

**BUREAU OF SUGAR EXPERIMENT STATIONS
QUEENSLAND, AUSTRALIA**

**FINAL REPORT
SRDC PROJECT BS119S
BEST LINEAR UNBIASED PREDICTION
AS A METHOD FOR PREDICTING
CROSS POTENTIAL**

by

J K Stringer

SD99017

Principal Investigator: Miss J K Stringer
Research Officer
BSES
PO Box 86
Indooroopilly Qld 4068

Phone: (07) 3331 3337

This project was funded by the Sugar Research and Development Corporation during the 1994/95, 1995/96, 1996/97 financial years.

**BSES Publication
SRDC Final Report SD99017**

September 1999

CONTENTS

	Page No.
ABSTRACT	
1.0 INTRODUCTION.....	1
2.0 OBJECTIVES	2
3.0 MATERIALS AND METHODS	2
3.1 Empirical Method of Estimating Breeding Value.....	2
3.1.1 Agronomic performance.....	2
3.1.2 Disease resistance status	2
3.1.3 Breeding performance	3
3.1.4 Inbreeding.....	3
3.2 BLUP analyses	3
3.3 Crossing in 1994 and 1995	4
3.3.1 BLUP and empirical crosses.....	4
3.3.2 Inbred crosses	4
3.3.3 Problems with crossing	5
3.4 TRIAL DETAILS.....	5
3.5 STATISTICAL ANALYSES.....	5
4.0 RESULTS AND DISCUSSION.....	5
4.1 Comparison of BLUP and empirical crosses	5
4.2 Inbred crosses	7
5.0 APPLICATION OF RESULTS TO THE INDUSTRY.....	8
6.0 CHANGES TO PROJECT OBJECTIVES	8
7.0 PUBLICATIONS	8
8.0 REFERENCES	8
9.0 ACKNOWLEDGMENTS	9

ABSTRACT

Until recently, BSES breeders used an empirical formula to assess the breeding potential of a parent (Hogarth and Skinner, 1986). Although this method uses data from many sources, it needs to draw on data accumulated for up to ten years to provide reliable estimates. An alternative is the statistical technique called Best Linear Unbiased Prediction (BLUP) which was developed to predict breeding values from dairy cattle data sets which are typically highly unbalanced and from diverse sources.

The main objective of this study was to compare the relative efficiencies of BLUP and the existing empirical method for identifying superior sugarcane cross combinations. There was little difference in the agronomic performance of crosses selected using either BLUP or the empirical method but BLUP proved to be the quicker method and should allow more rapid progress from breeding and selection.

1.0 INTRODUCTION

Up until the late 1990s, BSES breeders used an empirical formula to assess the breeding value of a parent. This method combines information on agronomic and breeding performance as well as disease ratings into an index, but needs to draw on data accumulated over ten years to provide reliable estimates. Sugarcane breeders were not satisfied with this method but found it difficult to make improvements.

Directly related to the problem of estimating breeding value is the design of the breeding system. The current system of breeding sugarcane has a very unbalanced mating design caused by sparse and unreliable flowering and poor pollen fertility at the main breeding station at Meringa in North Queensland (Berding and Skinner, 1987). This, coupled with the fact that up until the late 1980s, families were only visually assessed, made the application of statistical procedures inappropriate.

Each year, about 300 new clones enter the sugarcane parental collection. As crossing is expensive, only a limited number of crosses can be made and this makes it desirable to have a statistical method that could be applied to early stage families. This would allow superior parent clones to be identified more rapidly resulting in an increase in the rate of population improvement.

Sugarcane early stage family data are unbalanced because families are not equally represented across all years, locations and blocks. The classical approach to analysing such 'messy' data is to treat families as fixed effects to be estimated from the data. This results in ordinary or weighted least squares solutions or a procedure that approximates least squares (White *et al*, 1986). An alternative approach is to treat the breeding values as random effects to be predicted rather than fixed effects to be estimated. BLUP is an analytical technique that treats breeding values as random effects. It was developed specifically for use within the dairy industry to predict breeding values of sires and the performance of future offspring (Henderson, 1975).

