

# Evaluation of Moving Mean and Border Row Mean Covariance Analysis for Error Control in Yield Trials

George C.J. Fernandez<sup>1</sup>

Department of Plant Science, University of Nevada-Reno, Reno, NV 89.5574107

*Additional index words.* experimental error, fertility gradient, check plot design, local trend, micro soil variation

**Abstract.** The effectiveness of using moving mean covariance analysis (MMCA) rather than randomized complete-block design (RCBD) in experimental error control was compared in a large-scale mungbean [*Vigna radiata* (L.) Wilczek] yield trial. The MMCA was superior to the RCBD, since it significantly reduced the experimental error and the coefficient of variation (cv). Inclusion of five neighboring plots in the moving mean computation provided better error control. However, the estimation of optimum number of neighboring plots to be used and moving mean calculations were tedious. The feasibility of using border-row measurements such as mean plant height at 50% flowering or mean seed yield/m of row as a covariate in an analysis of covariance (BRMCA) was examined in a separate mungbean yield trial in which border rows were planted with a check cultivar. Both border-row measurements were equally effective in reducing the experimental error. However, plant height measurements were simpler than measuring seed yield. Because border-row measurements could be readily used as covariate in analysis of covariance without a need for moving mean computation from the response variable, BRMCA could be advantageous for error control in row crops yield evaluation.

Non-uniform fertility gradients caused by nonsystematic spatial soil variation in the experimental sites are impediments to the reliable estimation of treatment and/or cultivar effects (Baker and McKenzie, 1967; LeClerg, 1966; Mak et al., 1978). Experimental errors are usually escalated under these circumstances and high cv values are usually associated with the results.

Various experimental designs were considered for controlling local trend effects in experimental sites. The conventional way is to use RCBD or Latin square design for error control. The RCBD is effective in experimental sites with a uniform fertility trend. Latin square requires an equal number of treatments, rows, and columns, which can restrict the choice of sites. Lattice designs are effective for trials consisting of many treatments; however, the analysis is relatively complicated and is restricted to certain fixed numbers of treatment.

Check plot designs in which a systematic arrangement of check plots is superimposed on a RCBD have been proposed as an alternative method of error control in experiments with many treatments (Baker and McKenzie, 1967; LeClerg, 1966). The performance of a cultivar is expressed as the differences between it and the adjacent check plot or as a percentage of the check plots if a high correlation between check plot and test plot exists. If the check plot misrepresents the soil variation, improper adjustment might result (Mak et al., 1978; Townley - Smith and Hurd, 1973; Rosielle, 1980). The use of check plots was substantially less efficient than using lattice designs or appropriate incomplete designs (Besag and Kempton, 1986). Furthermore, the check plot design requires additional experimental units (Mak et al., 1978).

To prevent such an error in check plot design, MMCA has been recommended (Mak et al., 1978; Townley-Smith and Hurd, 1973; Rosielle, 1980). The mean of the neighboring plots, excluding the test plot in question, is assumed to be a measure of the fertility of the test plot and is used as a covariate to adjust

the treatment effect in an analysis of covariance. If there is no linear association between the test plot yield and the covariate, the regression coefficient of the covariate would be about zero and any overadjustment would be avoided. Although MMCA has been used successfully to adjust treatment effects (Mak et al., 1978; Townley-Smith and Hurd, 1973; Rosielle, 1980), the computation of moving means is cumbersome and the optimum number of neighboring plots included in moving mean calculation varies, depending on the variability of the experiment.

In yield trials of many row crops, plot yield is usually measured by excluding the two border rows and harvesting from the middle rows. If uniform plant types are to be evaluated in yield trials and a check cultivar is planted in the border rows of every experimental unit, then suitable measurements from the border rows could be substituted for the moving mean in the analysis of covariance to adjust the treatment effect for local soil variation. I found no reports on the use of border-row measurements as covariate in error control of large-scale field trials.

