Arunachalam, V.
The Fallacy Behind the Use of Modified Line x Tester Design.
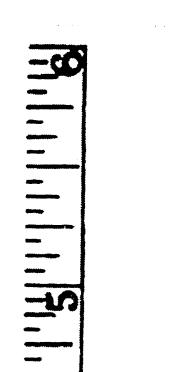*Indian Journal of Genetics and Plant Breeding,* 34 (1974): 280-287.

Record no. D-25

THE FALLACY BEHIND THE USE OF A MODIFIED LINE × TESTER DESIGN

V. Arunachalam

Division of Genetics, Indian Agricultural Research Institute, New Delhi 110012

The utility of the line × tester design (Kempthorne, 1957) in deciding about the relative capacity of a number of female and male parents to produce desirable hybrids is well-known. The salient differences between a diallel (DL) and a line × tester (LT) design are: (a) In DL, parents enter the design in their own right as the crosses $i \times i$, where i is one parent (diagonal entries), while in LT no such provision exists; (b) the same parents are used both as male and as female in DL, while the male and the female parents are invariably different in LT and (c) these considerations lead to the obvious result that an LT with the same number of male and female parents, say **n**, is not equivalent to a DL of the size **n** × **n**.

Plant breeders in India have recently been using a modified line × tester (MLT) design by including the parental lines also in a bid to obtain a single degree of freedom for the contrast, 'parents *vs.* hybrids'. The way the modification is employed at the field level and the mode of subsequent statistical analysis of the data collected lead to a large number of unreasonable and unscientific steps. This paper attempts a scrutiny of the procedures involved and employed, in addition to presenting the correct mode of utilizing an LT. A practical example of a line × tester design in *Brassica campestris* var. brown *sarson* is appended to illustrate the points in an effective manner.

Materials and Methods

The material for the practical example consisted of 10 female and 6 male parents selected, on the basis of the variability for a number of yield components, from populations generated by disruptive and stabilising selection in brown *sarson*. The F_1 hybrids were grown in a randomised blocks design (r.b.d.) with two replications during rabi 1972–73. Parents were grown separately in a contiguous block of field in a similar r.b.d. with two replications. The material was grown in ridges spaced 75 cm. apart and 10 cm. between plant to plant in 2·4 m. rows. For purposes of the present paper, the data on samples of 5 plants per line for plant height (cm.) are utilised.

Results and Discussion

We shall present first the model of LT (Kempthorne, 1957) for clarity and convenient reference and the modification (MLT) later.

The usual LT design is based on the model $Y_{ijk} = \mu + g_i + g_j + s_{ij} + r_k + e_{ijk}$ where $Y_{ij.}$ = height (or any character measured) of the cross **i** × **j** in the kth

replication; g_i=g.c.a. effect of the parent i; s_{ij}=s.c.a. effect of the cross $i \times j$; r_k=k^{th}. replication effect; e_{ijk}=environmental effect peculiar to $(ijk)^{th}$ individual and μ=population mean effect.

The estimation of genetic parameters follows from the full-sib, half-sib relationships that are inherent in the model. Plants within the cross $i \times j$, for example, in any one replication or more, are full sibs and plants of the cross $i \times j$ and $i \times l$, for example, in any one replication or more are half-sibs. The data can thus be put in the form of a two-way table with males and females forming the two arrays:

Female \ Male	1	2	3	..	..	..	j	..	..	s	Marginal Total
I											
II											
..											
..											
i							$Y_{ij.}$				$f_i \cdot$
..											
d											
							$m \cdot j$				$Y \cdots$

$$Y_{ij.} = \sum_{t=1}^{k} Y_{ijt}$$

$$f_{i.} = \sum_{j=1}^{s} \sum_{t=1}^{k} Y_{ijt}$$

$$m_{.j} = \sum_{i=1}^{d} \sum_{t=1}^{k} Y_{ijt}$$

$$Y_{...} = \sum_{i} \sum_{j} \sum_{t} Y_{ijt}$$

The combining ability effect of a female parent i, is judged by the performance of the s hybrids in which i is the female parent. In other words, the marginal totals f_i. determine the combining ability variance in females, and similar arguments hold for males. As one would easily expect, the precision of estimation would increase with increase in d and s. Further, when s and d are not equal, if $s > d$, the female effects would be estimated with greater precision than the male effects. The diallel, since $s = d$ always, this contingency does not arise. This is a reason why LT ranks only next to DL in this regard, though the advantages of LT over DL cannot also be underscored.