When the data are equally represented across blocks and locations and are equally precise for all sites and all parents, the procedures described above give the same genotype rankings (White *et al*, 1986). However, when the data are unbalanced, the fixed effects models can lead to imprecise family effects, and there is a tendency to select parents that have been in

fewer tests (Panter and Allen, 1995a). This contrasts with BLUP technique where a large proportion of the higher ranking parents will be based on higher quality and quantity of data.

In previous research, methodologies were developed to allow BLUP to be applied to BSES family selection data and preliminary results looked promising (Stringer *et al*, 1996). However, it was not possible to be sure that BLUP was as effective as the current BSES method for identifying superior parents without field research. This paper reports the results of this fieldwork and confirms that BLUP is at least as effective as the current BSES method for identifying superior parents.

2.0 OBJECTIVES

- (i) Compare BLUP estimates of breeding value based on multiple traits to the current methods used in the BSES breeding program.
- (ii) Compare performance of crosses with predictions based on BLUP estimates and the current methods.
- (iii) Determine the level of inbreeding in the BSES parent collection and the effect of inbreeding on breeding value and cross performance.

3.0 MATERIALS AND METHODS

3.1 Empirical method of estimating breeding value

The empirical method used by breeders within the Australian sugar industry to estimate the breeding value of parental clones incorporates the following information:

3.1.1 Agronomic performance

Clones are initially planted in the parental collection if they exhibit superior performance in yield trials conducted over a number of years and locations.

Clones are assessed relative to commercial standards and the results are tabulated in a selection index called net merit grade, NMG (Skinner, 1965). This is based on sugar yield, fibre quality and visual appearance grade. The mean NMG of the standards is always 10.

3.1.2 Disease resistance status

Information on the reaction of parental clones to various diseases is also taken into account when assessing breeding value. Disease ratings are assigned to a clone on a 1-9 scale where 1 is highly resistant and 9 is highly susceptible. In the empirical formula, disease ratings are in the form of an adjustment factor which is designed to equal zero if a clone has a rating of five for all diseases (Hogarth and Skinner, 1986). The form of the adjustment depends on the region for which the parental clone will be used; only important diseases in a particular area are considered.

3.1.3 Breeding performance

Many experimental crosses are evaluated each year in six selection programs. The aim is to identify those crosses with specific combining ability coupled with the high general

combining ability of the parent clones. These crosses are called *proven*, and more seedlings from such crosses are planted in subsequent years.

The system for identifying proven crosses is based on:

1. *Selection rate* - the percentage of original seedlings that are selected and replanted in more advanced stages of testing (Hogarth and Skinner, 1986).
2. *Family selection* - whole families are rejected or selected based on mean performance (Falconer and Mackay, 1996).

3.1.4 Inbreeding

Before a cross is made, the level of inbreeding is determined. Depending on the level of inbreeding, a proposed cross may not be made or the cross ratio value may be penalised. Four levels of inbreeding are recognised and these are given below together with their inbreeding coefficients (F), assuming $F = 0$ in the base population:

1. *Selfs* - Female and male clones are identical, $F = \frac{1}{2}$
2. *Line breeding or parent/offspring* - Female clone is same as either parent of male clone or vice versa, $F = \frac{1}{4}$.
3. *Half sibs* - Female and male clones have one parent in common, $F = \frac{1}{8}$
4. *Full sibs* - Female and male parents of both clones are identical, $F = \frac{1}{4}$.

Crosses are rejected for selfs and full sibs and, if line breeding is involved. In the case of half sibs, a value of one is subtracted from the cross ratio value.