The objectives of this study were to investigate the effectiveness of the MMCA in error control and to test the feasibility of using BRMCA when each test plot is bordered by a single check cultivar.

## Materials and Methods

Data used in the MMCA were obtained from mungbean yield trials consisting of 20 advanced breeding lines and a check cultivar conducted at the Asian Vegetable Research and Development Center (AVRDC) in Taiwan during Spring 1987. The experimental design was a RCBD with three replicates. The experimental units consisted of six 6-m rows, spaced 50 cm apart, 10 cm within rows. Two seeds per hill were planted and standard AVRDC mungbean cultural practices (Park, 1978) were employed. Mature pods were harvested manually from the middle four rows in each experimental unit three times, 65, 75, and 90 days after planting. The total seed yield harvested was used for this analysis.

*Statistical analysis.* RCBD analysis was performed on the data. Nonsignificant block and genotypic effects and high cv values indicated the presence of an inflated error term caused by local soil variation within the experimental site. A statisti-

Received for publication 27 Mar. 1989. Journal Paper no. JMS-5131. The cost of publishing this paper was defrayed in part by the payment of page charges. Under postal regulations, this paper therefore must be hereby marked *advertisement* solely to indicate this fact.

<sup>1</sup>Assistant Professor in plant breeding and biometrics.

cally significant trend was observed between residuals from the RCBD analysis and plot location represented by serially numbered plot (Fig. 1). The error term was, therefore, serially correlated or autocorrelated with the adjacent observations (Li and Keller, 1951; Neter et al., 1983). Testing for significance by Durbin-Watson D statistic (Neter et al., 1983) confirmed that the error term was positively autocorrelated (Fig. 1). Seed yields were high in columns C7 and C8, while lower yields were observed in columns C3 and C4 (Fig. 2). Genotypes planted in columns C7 and C8 produced higher yield, probably because they occupied the better plots in the experimental site. The error term was biased because the adjacent plot yields were correlated due to the local soil fertility. For such data, the assumption of

an uncorrelated or independent error term is not appropriate because the error term was positively correlated with the location of the plots.

To overcome the effects of serial correlation on the error term, MMCA has been suggested (Mak et al., 1978; Townley-Smith and Hurd, 1973; Rosielle, 1980). The statistical model of the MMCA is as follows:

$$Y_{ij} = \mu + \beta(X_{ij}) + B_i + G_j + \epsilon_{ij}$$

where:  $Y_{ij}$  = yield of  $j$ th genotype in  $i$ th block;  $\mu$  = grand mean;  $\beta$  = the regression coefficient, measure of linear relationship between the test plot yield and the moving mean covariate;  $X_{ij}$  = moving mean covariate, the mean yield of the neighboring plots of  $j$ th genotype in  $i$ th block;  $B_i$  = the effect of  $i$ th block;  $G_j$  = the effect of  $j$ th genotype; and  $\epsilon_{ij}$  = random error.

For each test plot, four moving means were manually computed from the adjacent plots within a column (Fig. 2). The number of adjacent plots included in the computation of the moving mean covariate varied from 2 (Cov. 1) to 5 (Cov.4). In each case, the moving mean was computed using an equal number of plots on each side of the plot in question, except at the end of the blocks, where the appropriate number of plots adjacent to the test plot within a column was used (Fig. 2).

Separate MMCA were performed using Cov. 1 to Cov.4. The MMCA was performed using the GLM procedure (SAS, 1988). The covariate, blocks, and genotypic terms were fitted sequentially to the model. The sequential (Type I) sums of squares (SS I), which correspond to the partitioning of total variation due to the individual effects as they are added sequentially to the model, were examined. The partial (Type IV) sums of squares (SAS, 1988) were used to test statistical significance of individual effects. Genotypic means adjusted for the covariate were estimated by the least square mean options in the GLM procedure (SAS, 1988). Dunnett's and  $t$  tests were performed to select genotypes that outyielded the check cultivar, VC1973-A, in RCBD and MMCA analysis, respectively.

