THE ESTIMATION OF MISSING PLOT VALUE IN SPLIT-PLOT AND STRIP TRIALS

BY S. A. KHARGONKAR

Institute of Plant Industry, Indore, C.I.

WHILE agricultural experimenters are familiar with formulæ for estimating missing plot yields in simple randomised block and Latin square layouts (Allan and Wishart, 1930; Yates, 1933; Hutchinson and Panse, 1937), no attempt seems to have been made so far to obtain formulæ for the missing sub-plot value in split plot and strip designs which are in common use in modern field experimental work. The object of the present paper is to present appropriate formulæ for calculating missing values in such trials.

The following two types of designs are considered:—

1. Split plot designs in randomised block and Latin square layouts.
2. Designs involving strip treatments.

The missing plot value in each case is derived by a straightforward extension of Yates' (1933) method of minimising the error variance.

In a simple split plot design consisting of main and sub-plots there are really two analyses of variance corresponding to the two types of plots. Thus there are two error variances, say, a for main plot analysis and b for sub-plot analysis. To estimate the missing value of a sub-plot we shall be concerned with minimising the error variance b. If more than one sub-plot are missing their yields can be estimated by the repeated application of the formula by substituting approximate values, equal to the general mean, treatment means or block means for all the other missing plots, and then obtaining a second approximation for each missing plot in turn in the manner explained by Yates (1933). The degrees of freedom allotted to the residual variance should, of course, be reduced by the number of missing values calculated.

For simplicity the case of a simple split-plot design having two plot sizes is considered with the help of an actual example although the method of deriving the formula for the missing sub-plot value is quite general.

Table I gives the yield data from an experiment on cotton carried out at Ganganagar Farm, Bikaner State. The experiment consists of four main treatments, A, B, C and D, being combinations of manuring and no manuring with heavy and moderate irrigations. The sub-treatments are two varieties V_1 and V_2. There were six replications.

TABLE I

Yield of kapas per plot (½ ch. units)

Treatments		Blocks						Total
Main treatments	Sub-treatments	I	II	III	IV	V	VI	
A	V_1	173	162	115	82	146	89	767
	V_2	196	252	204	208	257	142	1259
Total ..		369	414	319	290	403	231	2026
B	V_1	86	68	2	79	25	59	319
	V_2	165	159	45	232	102	98	801
Total ..		251	227	47	311	127	157	1120
C	V_1	110	152	57	126	63	97	605
	V_2	152	216	169	193	201	151	1082
Total ..		262	368	226	319	264	248	1687
D	V_1	91	87	8	80	64	63	393
	V_2	109	168	71	158	131	139	776
Total ..		200	255	79	238	195	202	1169
Grand Total		1082	1264	671	1158	989	838	6002

The various quantities relating to the analysis of variance will be denoted by the following symbols:—

Let

the number of blocks, *e.g.*, I, II, etc., in Table I q

the number of main treatments or the number of main plots per block, *e.g.*, A, B, C and D in Table I .. .. m

the number of sub-plots per main plot or the number of sub-treatments, *e.g.*, V_1, V_2 in Table I .. p

the number of sub-plots per block or the number of sub-plot treatments .. .. $m' = m \times p$

the total number of sub-plots .. .. .. $n = m \times p \times q$

Also let

the sum of known yields in the main plot containing the missing sub-plot .. .. .. M

the sum of known yields in the block containing the missing sub-plot .. .. B

the sum of known yields in the main treatment containing the missing sub-plot .. .. P

the sum of known yields in the sub-treatment containing the missing sub-plot .. .. .. Q

the sum of known yields in the sub-plot treatment containing the missing sub-plot .. .. R

the sum of all known yields .. .. .. T

Then if x is the value of a missing sub-plot, the terms in the sums of squares containing x will be

due to main plots, $\frac{1}{p}(M+x)^2 - \frac{1}{n}(T+x)^2$

due to main treatments, $\frac{1}{pq}(P+x)^2 - \frac{1}{n}(T+x)^2$

due to sub-treatments, $\frac{1}{mq}(Q+x)^2 - \frac{1}{n}(T+x)^2$

due to sub-plot treatments, $\frac{1}{q}(R+x)^2 - \frac{1}{n}(T+x)^2$

due to interaction main treatments $\times$ sub-treatments,

$$\frac{1}{q}(R+x)^2 - \frac{1}{pq}(P+x)^2 - \frac{1}{mq}(Q+x)^2 + \frac{1}{n}(T+x)^2$$