3.2 BLUP analyses

Before the start of the 1994 crossing season, BLUP estimates for parental clones from the Central, Southern and Burdekin selection programs were obtained. NMG plant crop (first harvest) data from the 1988 –1992 series seedling families were used for analysis. Each family plot contained 15-20 seedlings planted in 12 m plots with an intra-row plant spacing of about 60 cm and an inter-row spacing of approximately 1.5 m. There were at least two and usually four replicates of each family.

For each selection program, all accumulated family NMG data were analysed assuming the linear model for an incomplete diallel. BLUP estimates of parental general combining ability and specific combining abilities for crosses were obtained using the statistical program, GAREML (Huber, 1993). To assist in the inversion of matrices in the BLUP procedure, the year effect was removed before analysis by standardising the data.

After analysis, the BLUP breeding values were converted to a 0-9 scale to mirror the procedure for the empirical breeding values.

3.3 Crossing in 1994 and 1995

3.3.1 BLUP and empirical crosses

Assuming both BLUP and empirical breeding values were available for both parents in a cross, the classification of a cross was based on which average breeding value was the higher. Crosses from several categories were made:

1. **BLUPs**
 - Both parents had BLUP information. The best possible cross combinations were made.
2. **BLUP random**
 - Both parents had BLUP information. Crosses were made randomly and were rejected only on the basis of high disease status or high levels of inbreeding.
3. **Empirical**
 - Both parents had breeding values calculated using the empirical formula. The best possible cross combinations were made.
4. **Empirical random**
 - Both parents had breeding values based on the empirical formula. Crosses were made randomly and were rejected only on the basis of high disease status or high levels of inbreeding.

The same rules applied when making the BLUP, BLUP random, empirical and empirical random crosses. Not all types of crosses were made for all three selection programs.

3.3.2 Inbred crosses

Crosses with various levels of inbreeding were made to study the effects of inbreeding on cross performance. The aim was to make at least one self, two full sibs, two half sibs and two unrelated crosses from three or four chosen parents. Clones chosen were 79A147 and 80N314 (which are recognised as good male parents) and 75C35 (which can be used as either a female or a male). A few crosses were also made with 81S2354, which is a male.

A total of 52 crosses (including standard crosses) was made and these can be categorised according to their inbreeding coefficients (F) as

- | | |
|-----------------------|--------------------------------|
| • 4 selfs | F = 0.5 |
| • 9 full sibs | F = 0.25 |
| • 12 half sibs | F = 0.125 |
| • 24 crosses | F ranged from 0.0156 to 0.0313 |
| • 3 unrelated crosses | F = 0 |

Pedigrees were traced to great grandparents wherever possible.

3.3.3 Problems with crossing

Only about half the crosses made in the 1994 crossing season germinated due to lack of temperature control in the germination chambers and problems with the potting mix. These problems were rectified before the 1995 crossing season.

3.4 Trial details

After germination of seed from the 1994 and 1995 crossing seasons, the seedlings were planted into separate trials at the Central, Burdekin and Southern Experiment Stations in 1995 and 1996. The inbred crosses were planted only at the Central Experiment Station.

Depending on the selection program, each single-row family plot contained 15-20 seedlings planted in 12 m plots with each plant spaced approximately 60 cm apart and an inter-row spacing of 1.5 m. The number of replications of each family varied according to the number of seedlings available. However, there were at least two and usually four replicates of each family.

During harvest in 1996 and 1997, an eight-stalk sample (one from eight randomly chosen clones per family plot) was taken for commercial cane sugar (ccs) determination. The seedlings were also weighed as family units and tonnes cane per hectare, ccs, tonnes sugar per hectare and NMG for each family were calculated.

3.5 Statistical analyses

Analyses of variance were undertaken on the data from each selection program in each year. LSD values were calculated to examine the differences in means.

SAS Proc Inbred (SAS Institute Inc 1996) was used to calculate the inbreeding coefficients. As data were limited, it was desirable to combine data from the 1995 and 1996 harvest seasons. Six standard crosses were common to both years and this was used to standardise the data to remove year effects. After standardisation, the data were combined.

