A COMPARISON OF TWO STATISTICAL TECHNIQUES FOR ASSESSING YIELD DIFFERENCES BETWEEN EARLY GENERATION CANOLA SELECTIONS

N. Wratten, B. R. Cullis

Agricultural Research Institute, Private Mail Bag, Wagga Wagga, Australia, 2650

INTRODUCTION

The New South Wales Agriculture & Fisheries canola breeding programme at Wagga Wagga has used spatial analysis procedures for analysing yield data from variety trials since 1984. Between 1988 and 1990, five Brassica napus cultivars have been released with a cumulative yield advantage over Marnoo, the previous highest yielding Australian cultivar, of 31%. These cultivars are Maluka and Shiralee (released in 1988), Eureka (1989), and Barossa and Yickadee (1990).

At an early stage of the programme seed supply is limited and only one or two plots of each test line are sown in early generation yield trials. The test plots are generally interspersed with check plots, sown to standard cultivars and spaced at a regular frequency along each of several rows. A large number of test lines are usually involved, requiring large field experiments. In these experiments, it is important that an effective method of local yield control be applied to the yields of test plots to maintain reasonable genetic progress.

For replicated field trials, there has been considerable statistical research into efficient designs and analyses. In a study of 1019 variety trials (Cullis and Gleeson 1989) of a range of crops grown over three years in four states of Australia, the use of spatial analysis of field experiments (SAFE; Gleeson and Cullis 1987) resulted in an average reduction of 42% in variances of varietal yield differences compared to randomised complete block analysis. Of the 1019 trials, 219 were incomplete block designs and the average reduction in variances of varietal yield differences with incomplete block analysis and recovery of interblock information was 33%.

There were 59 canola trials involved in this study and use of SAFE resulted in a 21% reduction in variances of varietal yield differences relative to randomised complete block analysis.

In an extension of the SAFE procedure (Gleeson and Cullis 1987), a spatial method was developed for the analysis of early generation variety trials (EGVTs; Cullis et al. 1989). Their method involves residual maximum likelihood (REML) estimation of the variance parameters relating to genotype, trend and error, after which the best linear unbiased predictors (BLUPs) of the test line genotype effects are calculated. This procedure was

originally developed for EGVTs in which there is only one plot of each test line, thus obviating the use of SAFE. The main difference between SAFE and SAFEGVT is the assumption of fixed (SAFE) and random (SAFEGVT) test line effects. The assumption of randomness of test line effects is clearly valid for EGVTs in which the population of test lines are generally unselected for yield, and should lead to more accurate prediction of test line genotype effects.

This paper uses data from seven canola EGVTs, in which test lines were replicated twice, to provide a practical comparison between the SAFE and SAFEGVT procedures.

MATERIALS AND METHODS

A series of seven early generation yield trials was established in 1989 at the Agricultural Research Institute, Wagga Wagga. These involved a total of 483 F_3 single plant selections from a range of crosses and used Barossa and Yickadee as standard cultivars in check plots. Initial selection was based on resistance to the basal stem canker phase of blackleg (caused by the fungus Leptosphaeria maculans) and canola quality.

Each test plot was replicated twice with alternating standard cultivars every five plots. Each plot consisted of three rows (four rows were sown but one outside row was removed due to poor emergence), 6.0m long and a row spacing of 18 cm. Plots were separated by a 50cm gap. Trials were arranged in four rows each with either 33 or 41 columns. Establishment was recorded in all plots and used as a covariate in subsequent analyses.

Each dataset was subjected to three analyses. These were a complete row analysis (ROW), SAFE and SAFEGVT. It was not possible to use a randomised block analysis as there were an unequal number of plots of test lines and check cultivars. Both SAFE and SAFEGVT analysis were performed by removal of trend within rows and so row analysis was the most appropriate baseline analysis in which trend is assumed constant within each row. The efficiency of SAFE or SAFEGVT was calculated as the average variance of pairwise test line differences from a row analysis divided by the average variance of pairwise test line differences from a SAFE or SAFEGVT analysis. Each analysis was also compared by calculating the correlation coefficient between each set of means.

RESULTS AND DISCUSSION

Table 1 presents a summary of the row analysis of yield data for the seven EGVTs. These results are typical for EGVTs conducted by the canola breeding programme, but the coefficients of variation are generally lower than those obtained in later stages of replicated trials conducted by cooperators.

The arithmetic mean efficiencies of SAFE and SAFEGVT were 1.75 and 2.09, corresponding to an average 43%

reduction in average pairwise variance from SAFE analysis, compared with a 53% reduction for SAFEGVT. The median efficiencies of SAFE and SAFEGVT analyses were 1.33 and 1.60, respectively. The individual efficiencies are listed in Table 2.

Table 3 presents the correlation coefficients of the test line genotype effects for row analysis and SAFE analysis with those predicted by SAFEGVT analysis. The average correlations were 0.953 and 0.996 for ROW and SAFE analysis, respectively. There was a strong negative relationship between the correlation between ROW and SAFEGVT analysis and the efficiency of SAFEGVT (Fig 1.). That is, given more trend removal (and hence higher efficiency) there will be larger discrepancies between row analysis test line means and predicted test line genotype effects from SAFEGVT analysis. There was very close agreement between SAFE and SAFEGVT analyses, with the lowest correlation being 0.990.