Also the total s.s. will be $x^2 - \frac{1}{n}(T+x)^2$.

The contribution of terms containing x to the sub-plot error sum of squares will be

(due to total) — (due to main plots) — (due to sub-treatments) — (due to interaction main treatments $\times$ sub-treatments)

which after substitution of the above quantities, can be expressed after simplification as

$$x^2 - \frac{1}{p}(M+x)^2 + \frac{1}{pq}(P+x)^2 - \frac{1}{q}(R+x)^2.$$

Differentiating and equating to zero we get

$$pqx - qx + x - px - qm - pR + P = 0$$

which gives,

$$x = \frac{pR + qM - P}{(p-1)(q-1)} \qquad \text{(I)}$$

where p, q, R, M and P represent the quantities explained above.

Example 1.—In Table I assuming the value 82 in block IV in the sub-treatment V_1 in the main treatment A as missing and applying the formula, the value of x is obtained as

$$x = \frac{2 \times 685 + 6 \times 208 - 1944}{(2-1)(6-1)} = 134.8,$$

the value that would be substituted if the original value was really missing.

Similarly taking the value 204 in block III in the sub-treatment V_2 in the main treatment A as missing, its value is estimated from the formula as

$$x = \frac{2 \times 1055 + 6 \times 115 - 1822}{(2-1)(6-1)} = 195.6.$$

In split-plot designs with further subdivisions of sub-plots there are more analyses to be done, depending on the number of divisions. In the above example if the varietal plots V_1 and V_2 are split into two more plots, for, say, spacing treatments S_1 and S_2 then the ultimate plots will be those for S_1 and S_2. There will then be three error variances: (1) for main plots corresponding to manures and irrigation treatments, (2) for first order sub-plots corresponding to varieties and (3) for second order sub-plots corresponding to spacings. To estimate the missing value of a second order sub-plot, the third error variance will have to be minimised.

If x is a missing sub-plot value, then the formula for calculating x derived by a procedure exactly similar to that explained above is shown below:—

$$x = \frac{p'R' + qM' - P'}{(p'-1)(q-1)} \tag{II}$$

where p' = number of second order sub-treatments (spacing in the above illustration);

R' = sum of known yields in the second order sub-plot treatment, made up of second order sub-plots over all blocks and containing the missing sub-plot;

q = number of blocks;

M' = sum of known yields in the first order sub-plot containing the missing sub-plot;

P' = sum of known yields in the first order sub-plot treatment containing the missing sub-plot.

If the split-plot design is arranged in a Latin square instead of randomised blocks, the value of a missing sub-plot can be found by minimising the sub-plot error variance, by following a procedure similar to that used for split-plot designs in randomised blocks. Thus if x is the value of a missing sub-plot, the contribution of terms containing x to the sub-plot error sum of squares will be,

$$x^2 - \frac{1}{p'}(M+x)^2 - \frac{1}{p}(R+x)^2 + \frac{1}{pp'}(P+x)^2.$$

Differentiating and equating to zero the value of x is obtained as

$$x = \frac{p'R + pM - P}{(p'-1)(p-1)} \tag{III}$$

where p = number of main plots per row or column or number of main treatments;

p' = number of sub-plots per main plot or number of sub-treatments;

R = sum of known yields in the sub-plot treatment consisting of all sub-treatment plots in the main treatment containing the missing sub-plot;

M = sum of known yields in the main plot containing the missing sub-plot;

P = sum of known yields in the main treatment containing the missing sub-plot.

