An Integrated Modelling Approach to Enhancing Sugarcane Profitability

A. Higgins and **M. Haynes**

CRC for Sustainable Sugar Production, CSIRO Sustainable Ecosystems
120 Meiers Road, Indooroopilly, 4068, Australia
(Andrew.Higgins@cse.csiro.au and Michele.Haynes@cse.csiro.au)

Abstract: Increasing cost/price pressure has forced the Australian Sugar Industry to seek innovative avenues for increasing profitability. To address this, the industry saw opportunities for increasing productivity and hence profitability through optimising the harvest date of sugarcane, accounting for geographical differences in cane yield and the sugar content of cane for different harvest dates throughout the harvesting season. To optimise the harvest date of farm paddocks within a mill region, operations research tools were developed through an action research approach to successfully incorporate relevant industry capacities associated with harvesting, transport and milling. Through the use of several years of historical productivity data from each of the mill regions, a statistical model was also developed to estimate the relative sugar content of cane by harvest date which is a key driver for the optimisation model. Using the developed optimisation model to simulate various cane supply options, significant gains in industry profitability are possible. The simulated options include modification of farm management practices, regulated grower and harvester equity schemes as well as transport capacities. Achieving implementation of any of these options in practice required model outputs that are relevant to individual growers. Integration of statistical models with an optimisation model provided the tools to better capture the variation in sugar content among individual farms over a number of years. This approach resulted in increased confidence that implementation would deliver the potential benefits even with unknown future climate impacts. Grower participation in pilot studies to implement the alternative cane supply options in 2000 and 2001 was a significant move forward in the participatory research process between CRC Sugar, CSIRO and the Australian Sugar Industry.

Keywords: Sugar production; Operations research; Statistical models; Action research

1. INTRODUCTION

The harvest date of sugar cane is a key determinant of sugar yield and hence net revenue. Limited mill crushing and cane transport capacity dictate that the harvesting of cane be carried out over a harvest season of several months. In Australia, the harvest season currently extends from early winter to late spring, since this is when the CCS (that is, the percentage of extractable sugar from cane) is at its highest. A typical mill region in Australia, producing around 1.5 million tonnes of cane, consists of 200 - 300 privately owned farms (contained in 8 to 12 geographical districts), which are contracted to between 20 and 100 harvesting groups with about 3 to 12 farms to a group. Since the farms and harvesting groups are privately owned, regulated farm and harvesting group equity is enforced to ensure that harvesting progresses evenly on farms throughout the harvest season, thus preventing any farm from being unfairly disadvantaged. Each farm usually ranges in size between 20 and 100 paddocks, with crops differing in terms of cane variety, crop class and age at harvest (elapsed time since planted or last harvested). A crop class refers to whether the cane on a farm paddock is plant crop (around 1 year after planting), first ratoon (around 2 years after planting), second ratoon and so on. A ratoon is the re-growth of the plant from the established plant base (stool) after the top is removed at harvest, just like mowing the lawn. After each subsequent ratoon within a crop cycle, the sugar yield (equals CCS times cane yield) generally decreases until the low sugar yield makes it more economical to plough out and plant a new crop. After the cane is cut by the harvesting group in the field, it is transported to the mill by either road or dedicated narrow gauge rail transport owned by the mill. Transport can be a major restriction (along with harvesting capacity) to the amount of cane that can be supplied from different geographical districts to the mill in any time interval.

Research towards improving each of the above individual growing, harvesting, transport and milling components has historically provided significant enhancements to the sugar industry. However gains from this component based research has slowed, and this has increased the emphasis on the need for whole of system research for the Australian sugar industry [Muchow et al., 2000]. In light of this, a major opportunity exists to enhance profitability by harvesting cane at the optimal dates, accounting for the geographical and temporal differences in CCS and cane yield across the mill region at any harvest date. Achieving this has required consideration and integration of the growing, harvesting, transport and milling sectors which are all subject to capacities, costs of change and risk. In light of these considerations, the objective of this paper is to highlight an integrated modelling approach to optimise the harvest date of sugarcane for enhanced profitability.

