The R^2_{adj} for each stepwise model is shown in Table 1. For example, the R^2_{adj} values from the 1943–2012 linear models that contain (1) SONDJF rainfall, (2) SONDJF rainfall and July minimum temperature, (3) SONDJF rainfall and July minimum temperature and May maximum temperature are (1) 0.258, (2) 0.331 and (3) 0.369, respectively.

Table 1—The climate variables selected, R^2_{adj}, S^2 and final beta coefficients of the stepwise linear regression models explaining Tully detrended cane yields for eight different time blocks.

Years		Variables included in model	R^2_{adj}	S^2	Beta coefficient
1933–2012	80	1. JASOND rainfall	0.235	50.71	–0.236
		2. DJF rainfall	0.291	47.01	–0.260
		3. July minimum temperature	0.338	43.86	–0.263
		4. May maximum temperature	0.369	41.82	+0.197
1943–2012	70	1. SONDJF rainfall	0.258	54.33	–0.396
		2. July minimum temperature	0.331	49.00	–0.297
		3. May maximum temperature	0.369	46.18	+0.217
1953–2012	60	1. SONDJF rainfall	0.238	59.33	–0.509
		2. July minimum temperature	0.322	52.78	–0.307
1963–2012	50	1. SONDJF rainfall	0.295	58.94	–0.476
		2. July minimum temperature	0.357	53.74	–0.284
1973–2012	40	1. ONDJF rainfall	0.300	69.80	–0.486
		2. July minimum temperature	0.401	59.72	–0.346
1983–2012	30	1. ONDJF rainfall	0.268	67.37	–0.652
		2. NDJ radiation	0.382	56.86	–0.606
		3. May maximum temperature	0.461	49.63	+0.365
		4. ASON radiation	0.555	40.97	+0.524
		5. NDJF maximum temperature	0.624	10.62	–0.365
1993–2012	20	1. JASO rainfall	0.424	53.41	–2.030
		2. ASO rainfall	0.539	42.76	+1.432
		3. May minimum temperature	0.701	27.71	+0.556
		4. AM rainfall	0.769	21.42	–0.295
2003–2012	10	1. JAS rainfall	0.893	13.39	–0.695
		2. JAS minimum temperature	0.941	7.40	–0.339

(The number beside each climate variable indicates the step at which the variable entered the model).

The R^2_{adj} value for the final model increased as the length of time decreased. For the 80-, 70-, 60-, 50-, 40-, 30-, 20- and 10-year time blocks the model explained 36.9, 36.9, 32.2, 35.7, 40.1, 62.4, 76.9 and 94.1% of the variation in detrended cane yields, respectively.

The variables selected by the model differ depending on the historical time period analysed (Table 1). For example, JASOND rainfall accounted for 23.5% of the variability in detrended cane yields for the last 80 years. The combination of JASOND rainfall, DJF rainfall and July minimum and May maximum temperature accounted for 36.9% of the variability for the same time period. However, for the last 10 years JAS rainfall accounted for 89.3% of the variability in detrended cane yields and when combined with JAS minimum temperature, 94.1%.

Total rainfall for the six-month period July to December (JASOND) was the first variable selected for the longest time block analysed, 80 years. For the 50-, 60- and 70-year time blocks total, six-monthly rainfall was also important. However, the time of year shifted and it was the combined total of spring and summer (SONDJF) rainfall first selected by the model. The 30- and 40- year time blocks changed slightly and total rainfall for the five-month period October to February (ONDJF) was the first variable selected. The model entered July to September (JAS) and July to October (JASO) rainfall first for the 10- and 20-year datasets respectively.

The beta coefficients of the final stepwise model for each time block are also shown in Table 1. Rainfall, except for ASO rainfall in the last 20 years, always had a negative impact on detrended cane yields. Other variables having a negative impact on yield were July minimum temperature (80-, 70-, 60-, 50- and 40-year time blocks), NDJ radiation and NDJF maximum temperature (30 years) and JAS minimum temperature (10 years). Variables having a positive impact on yield included May maximum temperature (80-, 70- and 30-year time blocks), May minimum temperature (20 years) and ASON radiation (30 years).