As the inbreeding coefficients were not continuous, and hence, not normally distributed, a Spearman's Rank correlation was used to examine the relationships in the data.

4.0 RESULTS AND DISCUSSION

4.1 Comparison of BLUP and empirical crosses

One of the major advantages of BLUP is its ability to obtain more precise estimates of breeding values than traditional least squares approaches, when not all genotypes are equally represented across all years, locations and blocks (Henderson, 1975; Hill and Rosenberger, 1985). However, in the current research we found no significant differences for NMG between the BLUP and empirical crosses except for the Burdekin selection program in the 1995 harvest season (Table 1). It was difficult for sugarcane breeders to select the lists of parents independently, and this may have caused the true potential of BLUP to be underestimated.

TABLE 1

Mean net merit grades for BLUP, BLUP random, empirical and empirical random sugarcane crosses planted at the Burdekin, Central and Southern Experiment Stations in 1995 and 1996.

Year of planting	Type of cross	Mean NMG
Burdekin		
1995	Blup	11.8 a
	Blup random	9.4 bc

	Empirical	10.3 c
1996	Blup	7.7 a
	Blup random	8.4 a
	Empirical	8.1 a
	Empirical random	7.2 a
Central		
1995	Blup	9.5 a
	Blup random	8.6 bc
	Empirical	9.2 ac
1996	Blup	10.4 a
	Blup random	8.8 bc
	Empirical	10.4 a
	Empirical random	9.5 ac
Southern		
1995	Blup	10.0 a
	Blup random	10.3 a
	Empirical	9.2 ac
	Empirical random	8.5 bc
1996	Blup	10.3 a
	Blup random	8.7 b
	Empirical	9.7 a

Note: Means followed by the same letter are not significantly different.

Although there was no significant difference in the agronomic performance of crosses based on the two methods except for the 95 series trial in the Burdekin, BLUP breeding values were obtained using fewer years of data (maximum of six years) in comparison with the empirical method (ten years). The economic savings associated with using BLUP and a reduced generation interval were the major considerations that resulted in changes to the BSES crossing program. In 1998, BLUP was implemented as the standard technique to select parents in the BSES core breeding program except for new experimental varieties for which there is little agronomic and breeding information.

The power of BLUP as a predictive tool in sugarcane reinforces the decision by BSES to use BLUP as the technique to select parents for the breeding program (Stringer *et al* 1996). Using family selection data from Mackay and Bundaberg, predicted performance of BLUP crosses was correlated with actual trial performance over two years. The correlation coefficients for BLUP vs NMG was always greater than for the empirical method vs NMG.

As the BLUP method of estimating parental breeding values within the Australian sugar industry becomes more refined, its superiority will increase. Software limitations in the current research restricted the complexity of the BLUP models that were considered. Future research within the Australian industry will consider the genetic variance/covariance structure among individuals. This would allow related individuals to

contribute to the predicted values of one another which may enhance the precision of predictions (Panter and Allen, 1995ab).

4.2 Inbred crosses

Rank correlations between the level of inbreeding and harvest variables were negative and highly significant in all cases except for ccs (Table 2).

The inverse relationship between the level of inbreeding and NMG justifies the current strategy that is adopted within the BSES crossing system. Crosses are penalised crosses depending on the severity of inbreeding. Crosses that are selfs, full sibs or involve line breeding are rejected. In the case of half sibs, a value of one is subtracted from the cross ratio value.

TABLE 2

Spearman rank correlation between the level of inbreeding and harvest variables

Comparison	Inbreeding coefficient
Tonnes cane per hectare vs inbreeding coefficient	-0.40
Ccs vs inbreeding coefficient	-0.16
Tonnes sugar per hectare vs inbreeding coefficient	-0.45
NMG vs inbreeding coefficient	-0.45

5.0 APPLICATION OF RESULTS TO THE INDUSTRY

Given the positive results emanating from this project, a decision was made in April 1998 to incorporate BLUP as the standard technique for selecting clones for experimental crosses in the BSES core breeding program. However, when no BLUP information is available, parental choice will be based on the empirical formula.