2. INTEGRATED MODEL DEVELOPMENT

2.1 The Optimisation Model

The underlying model to optimise the harvest date of farm paddocks for enhanced industry profitability is a large-scale constrained integer program subject to capacity constraints. While the complete details of the integer program has been reported by Higgins [1999], the basic description needed for this paper is as follows: Assuming a one year planning horizon, the key decision variable is:

$$x_j^i = \begin{cases} 1 & \text{if paddock } i \text{ is harvested in time period } j \\ 0 & \text{otherwise} \end{cases}$$

where a time period is either a week or a fortnight during the harvest season, given that the harvest season is up to 26 weeks long. Given that a mill region can contain up to 12000 farm paddocks, the objective is to maximise profitability to a mill region:

$$\text{Max } Z = \sum_i \sum_j SP * s_j^i * c_j^i * x_j^i \text{ - milling, harvesting and growing costs} \quad (1)$$

where:

SP international sugar price per tonne

s_j^i estimated CCS of farm paddock i if harvested in time period j

c_j^i estimated tonnes of cane per hectare if paddock i is harvested in time period j

In terms of constraints on the system:

A paddock is harvested in only one time period ie.

$$\sum_j x_j^i = 1 \quad \forall i \quad (2)$$

Milling capacity limits the total amount of cane that can be harvested in a time period. This capacity varies throughout the harvest season due to cane quality and unexpected mill stoppages. In analytical form:

$$\sum_i c_j^i * x_j^i < U_j \quad (3)$$

where U_j is the estimated milling capacity in time period j

Transport capacity limits the amount of cane that can be shifted along a rail transport branch line or from a road transport loading point in a time period. In analytical form:

$$\sum_{i \in T_a} c_j^i * x_j^i < P_a \quad (4)$$

where T_a is the set of farm paddocks allocated to branch line or loading point a, and P_a is the capacity in a time period.

Harvesting capacity limits the amount of cane that a harvester can cut or is permitted to cut in a time period, ie.

$$\sum_{i \in H_b} c_j^i * x_j^i < G_b \quad (5)$$

where H_b is the farm paddocks cut by harvester b, and G_b is the capacity in a time period.

Constraints in (3) to (5) were developed with expert knowledge from milling staff through a participatory research process [Martin and Sherington, 1997]. The derivation required several revisions in an action learning cycle [McGill and Beaty, 1995]. As a consequence of this approach, the final set of constraints was approved by the industry partners as being fully credible from a practical perspective.

The resulting model is an extension of the generalised assignment problem and can only be solved optimally for very small size problems (that is, less than about 500 paddocks). In light of this, a heuristic solution technique was developed to find a near optimal solution but within a CPU time of about 12 hours on a Pentium III PC. The heuristic adopted [Higgins, 2001] was an extension of tabu search [Glover, 1993], which is based on neighbourhood search.

Three different cane supply options were agreed to by industry for implementation in the 2000 and 2001 harvest seasons. Each of these options involve optimising the harvest date of farm paddocks, but at different levels of farm and harvesting group equity:

Option 1: The amount of cane that a harvesting group cuts in any one day is only restricted by road/rail transport and milling capacities. If a harvesting group contains farms with a CCS rising relative to the mill average, then an optimal harvest schedule will suggest cutting more cane per day during the later part of the harvest season, thus increasing profitability for those farms.

Option 2: Within a harvesting group, there are no limits on the amount of cane that can be harvested from a farm at any time period of the harvest season. However, unlike Option 1, harvesting equity requirements for the group must be maintained throughout the season.

Option 3: The harvest dates of the paddocks on a farm are optimised with regards to variety and crop class without affecting the equity of other farms within the harvesting group.

The optimal harvest dates for all farm paddocks in the case study mill regions of Mackay and Mossman were simulated, with the objective of maximising industry profitability, for each of these three cane supply options.