The final R^2_{adj} value for each of the different time blocks analysed is shown in Figure 1a. After decreasing rapidly over the short term (last 10 to 40 years) the adjusted R^2_{adj} value reached a plateau when 40 or more years' of data were supplied to the model. Changes to the final estimate of the residual mean square (S^2) for each of the different time blocks analysed is shown in Figure 1b. After increasing rapidly over the last 10 to 40 years the S^2 also reached a peak when blocks of 40 or more years were used by the model and then started to decrease slowly.

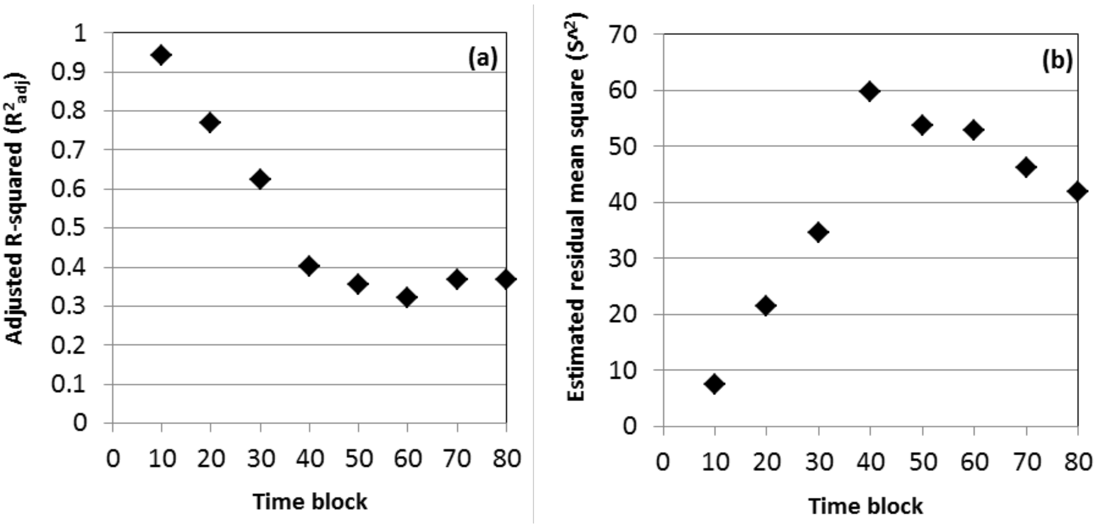

Fig. 1—Changes in the R^2_{adj} (a) and S^2 (b) values for each time block analysed.

The predicted detrended cane yield anomaly for the different time blocks was calculated using the corresponding regression model. The spread of the predicted yield anomalies about the actual yield anomalies for the 80-, 70-, 60-, 50-, 40-, 30-, 20- and 10-year models are shown in Figures 2 a, b, c, d, e, f, g and h, respectively. The plots with a greater sample size (Figure 2a-e) tended to have a higher amount of scatter than those with a small sample size (Figures 2f–h).