Using the principles mentioned above, Kempthorne (1957) has presented an analysis of variance and expectations of the mean sum of squares, using which

TABLE 1

Analysis of variance for combining ability based on LT (Plot Means)

Source	d.f.	m.s.	E (m.s.)
Males (M)	s–1	M_1	E+ rI+ rdy
Females (F)	d–1	M_2	E+ rI+ rsy
M × F	(s–1) (d–1)	M_3	E+ rI
Error	(r–1) (sd–1)	M_4	E

r=no. of replications; s=no. of male parents;
d=no. of female parents; y=cov (half-sibs)=Var. (g.c.a.);
2y+ I=cov (full sibs); E=environmental variance; I=Var. (s.c.a.)

TABLE 2

Analysis of variance based on MLT (Plot Means)

Source	d.f.
Parents+ Hybrids	sd+s+d–1
Parents	s+d–1
Males	s–1
Females	d–1
M *vs* F	1
Hybrids	sd–1
Males	s–1
Females	d–1
M × F	(s–1) (d–1)
Parent *vs* Hybrids	1
Error	(r–1) (sd+s+d–1)

r=no. of replications; s=no. of male parents and d=no. of female parents.

the components of the combining ability variances can be estimated (Table 1; for problems associated with estimation, see discussion).

The LT as modified by plant breeders (MLT) would result in a different ANOVA. Here the male and the female parents are also grown along with the hybrids in a r.b.d. by completely randomising the parents and hybrids within a replication. It is the usual practice to analyse the data using MLT as shown in Table 2. All the components of variance are tested for their significance against the error m.s. using an F test. The significance of the 1 d.f. comparison, parents **vs** hybrids, is taken to indicate the presence of heterosis for this character. It may be pertinent to point out here, that this comparison tests only the difference between the parental and hybrid mean. Significance

would result even when the hybrids are consistently inferior in performance to parents. Further, when there is significance, it is quite possible that many hybrids do not show heterosis and significance is caused by only a few highly heterotic hybrids. In addition, tests using an F-ratio with 1 d.f. are known to be less efficient than those with large d.f. However, the method of analysis shown in Table 2 is statistically valid.

Things are, however, different with the practice of re-analysing the data based on **rsd** entries only taken from the total of **r(sd+s+d)** entries of the MLT and to use the expected m.s. given by Kempthorne's LT (Table 1) to get estimates of g.c.a. and s.c.a. variances. This practice is resorted to, perhaps due to the immediate non-availability of the expectations of the mean squares for the ANOVA for MLT (Table 2).

An effective scrutiny of the procedures adopted in practice while using both LT and MLT designs would demand as a first step, a comparison of the statistical models inherent in these designs. We have already mentioned the model for LT; the model for MLT which inolves both the parents and the hybrids can be put in the following form:

$$Y_{ijk} = \mu + \alpha f_i + \beta m_j + \gamma(g_i + g_j + s_{ij}) + r_k + e_{ijk}$$

where μ, g_i, s_{ij}, r_k and e_{ijk} have the same interpretations as in the model for LT, f_i = effect of the ith female parent, m_j = effect of the jth male parent, and $\alpha = 1$; $\beta = 0$; $\gamma = 0$ for combinations (iik), $\alpha = 0$; $\beta = 1$; $\gamma = 0$ for combinations (jjk), $\alpha = 0$; $\beta = 0$; $\gamma = 1$ for combinations (ijk), $i \neq j$.

These restrictions are necessitated in order to accommodate the parental effects in the model. These very restrictions differentiate this model from that of the LT. One can easily appreciate that, while the model for LT is simple with g_i and g_j giving the combining ability effects of the parents, this model has two parameters each—f_i and g_i for females and m_j and g_j for males—relevant to the g.c.a. effects of the parents. It would be necessary to use the expectations of mean squares of this model only, in order to estimate the combining ability variances, if one has laid out the experiment according to this model. It is obvious that the procedure of analysing the data based on a different model using only the entries relating to the hybrids from this experiment is completely invalid.

An alternative argument can also be appended. Having laid out an experiment with both the parents and the hybrids randomised within each replication, if one uses the data on hybrids alone, this would be equivalent to a field design with gaps corresponding to the parental lines in each replication. This will obviously not conform to the basic requirements of a layout of a design. The analogy can also be understood if one realises the fact that the error in MLT is based on (r–1) (sd+s+d–1) d.f. while its counterpart in LT is based on (r–1)(sd–1) d.f. This may lead to an argument that the error in MLT can be split into that due to hybrids and a residual component. In fact, the residual component is the sum of (reps ×parents) interaction with (r–1) (s+d–1) d.f. and (reps × parents *vs* hybrids) with (r–1) d.f. Though this split-up is technically

valid, apart from the question of how to use these components to test the s.s. due to several sources, the fact remains that the corresponding component of error based on hybrids alone, cannot be used to estimate the variances using the expectations of LT, since the expectations of m.s. based on MLT would definitely differ from those based on LT.