Through participatory research with industry, formulating and solving the integer programming model defined by (1) to (5) was achieved to a satisfactory degree of credibility from both a scientific and industry perspective. This is because it only required a process to overcome the logistical complexities involved. However, solutions to the optimisation model also rely on estimates of average paddock CCS, s'_j, and cane yield per hectare, c'_j. Accuracy of the simulated model solutions, in terms of the ability for growers and the industry to achieve the profitability benefits when implemented, almost entirely depended on how well s'_j and c'_j are estimated. This estimation is complicated by year to year climate variability, spatial variation, and the need to rely on historical productivity data containing potential errors and unquantifiable variation. Therefore, a considerable amount of emphasis has been placed on the development of statistical models to more accurately estimate trends in CCS. For the purpose of the implementation of pilot studies in 2001, pre-season cane yield estimates supplied by all growers to the mill were used to estimate c'_j.

2.2 Estimation of Relative CCS by Harvest Date

The essential model parameters of s'_j and c'_j have been traditionally estimated through field experimentation [see Rostron, 1972; and McDonald and Wood, 2001]. However, this method is limited in that it is unable to account for a wide range of geographical and year to year effects. Instead, s'_j was estimated using historical block productivity data [Chardon and Smith, 1993] for all farm paddocks within the mill region. Block productivity data has been maintained by most mills in Australia since the early 1990's and contains information for each farm paddock which includes year of harvest, area of farm paddock, geographical district, harvest date, cane variety, crop class and yield attributes.

Growers are traditionally paid for the sugar that they supply to the mill through a relative payment scheme. The return to the grower is computed on a daily basis and is greatest when the average CCS for the farm (weighted by tonnes of cane supplied to the mill) relative to the average CCS for all cane supplied to the mill (farm relative CCS) is at its highest point during the season. The potential benefit to a farm of participating in the cane supply arrangements presented as Options 1 and 2, will largely depend on how its average CCS relative to the mill varies throughout the season. Therefore, the simulation of optimal harvest dates requires the estimation of farm relative CCS at each harvest date to reflect trends that are consistent for most seasons. For a mill region, the trend in average CCS with harvest date will typically vary from season to season, depending largely on the climatic conditions during the growing and harvesting seasons. Figure 1 shows that, for the Mossman sugarcane region, the mill CCS trend in 1998 took a downward turn at around the 11th week of harvest while the mill CCS trend in 1999 continued to rise beyond the 20th week of harvest. Due to the difficulties associated with forecasting climate patterns and, hence, CCS trends for an approaching season, we have focussed instead on deriving a method for estimating farm CCS relative to the mill, which may be less variable from season to season than actual CCS. For the purpose of optimising the harvest dates of farms across a mill region, it suffices to generate estimates of CCS relative to the mill.

To capture the variation in farm relative CCS trends over a mill region, the approach taken was to estimate a linear trend in relative CCS with harvest week using data over all seasons, for each farm. As the variety and crop class composition of crops growing on a farm may change considerably from one season to the next it was also appropriate to remove average variety and crop class effects before combining the seasonal relative CCS data. The farms were then assigned to groups pre-defined on the basis of the size of the slope and intercept attributes of the linear model for farm relative CCS. These groups capture some of the variation in relative CCS with harvest week that, is associated with the confounded effects of geographical location and farm management procedures.

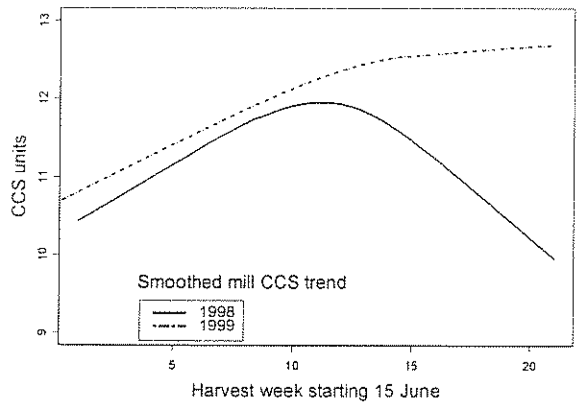


Figure 1. Smooth CCS trends for sugar cane supplied by all farms to the Mossman mill during the 1998 and 1999 harvest seasons.