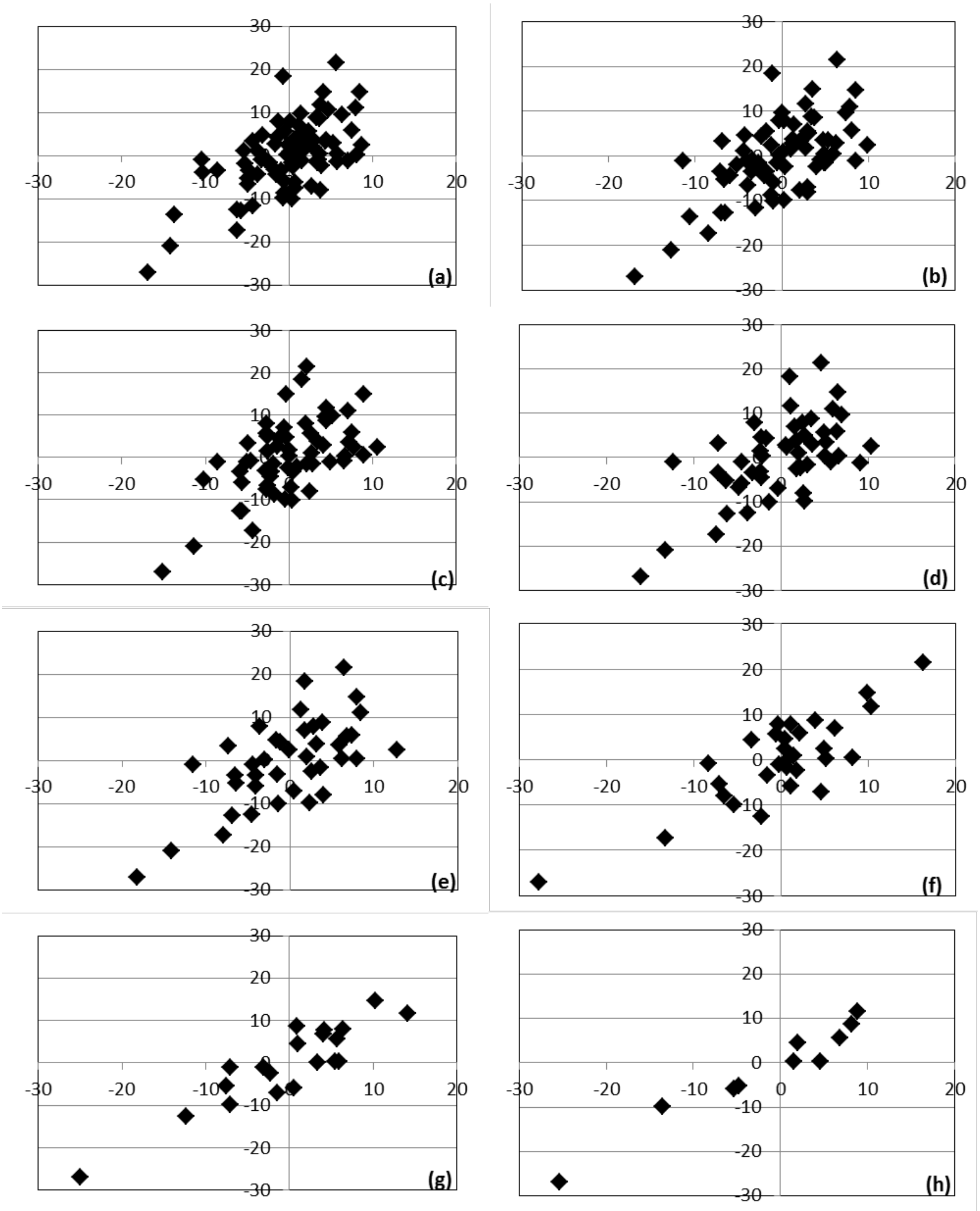

Fig. 2—Actual (y axis) vs. predicted (x axis) yield anomalies from the regression models for each of the eight historical time blocks analysed. (a) 1933–2012, (b) 1943–2012, (c) 1953–2012, (d) 1963–2012, (e) 1973–2012, (f) 1983–2012, (g) 1993–2012 and (h) 2003–2013.

Discussion

Rainfall has been selected as the first explanatory factor in all models and usually reduced detrended cane yields. Excessive rainfall coincides with low solar radiation and extreme waterlogging which adversely affects crop growth, increases nutrient losses (especially nitrogen) and may prevent crop production practices being completed in a timely manner (which may increase weed competition, delay fertiliser application or hilling up of plant cane).

Young roots of early ratoon cane can also be permanently injured by relatively short (approximately one week) periods of waterlogging (Rudd and Chardon, 1977). In addition, the productivity review conducted by Leslie and Wilson (1996) mentioned an environment of extreme soil wetness was having a major influence on cane growth, especially early ratoon cane up to 1 m high, for Babinda.