The formula for the missing sub-plot value in a Latin square split-plot design involving three plot sizes is

$$x = \frac{p''R' + pM' - P'}{(p''-1)(p-1)} \tag{IV}$$

where p = number of main plots per row or per column or number of main treatments;

p'' = number of second order sub-plots per first order sub-plot or number of second order sub-treatments;

R' = sum of known yields in the second order sub-plot treatment consisting of all second order sub-treatment plots in the first order sub-plot treatment containing the missing sub-plot;

M' = sum of known yields in the first order sub-plot containing the missing sub-plot;

P' = sum of known yields in the first order sub-plot treatment containing the missing sub-plot.

Estimation of yield of a missing sub-plot value in split-plot designs involving strip treatments is considered next. The procedure can be explained better with the help of an illustration though here again the method is quite general.

Table II shows the yields in a paddy experiment carried out at Dacca. There were four replications. A and B are the seedling treatments representing row-strip treatments, s_1, s_2, s_3 and s_4 are the spacing treatments representing the column-strip treatments and V_1 and V_2 are the varietal treatments representing sub-row strip treatments. Thus there are five different plot sizes, namely, row-strip plots, sub-row strip plots, column-strip plots, interaction plots formed by the intersection of row-strips and column-strips and the ultimate sub-plots formed by the intersection of sub-rows and columns. There are thus five errors a, b, c, d, e corresponding to these five plot sizes.

TABLE II

Yield of grain in the paddy experiment at Dacca

	Block I					Block II					Block III					Block IV					Total
	s_1	s_2	s_3	s_4	Total	s_1	s_2	s_3	s_4	Total	s_1	s_2	s_3	s_4	Total	s_1	s_2	s_3	s_4	Total	
A V_1 ..	17	21	55	65	158	16	34	48	65	163	15	27	50	61	153	15	34	46	76	171	645
V_2 ..	16	30	54	58	158	21	29	57	72	179	19	32	45	64	160	22	36	50	67	175	672
Total ..	33	51	109	123	316	37	63	105	137	342	34	59	95	125	313	37	70	96	143	346	1317
B V_1 ..	12·5	24	44	66	146·5	15	19	45	67	146	15	33	51	64	163	16	26	48	68	158	613·5
V_2 ..	14	30	59	73	176·0	15	26·5	51	65	157·5	19	32	47	63	161	16	30	43	66	155	649·5
Total ..	26·5	54	103	139	322·5	30	45·5	96	132	303·5	34	65	98	127	324	32	56	91	134	313	1263
Grand Total ..	59·5	105	212	262	638·5	67	108·5	201	269	645·5	68	124	193	252	637	69	126	187	277	659·0	2580

Seedling treatments (A, B) = Row-strip treatments.

Varietal treatments (V_1, V_2) = Sub-row strip treatments.

Spacing treatments (s_1, s_2, s_3, s_4) = Column-strip treatments.

where p = number of main plots per row or column or number of main treatments;

p' = number of sub-plots per main plot or number of sub-treatments;

R = sum of known yields in the sub-plot treatment consisting of all sub-treatment plots in the main treatment containing the missing sub-plot;

M = sum of known yields in the main plot containing the missing sub-plot;

P = sum of known yields in the main treatment containing the missing sub-plot.

The formula for the missing sub-plot value in a Latin square split-plot design involving three plot sizes is

$$x = \frac{p''R' + pM' - P'}{(p'' - 1)(p - 1)} \qquad \text{(IV)}$$

where p = number of main plots per row or per column or number of main treatments;

p'' = number of second order sub-plots per first order sub-plot or number of second order sub-treatments;

R' = sum of known yields in the second order sub-plot treatment consisting of all second order sub-treatment plots in the first order sub-plot treatment containing the missing sub-plot;

M' = sum of known yields in the first order sub-plot containing the missing sub-plot;

P' = sum of known yields in the first order sub-plot treatment containing the missing sub-plot.