The adjusted relative CCS for each farm paddock, rc^i_{jy}, was computed separately for each season, y $(y=1,\ldots,n_m)$, where n_m is the number of seasons represented in the block productivity data for mill region m ($n_m > 5$). A quartic polynomial model was fitted (using least squares regression) to paddock CCS, weighted by the tonnes of cane harvested, separately for each season so that the residuals, e^i_{jy}, represent paddock relative CCS computed using a smooth mill CCS trend that was relevant to the season (see Figure 1). Average variety and crop class were also removed from the e^i_{jy}, so that the rc^i_{jy} were computed from the analysis of variance model

$$e^i_{jy} = \mu + v_y + cl_y + (v.cl)_y + rc^i_{jy} \qquad (6)$$

where the rc^i_{jy} are the residuals of the model fitted to the e^i_{jy} for paddock i harvested in week j, with effects for variety (v_y), crop class (cl_y), their interactions during season y, and overall mean μ.

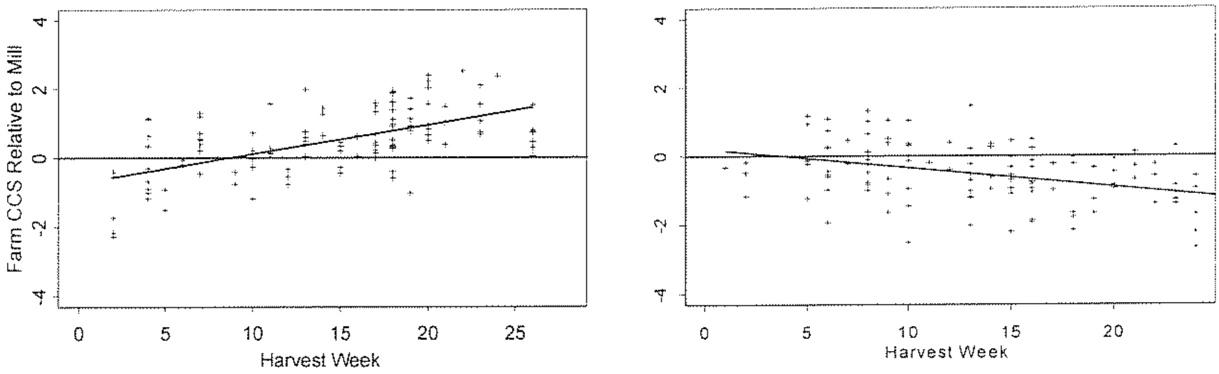
The linear model fitted to the adjusted relative CCS for all paddocks on a farm f is

$$rc^i_{jy} = \alpha_f + \beta_f * j + \eta^i_{jy} \qquad (7)$$

where $i=(1,\ldots n_f)$ and n_f is the number of paddocks belonging to farm f, $j=(1,\ldots,w)$ is the harvest week where w is the length of the harvest season in weeks, a_f and b_f are the estimated coefficients for the intercept, α_f, and slope (or gradient), β_f, of the linear model and the η^i_{jy} are the residuals of the model fit and are assumed to have a N(0,1) distribution. Figure 2a and 2b shows graphs for two farms with positive and negative trends in farm relative CCS with harvest date, respectively, in the Mackay sugar region.

Farm relative CCS with harvest week was modelled as a simple linear trend for the purpose of detecting whether average relative CCS for a farm is consistently higher either early or late in the season. A more complex trend was not considered, due to the large amount of variation in the relative CCS data for paddocks. A farm f was considered to be a potential candidate for Options 1 and 2 if $\beta_f \leq -0.02$ or $\beta_f \geq 0.02$. These values correspond to a change of at least 0.4 units in CCS over a period of 20 weeks. This was regarded by our industry partners to be the minimum gain in relative CCS that a grower would consider to be worthwhile pursuing. Farms were grouped according to the criteria that ($b_f \leq -0.02$ or $b_f \geq 0.02$) and ($a_f < 0$ or $a_f > 0$) with a 10% level of significance. This resulted in the formation of nine groups of farms. For the three groups in which b_f was not significantly different from zero, there was no evidence that the farm average CCS followed a different trend in CCS to the mill. Hence, the grower would not be encouraged to participate in the cane supply arrangements of Options 1 and 2.