Previous research has also linked rainfall to cane yield variability. Smith (1991) used stepwise linear regression analysis to identify the main weather parameters (total rainfall and number of wet days from July to June and monthly rainfall) associated with changes in cane yield for mill areas including Tully.

A rainfall model combining December and January rainfall was shown to account for 39% and 47% of the variability in plant and ratoon cane yields respectively for the Tully mill area over the 20 years analysed (1969 to 1988). Everingham *et al.* (2002) inferred that the link between October–November SOI phase and cane yields could be due to an association between the October–November SOI phase and summer rainfall (i.e. deeply negative October–November SOI phase is associated with lower summer rainfall).

Most recently a productivity review of the Herbert region found a strong correlation between November rainfall and final cane yields using linear regression analysis. November rainfall accounted for 43.4% of the annual cane yield variation experienced in the Herbert region over an 18 year period, although there were large differences between productivity zones (Garside, 2013). However, the rainfall variables identified in this analysis were not the same as in previous research.

Previous research identified November or December and January or summer rainfall as having the greatest impact on yields in the Wet Tropics region whereas in our analysis the models commonly entered rainfall around spring and summer as the first variable.

It was surprising to find that rainfall earlier in the growing season (late winter, early spring) was more important in the 10- and 20-year models than rainfall later in the growing season (around spring and summer) which was important for the 30-, 40-, 50-, 60- and 70-year datasets. This may be due to the fact that this analysis considered different historical time periods, climate variables other than rainfall and much longer time blocks (Smith, 1991; Leslie and Wilson, 1996; Garside, 2013).

Where radiation has been selected we suspect this is because of its association with rainfall (high solar radiation = low rainfall and vice versa) and its importance in physiological processes (photosynthesis). However, we cannot confidently explain the physiological phenomenon associated with the selection of other common variables (i.e. May maximum temperature and July minimum temperature). May coincides with the very end of the growing season and only a small proportion of the next crop is exposed to conditions in July.

Our analysis focused on trying to quantify the impact of atmospheric variables on detrended cane yields and if the same atmospheric variables (and time of year) remained important irrespective of the historical time period analysed. The amount of variability in detrended cane yields attributable to climatic conditions ranged from 32.2% (1953–2012) to 94.1% (2003–2012). There are obviously other factors such as mechanisation, time of ratooning, land expansion, changes to farming systems and growing inputs (e.g. N fertiliser, herbicides) influencing detrended cane yields that we did not incorporate into our models.

The data presented in Table 1 show there was some commonality in the variables entered early (i.e. rainfall around spring and summer) in the model. July minimum temperature was also commonly selected as a late entry in models with 40 or more years of data. However there were no other variables consistently entered late in the model across the different time blocks.

The stepwise approach was sensitive to the length of the time block. The R^2_{adj} steadily decreased and the S^2 steadily increased until the time interval reached 40 years. Once the time interval reaches 40 years and beyond there is little change in the R^2_{adj} or S^2 values.

Although more research is needed, it is reasonable to hypothesise that the true amount of variability explained by atmospheric variables via a simple linear regression approach is between 30 and 40%. Model confidence is clearly dependent on the length of the time block.

Conclusion

In summary, the key findings resulting from this research are:

- The amount of variability in detrended cane yields explained by the climate variables was highly dependent on the length of the time block. The R^2_{adj} ranged from 32.2% (1953–2012) to 94.1% (2003–2012).
- The R^2_{adj} steadily decreased and the S^2 increased until the time interval reached 40 years of data. This suggests model confidence may have been inflated when less than 40 years of data was entered.
- Model confidence depends on the length of the time block.
- Rainfall mostly had a negative impact on detrended cane yields and was the first variable selected in all models. However, there has been a shift in the time of year having the greatest influence on detrended cane yields. In the 10- and 20-year analysis rainfall earlier in the growing season (late winter, early spring) was more important than rainfall later in the growing season (spring and summer/ late spring and summer).
- July minimum temperature featured as a late entry in models with 40 or more years of data (the 1933-, 1943-, 1953-, 1963- and 1973-2012 models). May maximum temperature was also a late entry in the 80-, 70- and 30-year models. However, other variables (i.e. NDJ and ASON radiation, NDJF maximum temperature, May and JAS minimum temperature, AM rainfall) entering late into the models were not common, suggesting that they might be unstable predictors.

