Forecasting Australian Sugar Yields Using Phases of the Southern Oscillation Index

Y.L. Everingham[a], R.C. Muchow[b] and R.C. Stone[c]

[a] *CSIRO Sustainable Ecosystems, Davies Laboratory, Townsville, Queensland 4814, Australia (Yvette.Everingham@cse.csiro.au)*

[b] *SRDC, PO Box 12050, Brisbane Elizabeth St, Queensland, 4002, Australia (rmuchow@srdc.gov.au)*

[c] *Queensland Department of Primary Industries, 203 Tor Street, Toowoomba, Queensland 4350, Australia (stoner@.dpi.qld.gov.au)*

Abstract: Yields for the Australian sugar industry can vary seasonally and regionally. Advanced knowledge of likely sugar productivity levels for mill regions in a particular season would assist marketers to forward sell Australian sugar, and allow mill managers and harvest operators to better plan for the coming season. Given that climate is a key driver of productivity, the purpose of this paper is to investigate the potential usefulness of a climate forecast system which incorporates five patterns or phases of the Southern Oscillation Index (SOI) to forecast sugar yields. The chance of obtaining a sugar yield above the long-term median was computed for each SOI phase across eight regions which span the coastline of Queensland, where most of Australia's sugarcane is grown. Results indicate that for certain regions, the chance of obtaining an above average crop can be greatly increased, and in some cases decreased depending on the phase of the SOI. Since many decisions in the Australian sugar industry are based on crop size, the SOI phases provide a useful tool for enhancing decision making and risk management capability for the industry.

Keywords: Sugar; Yield; Forecasting; Southern Oscillation Index

1. INTRODUCTION

The Australian sugar industry contributes between one to two billion dollars annually to the Australian economy with more than 5 million tonnes of sugar produced from in excess of 400000 hectares of farmland harvested. Most of this farming land occupies the narrow coastal strip of Eastern Australia which extends some 2100 km between the latitudes of 15 to 30 degrees south. Climate across this coastal strip is remarkably diverse which leads to regional variation in production. Productivity can also vary seasonally.

The typical harvest season for the Australian sugar industry commences around June and extends to November/December of the same year. Estimates of productivity levels for the next season are usually made at the end of the previous harvest season around November. This allows industry decision-makers to begin planning for the following season.

Knowledge of sugarcane productivity levels is important for industry decision-makers from each component of the value chain i.e. from the farming and harvesting sectors through to milling and marketing sectors of the industry.

Many decisions are made based on the likely size of the crop. One such decision relates to timing the start of the harvest season. Underestimating the size of the crop could cause the start of the season to be delayed with risk of the harvest season extending into the wet season increased [Muchow and Wood, 1996]. Overestimating the size of the crop could cause the season to start unnecessarily earlier than required. With sugarcane levels typically being higher towards the end of the harvest season [Muchow et. al., 1997], harvesting earlier than required could result in a lowering of the overall sugar content thereby reducing profit levels. Decisions also made at the marketing end of the value chain relate to the selling and storage of sugar. Overestimates pose problems in meeting forward export commitments and underestimates can lead to potential shortfalls in managing limited storage requirements.

It is clear that if more accurate yield forecasting methods existed then such decisions could be better managed. Whilst yield estimates are made progressively throughout the year, this paper focuses on improving the accuracy of initial estimates of productivity made at the end of November.

The El Niño – Southern Oscillation (ENSO) phenomenon [McBride and Nicholls, 1983] is potentially one factor that contributes to the variability in sugarcane yields between seasons. Kuhnell [1994] found an inverse relationship between the Southern Oscillation Index (SOI), (a measure of the strength of ENSO), and productivity measured in both tonnes of cane per hectare (TCPH) and tonnes of sugar per hectare (TSPH) seven to eleven months prior to the commencement of harvest. This was especially the case for northern and southern sugar growing districts along the coast of Queensland. A weak positive relationship for the central coast of Queensland was found but this relationship did not become evident until just before the start of the harvest. Kuhnell [1994] noted that the relationship between productivity and the SOI was stronger for cane productivity (TCPH) than for sugar productivity (TSPH), but in either case only a small proportion of the variation of production can actually be explained by this simplistic relationship with the SOI.