The significance of trends in farm relative CCS with harvest week, and the model's ability to estimate farm CCS accurately, varied considerably for a number of reasons. For example, large time intervals between visits by the harvester to a smaller farm may result in a sparsity of CCS entries for every harvest week in the block productivity data, even over six or more seasons. If the timing of the visits is not consistent over several seasons or if a strong linear trend was not present in the majority of seasons for which data is available, then the evidence of a linear trend with harvest week may be weak. Also, there may be variations in the data due to unknown sources that have not yet been accounted for in the model which, consequently, distort any trend with harvest week. For these reasons it was acknowledged that some farms will have been incorrectly allocated to groups, and hence estimation for these farms will not be as accurate as for others. However, for the purposes of encouraging participation in a pilot study, a conservative approach was taken in which farms were not allocated to a "high early" or "high late" group unless the evidence was reasonably strong.



a) Increasing trend with harvest week for farm X   b) Decreasing trend with harvest week for farm Y

Figure 2. Graphs showing trends in farm relative CCS with harvest week, derived from paddock (block) productivity data recorded over 6 seasons, for two farms in the Mackay sugar region.

Due to the variation in farm size and the unbalanced structure of data by variety, crop class and harvest week for a single farm, a linear model as represented by (7) was fitted to the relative CCS data for all farms within a group to provide an estimate of the average linear trend in farm relative CCS with harvest week for the group. The relative CCS for a paddock is then estimated by combining the average effects of variety and crop class with the linear trend by harvest week effect contributed by the farm (confounding location and management effects).

For application in the 2001 harvest season, the statistical model developed for farm relative CCS with harvest week has been able to capture some of the variation in the data contributable to variety, crop class, farm and the trends with harvest week. This has enabled the optimisation model to generate cane supply by harvest date solutions which are consistent with a farm's ability to benefit by 1) harvesting a higher percentage of cane either earlier or later during the season (Options 1 and 2), and 2) harvesting according to a schedule which optimises the order of paddocks by variety and crop class with restrictions due to crop age (Option 3). The farm grouping structure, along with the graphs depicted in Figure 2, has provided a valuable tool for assessing whether a grower may be likely to benefit from participating in an Option 1 or 2 cane supply arrangement and, hence, encouraging growers to participate in the pilot studies.

3. APPLICATION TO CASE STUDY SUGARCANE REGIONS

As part of the participatory research process, the integrated model was implemented in the Mackay and Mossman mill regions, which are located on the north-east coast of Australia. The Mackay region has four mills and about 1200 growers supply cane to these mills and cover an area of about 85000 hectares. Mossman has one mill and about 250 growers covering an area of about 13000 hectares.

Implementation took place for at least two of the three options during the 2000 and 2001 harvest seasons. The whole process of implementation was a challenge as:

- Individual farms are privately owned and growers have for a long time been governed by regulated equity.
- Profitability gains for the growers can mean losses for the milling sector and vice versa.
- There are difficult to quantify operational constraints associated with harvesting, cane transport (road and rail) and milling.
- Socioeconomic barriers for change from the current system in harvest management are strong in the Australian sugar industry, and acceptance of change varies significantly across farmers and harvesting groups.

Overcoming these potential barriers for implementation required the following process linked with an action learning cycle:

- Creative solutions workshops with industry to plan paths to best overcome key barriers for implementation. One outcome from the workshops was that implementation was to be evolutionary rather than revolutionary.
- Establishment of a process to recruit growers and harvesting groups for pilot studies.
- Development of decision support tools to allow optimal harvest schedules developed by the

model to be integrated with expert knowledge from the local grower.