**Border-row mean covariance analysis.** To evaluate the effectiveness of BRMCA, a mungbean yield trial consisting of 30 genotypes was conducted at AVRDC during Fall 1987. The same experimental procedures used in the previous study were followed. In each experimental unit, the two border rows, (first and sixth rows) were planted with the standard check cultivar VC1973A. The fertility levels of each plot were estimated by plant height and seed yield of the border rows. Mean plant height at 50% flowering (PHT) was estimated from three representative plants in each border row. Mean seed yield/1 m of row (YLD/m) was estimated from plants harvested from 1 m of linear row in each border row 65 days after planting.

These border row means were used in an analysis of covariance to control the experimental error associated with the three response variables, total seed yield, first harvest yield, and mean maturity index (MI). Total seed yield ( $t\text{-ha}^{-1}$ , YLD) was estimated from the pods manually harvested three times from the middle four rows and adjusted for a 12% moisture content. First harvest percentage (FHP) was the proportion of total yield obtained from the first harvest. Mean maturity index was estimated as follows:

$$MI = \frac{\sum(\text{days to each harvest} \times \text{yield at each harvest})}{\text{total yield}}$$

Durbin-Watson D = 1.36, auto-correlation is significant at 1%

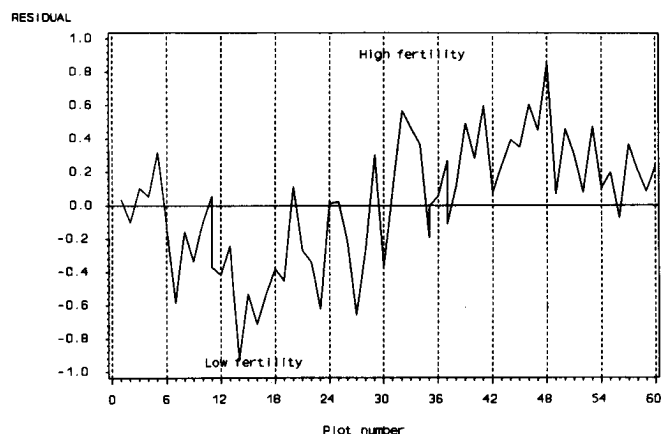


Fig. 1. Autocorrelation between the residuals from the RCBD and the plot arrangement (plot number).

6	7	18	19	30	31	42	43	54	55	→Plot No.
1.7	1.4	1.2	1.7	1.4	2.1	2.3	2.4	2.5	1.7	
5	8	17	20	29	32	41	44	53	56	→Yield (t)
3.0	2.3	2.2	1.8	2.2	2.3	3.3	3.2	2.7	2.0	
4	9	16	21	28	33	40	45	52	57	Block 2
3.0	2.2	2.2	2.1	2.4	2.4	3.1	3.1	2.7	2.3	
3	10	15	22	27	34	39	46	51	58	Block 1
2.5	2.4	2.0	1.9	2.0	2.5	2.6	3.6	2.4	2.7	
2	11	14	23	26	35	38	47	50	59	Block 3
2.2	2.5	1.9	2.2	2.0	2.4	2.9	2.9	2.7	2.7	
1	12	13	24	25	36	37	48	49	60	Block 1
2.1	1.5	1.8	1.7	2.3	2.5	2.8	3.0	2.0	2.7	
c1	c2	c3	c4	c5	c6	c7	c8	c9	c10	Columns

Fig. 2. The plot layout and the moving mean computations.<sup>2</sup>

Test plot	Cov.2	Cov.3	Cov.4	Cov.5
Plot no. 3	$(\#2^2 + \#4)/2$	$(\#1 + \#2 + \#4) + (\#2 + \#4 + \#5)/6$	$(\#1 + \#2 + \#4 + \#5)/4$	$(\#1 + \#2 + \#4 + \#5 + \#6)/5$
Plot no. 1	$(\#2 + \#3)/2$	$(\#2 + \#3 + \#4)/3$	$(\#2 + \#3 + \#4 + \#5)/4$	$(\#2 + \#3 + \#4 + \#5 + \#6)/5$

<sup>2</sup>Numerals following # refer to yield of plots indicated.