Instead of using actual values of the SOI, an alternative approach is to consider phases of the SOI [Stone and Auliciems, 1992; Stone et. al., 1996]. The SOI 'phases' represent the relative importance of both the consistency and the change in month to month values of the SOI, and were derived by a principal components analysis and cluster analysis of a time series of the SOI. The five clusters or phases are referred to as 'consistently negative', 'consistently positive', 'rapidly falling', 'rapidly rising' and 'consistently near zero'. Each month since 1887 can be classified into one of these phases.

The five phase SOI climate forecasting system has previously been used for forecasting agricultural-based responses for the peanut [Meinke et. al., 1996] and wheat industries [Hammer et. al., 1996]. For the sugar industry, Singels and Bezuidenhout [1998 and 1999] identified, by examining graphical displays and simple summary statistics, that South African sugarcane yields tended to be reduced in years when the November SOI phase, was 'consistently negative'. Climate conditions occurring in Australian sugarcane growing regions can be different to climate conditions occurring in cane growing regions of South Africa. The work by Singels and Bezuidenhout [Singels and Bezuidenhout, 1998], [Singels and Bezuidenhout, 1999] was extended by Everingham et. al. [2001] by examining the effect that each of the five SOI phases has on Australian sugarcane yields. Everingham et. al. [2000] made use of Monte Carlo procedures [Good, 1997] to determine which of the five SOI phases were most useful for indicating when Australian sugarcane yields (TCPH) are likely to be above (or below) the long term median (detrended) for eight mill locations of relevance to the Australian sugar industry – Mulgrave, Tully, Victoria, Inkerman, Proserpine, Marian, Isis, Moreton. Results indicated that

- A 'consistently negative' November SOI phase offers the potential to improve yield forecasting for Mulgrave, Tully and Isis.
- A 'consistently positive' November SOI phase offers the potential to improve yield forecasting for Tully, Isis and Moreton.
- A 'consistently falling' November SOI phase offers the potential to improve yield forecasting for Tully and Isis.
- A 'rapidly rising' November SOI phase offers the potential to improve yield forecasting for Tully.
- A 'consistently near zero' November SOI phase offers the potential to improve yield forecasting for Marian.

Thus, the five phase SOI system offers the potential to improve sugarcane estimates but success is likely to vary with geographical location and SOI phase.

Since sugar is the actual quantity that marketers sell, this paper seeks to identify the relationship between November SOI phases and sugar yields measured in tonnes of sugar per hectare for the next season. This information can then provide a more direct linkage for marketers to begin planning export commitments for the next season.

2. DATA and METHODS

2.1 Sugar Productivity Data

The sugar productivity data considered in this paper consist of sugar tonnages per hectare harvested for eight mill locations which span the eastern coast of Queensland (Mulgrave, Tully, Victoria, Inkerman, Proserpine, Marian, Isis, Moreton). These mills along with the start and finish times of the TSPH data are displayed in

Table 1. Table 1 also contains information regarding the approximate locations of the mills in decimal degrees as well as the mean annual rainfall (mm) recorded at nearby Bureau of meteorological sites. These summary rainfall data were extracted from Rainman version 3.2 [Clewett et. al. 1999].

Table 1. Latitude and longitude of selected mills and mean annual rainfall. The last two columns show the start and finish times of actual yield data obtained for each mill.

Mill	Latitude	Longitude	Rainfall	Start	Finish
MUL	-17.09	145.79	1940	1942	1999
TUL	-17.94	145.93	4056	1945	1999
VIC	-18.65	146.18	2064	1900	1999
INK	-19.77	147.46	903	1942	1999
PRO	-20.55	148.66	1672	1918	1999
MAR	-21.16	148.87	1464	1942	1999
ISI	-25.10	152.56	979	1942	1999
MOR	-26.62	152.97	1695	1942	1999

2.2 Detrending Process

Owing largely to modifications in farming and harvesting practices, the time series of TSPH data contains trends. Prior to any formal analyses it is necessary to remove such trends from the data whilst at the same time capturing year-to-year variability in yields that can be largely attributed to climate variation. The detrending procedure applied in this paper consists of two steps. The first step fits a robust smoother which incorporates a running median procedure [SPLUS 2000 Guide to Statistics, 1999] to each time series of TSPH values. The TSPH data fitted by the smoother are then subtracted from the observed TSPH values to produce a TSPH anomaly.