- Development of a process to provide adequate support for the participants during the harvest season and to prevent deviations from the model generated schedules due to unforeseen difficulties
- Evaluation of the pilot studies for actual profitability gains to the growers as well as the findings to enhance the process for 2001.

As a result of this pilot implementation process, 93 growers piloted one of the three options (72 for Option 1, 3 for Option 2, and 16 for Option 3) in the 2000 harvest season. Unfortunately, the harvest season suffered very adverse conditions in wet weather, poor cane growth, disease and rat damage. This meant that most growers only partially followed the optimal harvest schedules. However, for those growers that were able to follow the harvest schedules closely, there were indications of some increase in profitability relative to previous years and control farms. For full details on the pilot study process, refer to Higgins et al. [2000]. Nearly all growers who participated in the 2000 harvest expressed a willingness to continue in 2001.

4. CONCLUSIONS

Sugar industries around the world are a complex integrated system of the growing, harvesting, transport, milling and marketing sectors. A major issue within the Australian sugar industry is the need to take advantage of geographical, temporal, and crop characteristic differences in sugar yields, while accounting for the complex integrated system. In addressing this issue, the key contribution of this paper has been an overview of an integrated optimisation and statistical modelling approach to provide alternative harvest schedules for enhanced industry profitability. The integrated modelling approach was developed through a participatory research approach with the sugar industry and required an action learning cycle to ensure credibility for the growing and milling sectors within the industry.

Through participatory research with five case study mill regions, the integrated modelling approach has incorporated local industry constraints, costs, and CCS and cane yield inputs. Application to the mill regions showed significant average potential profitability gains from optimising the harvest date of sugarcane with consideration of geographical differences in CCS and cane yield. Through a process to overcome barriers for implementation, 93 growers piloted optimal harvest schedules in the 2000 harvest season. While most growers were not able to follow the schedules throughout the whole harvest season, due to very adverse climatic conditions, there were some significant gains in profitability for those who did follow the schedules closely. As a result of evaluating the pilots in 2000, several issues have been incorporated into the modelling and implementation process for 2001. Such issues include the way in which constraints are modelled and the level of support required during the harvest season.

5. REFERENCES

Chardon, C.W., and M.A. Smith, Improved extension material from block recording schemes, *Proceedings of the Australian Society of Sugar Cane Technologists, 15*, 60 – 65, 1993.

Glover, F., A user's guide to tabu search, *Annals of Operations Research, 41*, 3 – 28, 1993.

Higgins, A. J., Optimising cane supply decisions within a sugar mill region, *Journal of Scheduling, 2*, 229 – 244, 1999.

Higgins, A. J., A dynamic tabu search for large scale generalised assignment problems, *Computers and Operations Research, 28*, 1039-1048, 2001.

Higgins, A. J., M. A. Haynes, and R. C. Muchow, Pilot study evaluation of alternative cane supply options in Mackay sugar region, Internal report for the CRC for Sustainable Sugar Production, James Cook University, Townsville, 2000.

Martin, A., and J. Sherington, Participatory research: Methods—Implementation, effectiveness and institutional context, *Agricultural Systems, 55*, 195-216, 1997.

McDonald, L. and A. W. Wood, Pre-harvest season productivity of sugarcane in different growing regions of Queensland, *Proceedings of the Australian Society of Sugar Cane Technologists, 23*, 90-96, 2001.

McGill, I., and L. Beaty, Action Learning: A guide for professional, management, and educational development, Kogan Page, London, UK, 1995.

Muchow, R.C., A.J. Higgins, W.J. Andrew, and M.A. Haynes, Towards improved harvest management using a whole industry systems approach, *Proceedings of the Australian Society of Sugar Cane Technologists, 22*, 30 – 37, 2000.

Rostron, H., The effects of age and time of harvest on the productivity of irrigated sugar cane, *Proceedings of the South African Sugar Technologists Association*, 142 – 150, 1972.