Table 1. Variability attributed to various effects in randomized complete-block design (RCBD) and moving mean covariance analysis (MMCA) in mungbean yield trial.

Type of analysis	Covariate	Block	Genotype	Error	CV (%)
	Sequential sums of squares <sup>z</sup>				
RCBD		0.84 <sup>NS</sup> (5.8)	5.30 <sup>NS</sup> (36.7)	8.31 (57.5)	19.9
MMCA					
Cov.1	4.13 <sup>***</sup> (28.7)	0.45 <sup>NS</sup> (3.0)	5.51 <sup>**</sup> (38.1)	4.37 (30.2)	14.6
Cov.2	4.76 <sup>***</sup> (32.9)	0.78 <sup>*</sup> (5.3)	5.36 <sup>**</sup> (37.1)	3.55 (24.7)	13.2
Cov.3	4.98 <sup>***</sup> (34.4)	1.04 <sup>**</sup> (7.3)	5.25 <sup>**</sup> (36.3)	3.12 (22.0)	12.5
Cov.4	5.28 <sup>***</sup> (36.6)	1.14 <sup>**</sup> (7.8)	5.16 <sup>**</sup> (35.8)	2.86 (19.8)	11.9

<sup>z</sup>In parentheses: the percent of the total sums of squares attributed to the effects.

<sup>y</sup>Cov.1–Cov.4 = moving means computed from two, three, four, and five adjacent plots, respectively.

<sup>NS,\*,\*\*,\*\*\*</sup>Significant at  $P = 0.05, 0.01, \text{ or } 0.001$ , respectively.

Table 2. Mean seed yields in randomized complete-block design (RCBD) vs. the adjusted seed yield in the moving mean covariance analysis (MMCA).

Genotype	RCBD	MMCA <sup>z,y</sup>	
		Cov.3	Cov.4
V2984	2.08	2.16	2.16
V3476	1.95	1.76	1.78
V3726	2.02	2.06	2.09
VC1973A (check cultivar)	1.76	1.97	1.93
VC2768A	2.31	2.32	2.35
VC2802A	2.42	2.55*	2.55*
VC2917A	2.51	2.39	2.43
VC3004A	2.18	2.13	2.16
VC3012A	1.91	1.90	1.83
VC3012B	2.39	2.19	2.21
VC3061A	2.66	2.69**	2.68**
VC3117A	2.28	2.43	2.39
VC3178A	2.45	2.30	2.33
VC3300A	2.75	2.61*	2.58*
VC3301A	2.61	2.52*	2.51*
VC3664A	2.19	2.34	2.35
VC3737A	2.83	2.92***	2.88***
VC3738A	2.82	2.77***	2.80***
VC3746A	2.45	2.64**	2.63**
VC3888A	2.32	2.22	2.24
Genotypic effect	NS	**	**
SE	0.27	0.16	0.16

<sup>z</sup>Cov.3–Cov.4 = moving means computed from four and five adjacent plots, respectively.

<sup>y</sup>Significance: different from the check cultivar by  $t$  test.

<sup>NS,\*,\*\*,\*\*\*</sup>Nonsignificant or significant at  $P = 0.05, 0.01, \text{ or } 0.001$ , respectively.

The relative efficiency (RE) of BRMCA and MMCA over RCBD was computed as follows:

$$RE = \frac{\text{error mean square}_{(RCBD)}}{EMS_z} \times \frac{(F_1 + 1)(F_2 + 3)}{(F_2 + 2)(F_1 + 3)}$$

Table 3. Variability attributed to different effects in randomized complete-block design (RCBD) and border-row mean covariance analysis (BRMCA) in a large-scale mungbean trial.