Estimation of yield of a missing sub-plot value in split-plot designs involving strip treatments is considered next. The procedure can be explained better with the help of an illustration though here again the method is quite general.

Table II shows the yields in a paddy experiment carried out at Dacca. There were four replications. A and B are the seedling treatments representing row-strip treatments, s_1, s_2, s_3 and s_4 are the spacing treatments representing the column-strip treatments and V_1 and V_2 are the varietal treatments representing sub-row strip treatments. Thus there are five different plot sizes, namely, row-strip plots, sub-row strip plots, column-strip plots, interaction plots formed by the intersection of row-strips and column-strips and the ultimate sub-plots formed by the intersection of sub-rows and columns. There are thus five errors a, b, c, d, e corresponding to these five plot sizes.

The formula for estimating a missing value in the ultimate sub-plot can be obtained by minimising the error variance e. The contribution of terms containing x, which represents the missing value to this error sum of squares is given by

(Total s.s. due to all sub-plots) — (due to row-strip plots) — (due to column-strip treatments) — (due to erorr b) — (due to row-strip treatments × column-strip treatments) — (due to error c) — (due to sub-row strip treatments) — (due to sub-row strip treatments × row-strip treatments) — (due to error d) — (due to sub-row strip treatments × column-strip treatments) — (due to sub-row strip treatments × row-strip treatments × column-strip treatments).

Substituting the appropriate algebraic quantities and simplifying, the following expression is obtained:—

$$x^2 - \frac{1}{p_1}(M''+x)^2 - \frac{1}{q_1}(m+x)^2 + \frac{1}{p_1 q_1}(M+x)^2 - \frac{1}{q}(Q''+x)^2$$

$$+ \frac{1}{qq_1}(Q+x)^2 - \frac{1}{pqq_1}(P+x)^2 + \frac{1}{qp_1}(R+x)^2$$

Differentiating and equating to zero gives,

$$x = \frac{p_1(q_1 Q'' - Q) + q(p_1 m + q_1 M'' - M) - q_1 R + P}{(p_1 - 1)(q-1)(q_1 - 1)} \qquad \text{(V)}$$

where p_1 = number of sub-row strip treatments;
q = number of blocks;
q_1 = number of column-strip treatments;
m = sum of known yields in the sub-row strip plot containing the missing sub-plot;
M = sum of known yields in the row-strip plot containing the missing sub-plot;
M'' = sum of known yields in that portion of the column-strip plot which is cut by row-strip plot and containing the missing sub-plot;
Q = sum of known yields in the sub-row strip treatment from a row-strip treatment and containing the missing sub-plot;
Q'' = sum of known yields in the sub-row strip plots from a column-strip treatment in a row-strip treatment containing the missing sub-plot;
R = sum of known yields in the column-strip treatment from a row-strip treatment containing the missing sub-plot;
P = sum of known yields in the row-strip treatment containing the missing sub-plot.

Example 2.—Taking the value 48 in Block II in column strip treatment s_3 in the sub-row strip treatment V_1 from the row-strip treatment A as missing and applying the formula, we get its value as

$$x = \frac{2(4\times151-597)+4(2\times115+4\times57-294)-4\times357+1269}{1\times3\times3} = 56.8$$

Similarly, assuming the value 57 in the block II in the same column-strip treatment s_3 in the sub-row strip treatment V_2 from the row-strip treatment A, as missing and applying the same formula we get its value as

$$x = \frac{2(4\times149-615)+4(2\times122+4\times48-285)-4\times348+1260}{1\times3\times3} = 48.2$$

If instead of row-strips, column-strips are subdivided for introducing further treatments, the same formula (V) will be applicable for calculating a missing value in the ultimate sub-plot. The only change necessary will be to replace row and sub-row strips in the formula by column and sub-column strips.