A heuristic argument is generally advanced that there would not be much of a difference in the actual values of error variance between LT and MLT design. Even assuming that the error variances in both these cases are equal, one cannot still justify the procedure as there are technical flaws relating to the statistical models as explained earlier.

For illustration, the analysis of data on height from the experiment on brown *sarson* was done using both the LT and MLT designs (Table 3). In order to take into consideration the plant to plant variation, the analysis was based on data from single plants (Table 4).

It is seen from Table 3 that only the error variance in LT has not differed much from that in MLT but the other components did vary resulting in discrepancies in the estimates, those of cov (F.S.) and of σ^2_{sca} in particular. A large number of published papers can be found in which one would come across instances where the differences were pronounced and those where the differences were nominal (Goud, 1971; Chandra, Sidhu and Arora, 1969; Anand, Rana and Jain, 1972; Murty, Tiwari and Harinarayana, 1967). For example, Goud (1971) reported an error variance of 200·44 (117 d.f.) for plant height under LT and of only 57·41 (183 d.f.) under MLT in *Sorghum*. This resulted in non-significant female variance in the former and significant female variance in the latter from partition of hybrids s.s. The MLT was further complicated by the inclusion of 3 local checks in the experiment by Chandra *et al.* (1969). For purposes of highlighting the conflicting results obtained in breeding experiments by the inappropriate use of LT and MLT, a list of some published papers is included under REFERENCES.

In addition to the above pitfalls, there are some problems in the estimation of cov (half-sibs), denoted by y for convenience and cov (full sibs), denoted by x in the LT design (Table 1). In general, the formulae provided by King *et al.*, (1961) are in vogue for estimation of these components by equating the observed mean squares to their expectations. We give them below for convenient reference:

$$y = \text{cov (half-sibs)} = (M_1 + M_2 - 2M_3)/[r(s+d)]$$
$$x = \text{cov (full sibs)} = (1/3r) [M_1 + M_2 + M_3 - 3M_4 + ry(6-s-d)]$$

We shall call these as the Current Estimates (CE). A close look at the expectations of the m.s. of the various sources under LT (Table 1) would reveal that there are only 2 variables, x and y, but there are 3 equations corresponding to (M_1-M_4), (M_2-M_4), (M_3-M_4). A unique set of solutions cannot result by solving these equations. Any two or any possible combination of the 3 equations into two would result in feasible solutions. Thus the CE given above are only one set out of all the feasible sets of solutions.

TABLE 3

ANOVA of plant height in brown sarson

Source	d.f. LT	d.f. MLT	m.s. LT	m.s. MLT			@	Est. LT	Est. MLT
Reps	1	1	812·0	1433·0**			1	229·9	303·2
Hybrids	59	59	1371·3**	1371·3**		x			
Males	5	5	1421·8**	1421·8**			2	46·0	
Females	9	9	816·9**	816·9**			1	−22·3	−22·3
M×F	45	45	1476·6**	1476·6**	C.E.	y			
Parents	—	15	—	1559·3**			2	−4·5	
Males	—	5	—	2853·4**			1	274·5	347·8
Females	—	9	—	968·7*		s			
M vs F	—	1	—	407·4			2	54·9	
Par. vs Hyb.	—	1	—	11874·0**					
Reps × Entries+	59	75	927·5**	780·9**	L.E.	x		178·3	251·6
						y		−6·9	−6·9
Error	480	608	242·0	250·3		s		192·0	265·3

+ Entries=Hyb for LT, =(Par+Hyb) for MLT; @ 1–Based on single plant data; 2–Based on plot means; x–cov (full sibs); y–Cov (half-sibs); s=σ^2s.c.a; σ^2g.c.a. not estimable; *–Significant at 5% level; **– Significant at 1% level; C.E.–conventional estimates; L.E.–Least Squares estimates.

TABLE 4

ANOVA of LT on single plant basis

Source	d.f.	m.s.	E (m.s.)
Males (M)	s−1	M_1	$E + nE_1 + rI + rdy$
Females (F)	d−1	M_2	$E + nE_1 + rI + rsy$
M × F	(s−1) (d−1)	M_3	$E + nE_1 + rI$
Rep × Hyb	(r−1) (sd−1)	M_4	$E + nE_1$
Error	rsd (n−1)	M_5	E

n—No. of plants sampled per line; E_1—variance associated with (reps × hyb); other symbols as in Table 1.