Variable	Sequential sums of squares <sup>z</sup>		
	RCBD	BRMCA-HT <sup>y</sup>	BRMCA-YLD <sup>x</sup>
	<i>Seed yield</i>		
Covariate		0.26 <sup>x</sup> (6.0)***	0.33 (7.3)***
Block	0.03 (0.8) <sup>NS</sup>	0.03 (0.8) <sup>NS</sup>	0.03 (0.8) <sup>NS</sup>
Genotype	2.82 (63.4)**	3.11 (70.0)**	3.06 (68.8)**
Error	1.59 (35.8)	1.03 (23.2)	1.02 (22.9)
SE	0.095	0.077	0.077
CV (%)	10.3	8.4	8.3
R.E (%) <sup>w</sup>	100	154	154
	<i>First harvest percentage</i>		
Covariate		3922.0 (22.9)***	4946.0 (28.9)***
Block	129.4 (0.8) <sup>NS</sup>	132.2 (0.8) <sup>NS</sup>	140.0 (0.8) <sup>NS</sup>
Genotype	8812.0 (51.4)***	6478.6 (37.8)***	5585.0 (37.6)***
Error	8178.0 (47.8)	6588.0 (38.5)	6450.0 (37.6)
SE	6.85	6.20	6.14
CV (%)	14.5	13.2	13.0
R.E (%) <sup>w</sup>	100.0	124.0	128.0
	<i>Mean maturity days</i>		
Covariate		39.2 (22.9)***	49.6 (30.0)***
Block	1.3 (0.7) <sup>NS</sup>	1.3 (0.8) <sup>NS</sup>	1.4 (0.8) <sup>NS</sup>
Genotype	88.1 (51.5)***	67.8 (36.3)**	55.8 (32.6)**
Error	81.8 (47.8)	65.9 (38.5)	64.5 (37.7)
SE	0.69	0.62	0.62
CV (%)	1.6	1.5	1.5
R.E (%) <sup>w</sup>	100.0	124.0	127.0

<sup>z</sup>In parentheses: the percent of the total sums of squares attributed to the effects.

<sup>y</sup>Covariate calculated from mean plant height of border rows.

<sup>x</sup>Covariate calculated from seed yield of 1-m border rows.

<sup>w</sup>R.E % = Error MS<sub>RCBD</sub>/Error MS<sub>BRMCA</sub> corrected for degrees of freedom.

<sup>NS,\*,\*\*,\*\*\*</sup>Nonsignificant or significant at  $P = 0.01 \text{ or } 0.001$ , respectively.

where: EMS<sub>z</sub> = error mean square of MMCA or BRMCA; F<sub>1</sub> = error df of MMCA or BRMCA; and F<sub>2</sub> = error df of RCBD.

## Results and Discussion

*Moving mean covariance analysis.* The highest cv for seed yield was observed in the RCBD design (Table 1). The block and genotypic effects were not significant. Without exception, all the covariance adjustments (Cov. 1 to Cov.4) reduced the error and cv. The covariate and genotypic effects were significant, suggesting that the precision of the analysis increased significantly.

The sequential sum of squares attributed to each individual effect and their percentages of the total sum of squares are presented in Table 1. The cv and the percentage variability attributed to the random error decreased as the number of plots included in the moving mean analysis increased. In similar studies, a maximum reduction in cv was obtained with eight to 18 neighboring plots in moving mean estimation (Mak et al., 1978; Rosielle, 1980). In the current study, five neighboring plots gave the lowest cv. Between five and four neighboring plots, the differences in mean comparisons were not different. The optimum number of adjacent plots to be included in the MMCA would logically vary, depending on the variability of the experimental area (Mak et al., 1978). However, too large a num-

ber should be discouraged, because the size of the block would be increased accordingly, making the estimates less sensitive to changes in the micro soil variation.