The detrending procedure can be summarised by

$$a_{ij} = y_{ij} - x_{ij} \qquad (1)$$

where

x_{ij} is equal to the actual TSPH value for mill i in year j, y_{ij} is equal to the smoothed TSPH values (y_{ij}=f(x_{ij}) where f() is the smoothing function) and a_{ij} is the TSPH anomaly for mill i and year j.

2.3 Climate Forecasting Technique

The 'five phase SOI system' is applied in this paper [Stone et al., 1996; Stone and Auliciems, 1992]. This climate forecast system utilises pre-determined clusters of the SOI which represent the relative importance of both the consistency and change in month to month values of the SOI. The five clusters or phases of the SOI are referred to as: 'consistently negative' (neg), 'consistently positive' (pos), 'rapidly falling' (fal), 'rapidly rising' (ris) and 'consistently near zero' (nz). By comparing the average SOI from one month to the next, it is possible to assign the consecutive months into this phase typology.

2.4 Analysis Methods

Since initial estimates of sugar tonnages are made during the November prior to the next harvest (which usually commences around June of the following year) we compute the probability of obtaining a TSPH anomaly greater than the median conditioned on the November SOI phases. By partitioning sugar yields into five different groups which correspond to each November SOI phase, it is possible to determine how many TSPH values exceed the overall median for each phase grouping. In general terms, the probability of exceeding the median is equal to 0.50, but for certain years which experience specific SOI patterns this probability may be conditional on the SOI phase and can shift accordingly.

Probabilities that differ from 0.5 are of particular interest to industry since this indicates that above (or below) average yields are more (or less) likely. To determine the approximate significance of each probability conditional on a specific SOI phase, a Monte Carlo procedure [Good, 1997] was employed. Ten thousand random permutations of the November SOI phases were generated using SPLUS 2000. For each SOI phase in each permutation, the probability of exceeding the median TSPH anomaly was computed. The proportion of probabilities (from the possible 10,000 values) that were as extreme or more extreme than the conditional probabilities computed from the actual SOI phases were computed. The smaller the proportion the less likely that the observed probabilities are due to chance. The actual cut-off value for determining whether the probability for exceeding the median is due to chance is somewhat arbitrary. We adopt a cut-off criterion of 0.10, or in other words, a significance level of 10%.

3. RESULTS AND DISCUSSION

Figure 1 displays the effect of the detrending process described in Section 2.3 when applied to yields from the Proserpine Mill. The original and smoothed TSPH time series are shown in Figure 1(a) and the detrended time series which produces the TSPH anomalies is displayed in Figure 1(b). For each mill the value of the median TSPH anomaly value is approximately zero.

The percent chance of exceeding the median TSPH anomaly for each mill conditioned upon each SOI phase is shown in Table 2. The value shown in brackets in each cell is the percentage of the 10000 probabilities generated from the Monte Carlo process that were as extreme or more extreme than the observed probability. Taking Proserpine, as an example, – when the November SOI phase was consistently negative 31% of cases resulted in a TSPH anomaly greater than the median and only 8.5% of the 10000 simulated probabilities generated from the permuted SOI phases were less than or equal to 31%. The approximate P-values computed from the Monte Carlo process are one sided, for example, when the November SOI phase was consistently near zero, 11.6% of the 10000 simulated probabilities were greater than or equal to 62%. Bracketed values less than 10% are indicated with an asterisk.

Results indicate that not all the November SOI phases are useful for identifying the likelihood of the TSPH anomaly exceeding the median. This result is also dependent upon location. Probabilities which deviate from 50% at an approximate significance level of 10% suggest that:

- A 'consistently negative' November SOI phase offers the potential to improve yield forecasting for Tully, Proserpine and Isis.
- A 'consistently positive' November SOI phase offers the potential to improve yield forecasting for Tully and Isis.
- A 'rapidly rising' November SOI phase offers the potential to improve yield forecasting for Tully.