The atmospheric data were from a point source but detrended cane yields were representative of all districts supplying Tully mill. Future research could include rerunning the analysis with plant and ratoon yield data to see if the model is sensitive to crop class and completing the analysis for different mill areas and districts within a mill area (where sufficient climate data is available).

This would allow the identification of spatial differences across a region (Wet Tropics) and within a mill area (i.e. Mossman, Mulgrave, Tablelands, South Johnstone, Tully), which may facilitate the fine tuning of yield forecasting and harvest scheduling. The time of ratooning effect on cane yields could also be incorporated into future investigations.

Lawes *et al*. (2002) identified year and the time of ratooning as having a major influence on cane yield variation in the Tully mill area and suggested that these factors may provide a surrogate measure of the conditions experienced when new ratoon crops are initiated.

Obviously the time of ratooning determines the timing of the crop-growth period. It would be interesting to investigate if crops ratooned early in the season (July to September) are less sensitive to rainfall around spring and summer than crops ratooned later (October to December). It is also possible that the SOI or sea surface temperatures (SST) may be better suited than atmospheric variables for the prediction of sugarcane yields in the Wet Tropics region and this is the focus of future research.

Acknowledgements

We are grateful to Tully Sugar Limited for supplying daily rainfall and mill average cane yield data and acknowledge financial contributions from the Sugar Research and Development Corporation (SRDC) and Sugar Research Australia (formerly BSES Limited).

REFERENCES

Anon. (2012) Tully district comprehensive area productivity analysis 2011. Tully Sugar Industry, Tully.

Coakes S, Steed L (2006) 'SPSS version 13.0 for Windows: analysis without anguish'. (John Wiley & Sons Australia, Ltd: Milton).

Everingham YL, Muchow RC, Stone RC, Coomans DH (2003) Using Southern Oscillation Index phases to forecast sugarcane yields: A case study for northeastern Australia. *International Journal of Climatology* **23**, 1211–1218.

Garside AL (2013) Review of productivity trends in the Herbert sugarcane growing region. Herbert Cane Productivity Services Limited Publication 1/2013.

Hurney AP, Bown PA (2000) Final Report SRDC Project BSS159: Farm assessment of productivity limits in the wet tropics. (Bureau of Sugar Experiment Stations, Queensland, Australia).

Lawes RA, Lawn RJ, Wegener MK, Basford KE (2001) Spatial variation of sugarcane productivity in the Tully mill district: is it worth worrying about? *Proceedings of the Australian Society of Sugar Cane Technologists* **23**, 97–101.

Lawes RA, McDonald LM, Wegener MK, Basford KE, Lawn RJ (2002) Factors affecting cane yield and commercial cane sugar in the Tully district. *Australian Journal of Experimental Agriculture* **42**, 473–480.

Leslie JK, Wilson GL (1996) Productivity trends in sugarcane in the Wet Tropics. Technical Report 1/96, Sugar Research and Development Corporation and Cooperative Research Centre for Sustainable Sugar Production.

Norušis MJ (1997) 'SPSS® 7.5 Guide to Data Analysis'. (Prentice-Hall, Inc.: New Jersey)

Rudd AV, Chardon CW (1977) The effects of drainage on cane yields as measured by water-table heights in the Macknade mill area. *Proceedings of the Queensland Society of Sugar Cane Technologists* **44**, 111–117.

Smith MA (1991) Recent production trends in the Queensland sugar industry. (Bureau of Sugar Experiment Stations, Queensland, Australia).

Wilson GL, Leslie JK (1997) Productivity trends in sugarcane in the Wet Tropics. *Proceedings of the Australian Society of Sugar Cane Technologists* **19**, 21–29.