Another interesting feature of the spatial analyses is presented in Table 4, which shows the estimated yield difference between the two check cultivars, Barossa and Yickadee. Overall, the estimated yield differences were comparable for the three analyses. However the most striking and reassuring difference is the stability of the SAFEGVT analysis. These experiments were sown in a similar environment and at similar times and so genotype differences should remain stable across the seven experiments.

The results from this study are based on a limited dataset, in the sense that trials from only one year, 1989, are involved. For various reasons, trials sown in 1990 were not harvested. A second dataset from 1991 EGVTs will be included in a subsequent comparison. However, the value of SAFEGVT in more accurately estimating test line genotype effects is clear. Continued use should see the development of even higher yielding cultivars for the Australian canola industry.

REFERENCES

CULLIS, B.R. and GLEESON, A.C. 1989. Efficiency of neighbour analysis for replicate variety trials in Australia. Journal of Agricultural Science, Cambridge 113:233-239.

CULLIS, B.R., LILL, W.J., FISHER, J.A and READ, B.J. 1989. A new procedure for the analysis of early generation variety trials. Applied Statistics 38:361-375.

GLEESON, A.C. and CULLIS, B.R. 1987. Residual maximum likelihood (REML) estimation of a neighbour model for field experiments. Biometrics 43:277-288

Table 1. Summary of the design, mean yield, coefficients of variations (C.V.) and average standard error of differences (SED) between test lines for row analysis.

Trial	Test Lines	Rows	Columns	Yield (kg/ plot)	C.V. (%)	SED (kg/ plot)
PT1A	60	4	37	1.745	9.8	0.172
PT1B	60	4	37	1.847	14.8	0.273
PT1C	60	4	37	1.751	10.9	0.191
PT1D	60	4	37	1.947	11.2	0.219
PT2A	66	4	41	1.548	12.0	0.186
PT2B	66	4	41	1.387	11.5	0.159
PT2C	66	4	41	1.245	13.4	0.167

Table 2 Efficiency of spatial analyses for seven EGVTs

Trial	SAFE	SAFEGVT
PT1A	1.33	1.58
PT1B	4.09	5.09
PT1C	1.16	1.37
PT1D	1.83	2.13
PT2A	1.40	1.60
PT2B	1.11	1.25
PT2C	1.31	1.60

Table 3. Correlation coefficients of test line genotype means from ROW analysis and SAFE analysis with predicted test line genotype effects from SAFEGVT

Trial	ROW	SAFE
PT1A	0.973	0.993
PT1B	0.819	0.997
PT1C	0.987	0.999
PT1D	0.938	0.994
PT2A	0.975	0.990
PT2B	0.994	0.999
PT2C	0.982	0.997

Table 4 Estimated yield difference (Barossa - Yickadee, kg/plot) for three analyses of seven EGVTs.

Trial	ROW	SAFE	SAFEGVT
PT1A	0.17	0.17	0.21
PT1B	0.32	0.30	0.30
PT1C	0.25	0.25	0.23
PT1D	0.38	0.35	0.35
PT2A	0.20	0.19	0.21
PT2B	0.26	0.24	0.25
PT2C	0.22	0.23	0.24
Mean	0.257	0.247	0.256
s.d.	0.072	0.062	0.052

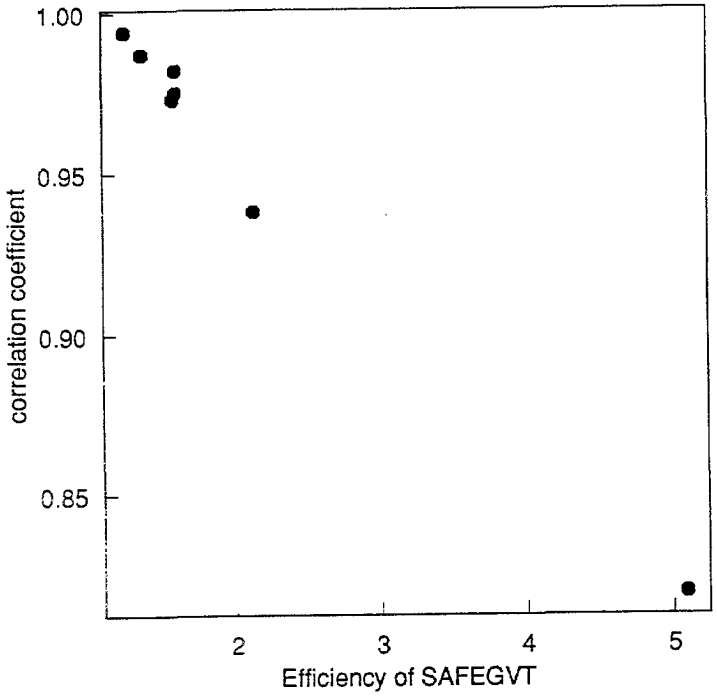

Fig. 1. Relationship between the correlation between Row analysis and SAFEGVT analysis and the efficiency of SAFEGVT