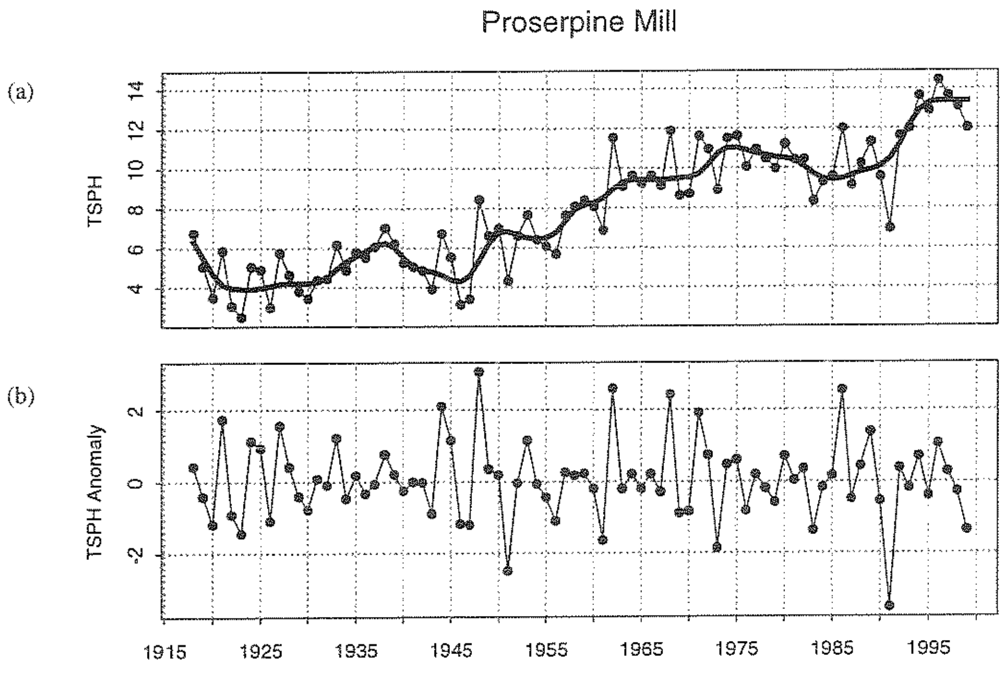

Figure 1. (a) Time series plots displaying the original and smoothed sugarcane yields measured in tonnes of sugar per hectare (TCPH) for Proserpine Mill. (b) A time series plot for Proserpine Mill displaying the TSPH anomalies which is the difference between the original and smoothed sugarcane yields (see Eq. 1).

Table 2. Percent chance (*P*) of exceeding the median TSPH anomaly for each of the 5 SOI phases for November. The number in brackets represents the approximate percentage chance of obtaining a value as extreme or more extreme than *P*. Bracketed values less than 10% are indicated by an asterisk.

Mill	Negative	Positive	Falling	Rising	Near Zero
Mulgrave	58 (37.3)	42 (37.9)	71 (20.9)	50 (62.4)	41 (28.9)
Tully	82 (1.8)*	27 (9.8)*	50 (65.0)	40 (40.0)	47 (53.6)
Victoria	55 (46.6)	64 (22.7)	33 (35.2)	20 (4.4)*	59 (25.5)
Inkerman	50 (61.8)	50 (61.9)	43 (49.5)	50 (63.9)	53 (50.2)
Proserpine	31 (8.5)*	43 (30.2)	56(49.6)	58 (38.0)	62 (11.6)
Marian	42 (36.1)	58 (37.7)	43 (50.6)	30 (15.4)	65 (12.6)
Isis	8 (0.1)*	75 (5.5)*	43 (50.2)	70 (14.9)	53 (49.7)
Moreton	50 (62.7)	33 (16.3)	57 (50.9)	70 (14.4)	47 (50.1)

The choice of the significance level is in many ways arbitrary. If the approximate significance level were relaxed from 10% to 20% then results would indicate that:

- Yield estimates for Moreton could also be improved when the November SOI phase is 'consistently positive'.
- Yield estimates for Marian, Isis and Moreton could also be improved when the November SOI phase is 'rapidly rising'.
- Yield estimates for Proserpine and Marian could also be improved when the November SOI phase is 'consistently near zero'.