The field trials in this research were conducted in three selection programs that service approximately 70% of the Australian sugar industry. The earlier recognition of superior parents by the BLUP method will result in increased genetic gains and improved varieties for the Australian sugar industry.

6.0 CHANGES TO PROJECT OBJECTIVES

CSR was initially involved in the project. However, CSR withdrew from the project early on since the crosses they were making from their existing methods were similar to those predicted from BLUP. This may be due to CSR having a much smaller and simpler breeding program than BSES. As they do not combine data from as many sources as BSES the potential benefit of BLUP would be less.

7.0 PUBLICATIONS

Stringer, J.K., McRae, T.A. and Cox, M.C.(1996). Best linear unbiased prediction as a method of estimating breeding value in sugarcane. In Sugarcane: Research towards efficient and sustainable production. J.R. Wilson, D.M. Hogarth, J.A. Campbell and A.L. Garside, eds., CSIRO Division of Tropical Crops and Pastures, Brisbane.

8.0 REFERENCES

- Berding, N. and Skinner, J.C.(1987). Traditional breeding methods. In Copersucar International Sugarcane Workshop. Copersucar, Piracicaba, Brazil.
- Falconer, D.S. and Mackay, T.F.C. (1996). Introduction to quantitative genetics, 4th ed. Addison Wesley Longman Limited, Essex.
- Henderson, C.R. (1975). Best linear unbiased estimation and prediction under a selection model. *Biometrics* 31:423-447.
- Hill, R.R. and Rosenberger, J.L. (1985). Methods for combining data from germplasm trials. *Crop Sci.* 25:467-470.
- Hogarth, D.M. and Skinner, J.C. (1986). Computerisation of cane breeding records. *ISSCT* 19:478-491.
- Huber, D. (1993). Optimal mating designs and optimal techniques for analysis of quantitative traits in forest genetics [Ph.D. Dissertation]. University of Florida, Gainesville, Florida.

- Panter, D.M. and Allen, F.L. (1995a). Using Best Linear Unbiased Predictions to Enhance Breeding for Yield in Soybean: I Choosing Parents. *Crop Sci.* 35:397-405.
- Panter, D.M. and Allen, F.L. (1995b). Using Best Linear Unbiased Predictions to Enhance Breeding for Yield in Soybean: II Selection of Superior Crosses from a Limited Number of Yield Trials. *Crop Sci.* 35:405-410.
- SAS Institute Inc (1996). SAS/STAT Software: Changes and Enhancements through Release 6.11. SAS Institute, Cary, North Carolina.
- Skinner, J.C. (1965). Grading varieties for selection. *ISSCT* 12:938-949.
- Stringer, J.K., McRae, T.A. and Cox, M.C.(1996). Best linear unbiased prediction as a method of estimating breeding value in sugarcane. In Sugarcane: Research towards efficient and sustainable production. J.R. Wilson, D.M. Hogarth, J.A. Campbell and A.L. Garside, eds., CSIRO Division of Tropical Crops and Pastures, Brisbane.
- White, T.L., Hodge, G.R. and Delorenzo, M.A.(1986). Best linear unbiased prediction of breeding values in forest tree improvement. In Workshop of the genetics and breeding of southern forest trees. Southern Region Information Exchange Group 40 Gainesville, Florida 25-26 June 1986. University of Florida, Gainesville.

9.0 ACKNOWLEDGMENTS

The funding support from the Sugar Research and Development Corporation for this project is gratefully acknowledged together with the contribution from technical staff at the Central, Burdekin and Southern Experiment Stations. BSES plant breeders Drs Mike Cox and Tony McRae, and Group Manager, Dr Mac Hogarth, were responsible for selection of experimental crosses and technical guidance in this project.