It is preferable, therefore, to obtain estimates of x and y by least squares technique by minimising the sum of squares of deviations of (M_1-M_4), (M_2-M_4) and (M_3-M_4) from their expectations. In the case of LT analysis on plot means basis, M_4 will refer to error m.s. and on single plant basis, M_4

will refer to m.s. due to (reps×hybrids). The least squares estimates (LE) are given below:

Let $a=(M_1-M_4)/r$, $b=(M_2-M_4)/r$, $c=(M_3-M_4)/r$. Then,

$$x(sd-s^2-d^2)=s(a+c-2b)+d(b+c-2a)-\tfrac{1}{2}[s^2(a+c)+d^2(b+c)-sd(a+b)]$$

$$y(sd-s^2-d^2)=\tfrac{1}{2}[s(a+c-2b)+d(b+c-2a)]$$

$\sigma^2{}_{gca}=y$ and $\sigma^2{}_{sca}=x-2y$

The estimates CE and LE can be shown to be equal when $s=d$.

When $s=d$, $y=(M_1+M_2-M_3)/2rs$ and $x=2y+(1/r)(M_3-M_4)$.

The differences in the estimates, CE and LE are clearly brought out in the example considered (Table 3).

At this stage, it would be useful to point out the need to take into account the plant to plant variation whenever it exists in LT analysis; otherwise, one would obtain biased estimates from the analysis based on plot means. This point also would be clear in the example considered (Table 3). This would emphasise the need for caution in deciding the size of the sample of plants that would be adequate for LT analysis. In fact, if a reasonably large samples of plants is used in the each line, an analysis on the basis of plot means may, sometimes, be justified, if variation observed is levelled both in the positive and in the negative side. But recording data on a large sample per line in a replicated experiment will be a laborious process. On the other hand, if a moderate sample size is used and if intra-sample variation is present, it would be fruitful to carry out the analysis of data on single plants.

In conclusion, the following points emerge from this study:

(i) In general, it is unnecessary to use MLT. The superior performance of hybrids or otherwise can be easily tested by t- test. The way an MLT is used and analysed by breeders is erroneous.

(ii) If circumstances are exceptional warranting the inclusion of parents, it would be beneficial to grow the hybrids together followed by parents in contiguous plots in a r.b.d. (randomising hybrids and parents separately) so that hybrid plots alone can be considered to form a separate r.b.d. When needed, pooled ANOVA over hybrids and parents can be done.

(iii) The ideal situation would, however, be to obtain the expected m.s. for MLT and use them for estimation of combining ability components.

SUMMARY

The genetic models governing a line × tester design and a modified design in which the parents are included are examined in detail. The errors involved in the procedure of analysis followed by breeders are highlighted with a practical example. The need to scrutinise plant to plant variation and to modify the analysis on single plant basis is stressed. Least squares estimates of the components of combining ability variances are derived and their superiority over the conventional estimates brought out.

REFERENCES

Anand, I. J., Rana, B. S. and Jain, O. P. (1972).   Estimation of genetic variances by full-sib and half-sib analysis in linseed.   *SABRAO Newsletter,* **4:** 33–37.

Chandra, S., Sidhu, G. S. and Arora, N. D. (1969).   Line × tester studies on some male-sterile and pollinator parents in forage *Sorghum. Indian J. agric. Sci.,* **39:** 690–698.

Goud, J. V. (1971).   Combining ability in *Sorghum. Indian J. agric. Sci.,* **41:** 924–931.

Kempthorne, O. (1957).   *An Introduction to Genetic Statistics.*   John Wiley and Sons, New York.

King, J. G., Quinby, J. R., Stephens, J. C., Kramer, N. W. and Lahr, K. A. (1961).   An evaluation of parents of grain *Sorghum* hybrids.   *Bull. Texas Agric. Exp. Sta.,* M.P. 510.

Murty, B. R., Tiwari, J. L. and Harinarayana, G. (1967).   Line × tester analysis of combining ability and heterosis for yield factors in *Pennisetum typhoides* (Burm.) S. and H. *Indian J. Genet.,* **27:** 238–245.

Rana, B. S. and Murty, B. R. (1971).   Genetic analysis of resistance to stem borer in *Sorghum. Indian J. Genet.,* **31:** 521–529.

Rao, N. G. P., Rana, V. K. S. and Tripathi, D. P. (1968).   Line × tester analysis of combining ability in *Sorghum. Indian J. Genet.,* **28:** 231–238.

Singh, B. B., Murty, B. R. and Butany, W. T. (1971).   Line × tester analysis of combining ability and heterosis for some yield and quality components in upland cotton.   *Indian J. Genet.,* **31:** 8–15.

Singh, K. B. and Singh, J. K. (1971).   Heterosis and combining ability in black gram.   *Indian J. Genet.,* **31:** 491–498.

Singh, T. H., Gupta, S. P. and Phul, P. S. (1971).   Line × tester analysis of combining ability in cotton.   *Indian J. Genet.,* **31:** 316–321.