In some instances a hypothesis for explaining physical phenomena associated with the variation in TSPH yields for different SOI phases across different locations can be forthcoming. It appears that in the wetter districts of the north such as Mulgrave, a 'consistently negative' SOI phase 'brings relief through drier conditions' in that yields are likely to be higher in such years. For drier areas such as Isis, and also where irrigation is limited, the 'consistently negative' SOI phase has the converse effect to Mulgrave. For Isis a 'consistently negative' November SOI phase indicates that the chances of having a TSPH anomaly greater than average is somewhat reduced. The relationship between TSPH anomalies for a 'consistently positive' SOI phase for Tully and Isis are also contrasting. A 'consistently positive' SOI phase for Tully *reduces* the chances of experiencing a TSPH anomaly greater than the median, whilst for Isis the chances are increased. A 'consistently negative' November SOI phases reduces the chance of experiencing above average rainfall over the critical growing period December-February for most of the locations listed in Table 1, whereas a 'consistently positive' SOI phase increases the chance of above median or average rainfall.

For the remaining situations a physical explanation detailing why the TSPH yields vary with SOI phase and location is somewhat of a challenge at this stage of the research. If one is willing to accept what has happened to TSPH anomalies historically, than some value can be placed on the SOI phases for assisting industry decision makers to plan for the coming season.

4. CONCLUSION

The five phase SOI system offers the potential to improve estimates of a yield anomaly which is based on the ratio of tonnes of sugar to hectares harvested, and to improve risk management for decisions which are based on knowledge of crop size. The usefulness of the system was found to vary with SOI phase as well as geographical location.

Future research should investigate whether the direction and magnitude of the yield anomaly can be predicted with greater accuracy by also considering other potential predictor variables such as sea surface temperatures. Industry has already used the results contained in this paper to assist with the planning of the Australian 2001 sugar season.

5. ACKNOWLEDGEMENTS

The authors wish to thank Bernie Milford, Owen Crees and Heidi Horan for their assistance in collating the yield databases. This research has been funded in part by the Sugar Research and Development Corporation (SRDC) and the Climate Variability and Agriculture R&D Program (CVAP).

6. REFERENCES

Clewett J.F., P.G. Smith, I.J. Partridge, D.A. George, A. Peacock, Australian Rainman Version 3: An integrated software package of Rainfall Information for Better Management. QI98071, Department of Primary Industries Queensland, 1999.

Everingham Y.L., R.C. Stone, R.C. Muchow, Enhancing Sugarcane Yield Forecasting Capability Using SOI Phases: A Case Study for North Eastern Australia. *International Journal of Climatology.* In prep 2001.

Good P.I., *Permutation tests: a practical guide to resampling methods for testing hypotheses* (3rd edn). Springer-Verlag New York Inc: New York, 1997.

Hammer G.L., D.P. Holzworth, R.C. Stone, The value of skill in seasonal climate forecasting to wheat crop management in a region with high climatic variability. *Australian Journal of Agricultural Research* **47:** 717-737, 1996.

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

McBride J.L. and N. Nicholls, Seasonal relationships between Australian rainfall and the Sourthern Oscillation. *Monthly Weather Review* 111:1998-2004, 1983.

Meinke H., R.C. Stone, G.L. Hammer, SOI phases and climatic risk to peanut production: A case study for northern Australia. *International Journal of Climatology* **16:** 783-789, 1996.

Muchow R.C., M.J. Robertson, and B.A. Keating, Limits to the Australian sugar industry: climatic and biological factors. In: Keating BA, and Wilson JR, (eds), Intensive Sugarcane Production: Meeting the Challenges Beyond 2000. CAB International, Wallingford, UK, pp. 37-54, 1997.

Muchow R.C., and A.W. Wood, Rainfall risk and scheduling the harvest of sugarcane. In: Wilson, JR, Hogarth DM, Campbell JA, and Garside AL, (eds). *Sugarcane: research towards efficient and sustainable production.* CSIRO Division of Tropical Crops and Pastures, Brisbane, 26-27, 1996.

Singels A. and C.N. Bezuidenhout, ENSO, the South African climate and sugarcane production. Proceedings of the South African Sugar Technologists Association 72: 10-11, 1998.

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

SPLUS 2000 Guide to Statistics, Volume 1, Data Analysis Products Division, MathSoft Seattle, WA, 1999.

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

Stone R.C., G. Hammer, T. Marcussen, Prediction of global rainfall probabilities using phases of the Southern Oscillation Index. *Nature* 384: 252-255, 1996.