With the covariance adjustment, the SE of the mean was reduced and the mean yield, adjusted to the local fertility trend, also differed significantly from the RCBD analysis (Table 2). Seven genotypes significantly outyielded the check cultivar VC 1973A in MMCA analysis, whereas genotypic differences were not significant in RCBD.

These results suggest that covariance adjustment using moving means of neighboring plots gives superior control of the experimental error in situations where many genotypes are evaluated in sites with large micro soil variation. A partially balanced lattice design and MMCA have been compared in error control (Mak et al., 1978). Both methods gave a similar level of error control. In cases where the number of genotypes is too large to accommodate in an incomplete block design, MMCA was found to be advantageous (Mak et al., 1978). Furthermore, use of MMCA has advantages over lattice designs in that the number of treatments is not restricted to a multiple of  $P^r$  ( $P$  = number of plots/block; and  $n > 2$ ) and it is easier to lay out treatments in MMCA (Rosielle, 1980). However, calculation of the moving mean and estimation of the optimum number of neighboring plots could be tedious.

*Border-row mean covariance analysis.* In the second yield trial, blocking was not effective in reducing the experimental error for YLD, FHP, and MM (Table 3). About 36% to 48% of the total variability was attributed to experimental error. Border-row mean covariates, PHT, and YLD/m reduced SE, CV, and the experimental errors in the YLD, FHP, and MI. The relative efficiency of BRMCA over RCBD in the error control varied from 124% to 154%.

These results indicated that BRMCA could be effectively used for error control in large-scale field experiments. Border-row measurements, PHT, and YLD/m reduced the experimental error similarly. Therefore, either of the measurements could be

used in the covariance analysis for error control. However, PHT is simpler to measure than YLD/m. These border-row measurements could be used as covariates without any further computations in the covariance analysis. Computation of moving means or finding the optimum number of neighboring plots in moving means calculations is not necessary. Unlike MMCA, where moving means are computed from the response variable, border-row measurements are not dependent on the response variable. Therefore, statistically, BRMCA could be considered a better analysis than MMCA, since the regression of the dependant variable on the covariate (border-row measurements) after measurement for the blocks and genotypic effect is linear and independent of their effects.

#### Literature Cited

- Baker, R.J. and R.I.H. McKenzie. 1967. Use of control plots in yield trials. *Crop Sci.* 7:335-337.
- Besag, J. and R. Kempton. 1986. Statistical analysis of field experiments using neighboring plots. *Biometrics* 42:231-251.
- LeClerg, E.L. 1966. Significance of experimental design in plant breeding. p. 243-313. In: K.J. Frey (ed.). *Plant breeding*. Iowa State Univ. Press, Ames.
- Li, J.C.R. and K.R. Keller. 1951. An application of serial correlation in field experiments. *Agron. J.* 43:201-203.
- Mak, C., B.L. Harvey, and J.D. Berdahl. 1978. An evaluation of control plots and moving means for error control in barley nurseries. *Crop Sci.* 18:870-873.
- Neter, J., W. Wasserman, and M.H. Kutner. 1983. *Applied linear regression models*. Richard D. Irwin, Homewood, Ill.
- Park, H.G. 1978. Procedure for mungbean evaluation trials. *International Cooperators Guide*. AVRDC, Shanhua, Tainan, Taiwan.
- Rosielle, A.A. 1980. Comparison of lattice designs, check plots, and moving mean in wheat breeding trials. *Euphytica* 29: 129-133.
- SAS Institute, Inc. 1985. *SAS user's guide: Statistics*. version 5. SAS, Cary, N.C.
- Townley-Smith, T.F. and E.A. Hurd. 1973. Use of moving means in wheat yield trials. *Can. J. Plant Sci.* 53:447-450.