When there is no subdivision of either rows or columns, there will then be only three plot sizes, namely, row-strip plots, column-strip plots and the ultimate plots formed by the intersection of column strips and row-strips. There will be three errors corresponding to these three plot sizes. The formula for a missing value x of an ultimate plot can be found in this case by minimising the third error variance. The formula so obtained is,

$$x = \frac{p(q_1R - P) + q(pM + q_1M' - B) - q_1S + T}{(p-1)(q-1)(q_1-1)} \qquad \text{(VI)}$$

where q, q_1, P, R and M represent the same symbols as given in the formula V and

p = number of row-strip treatments;

M' = sum of known yields in the column-strip plot containing the missing plot;

B = sum of known yields in the block containing the missing plot;

S = sum of known yields in the column-strip treatment containing the missing plot;

T = sum of all known yields.

Finally, in an experimental design where row-strips as well as column-strips are divided into sub-row strips and sub-column strips, there will be eight plot sizes: (1) row-strip plots, (2) column-strip plots, (3) plots formed by the intersection of row-strips and column-strips, (4) sub-row strip plots, (5) plots formed by the intersection of sub-row strips and column-strips, (6) sub-column strip plots, (7) plots formed by the intersection of sub-column strips and row-strips, and (8) plots formed by the intersection of sub-row strips and sub-column strips. There will thus be eight errors corresponding to these plot sizes. The formula for estimating a missing value x in the ultimate sub-plot can be obtained by minimising the eighth error variance. Applying the usual method of procedure the formula for the missing value is obtained as

$$x = \frac{p_1 (q_2 k_3 - Q'') + q\,(p_1 m'' + q_2 m''' - M'') - q_2 Q_1'' + R}{(p_1 - 1)(q - 1)(q_2 - 1)} \qquad \text{(VII)}$$

where p_1, q, M'', Q'' and R represent the same symbols as in the formulæ V, VI and

q_2 = number of sub-column strip treatments;

Q_1'' = sum of known yields in the sub-column strip plots in column-strip treatment from row-strip treatment containing the missing sub-plot;

m'' = sum of known yields in the plot formed by the intersection of sub-row strip and column-strip and containing the missing sub-plot;

m''' = sum of known yields in the plot formed by the intersection of sub-column strip and row-strip and containing the missing sub-plot;

k_3 = sum of known yields in the sub-row strip treatment in sub-column strip treatment from column-strip treatment in row-strip treatment and containing the missing sub-plot.

It is interesting to compare the formulæ derived in the present paper with the formulæ for simple randomised block and Latin square layouts. The formulæ, I to IV, for ordinary split-plot designs closely resemble the formula for simple randomised blocks, provided the main plot in the split-plot layout is considered analogous to the block. Where there is a further division of sub-plots, it is the immediately larger sub-plot that would take the place of the block, as in formulæ II and IV. In fact, as far as the estimation of a missing sub-plot is concerned, the experiment can be properly regarded as only consisting of the replications of the particular main plot, or, in case of further sub-division, of the larger sub-plot, containing the missing sub-plot and is thus reduced to a simple randomised block layout. The formula for the missing plot can then be set down immediately. It should be noted from formulæ III and IV that this analogy holds even when the main plots are arranged in a Latin square, since the Latin square structure is confined to the main plots, while for sub-plot analysis the particular main plot containing the missing plot assumes the place of a block with a number of replications equal to the rows or columns of the Latin square. The missing plot formulæ V, VI and VII for strip treatments are also expressed in a form similar to that of the missing plot formula in randomised blocks; but no simple relationship between these is evident, nor does it appear possible to deduce the formulæ for strip treatments directly from the randomised block formula.

Anderson* (1946) has derived formula I above for a missing plot in a split plot design by the covariance technique (Bartlett, 1937). The method of covariance furnishes an easy means for estimating the bias in the treatment sums of squares and for calculating the standard errors of the treatment means containing the missing value. This method was applied to the designs described in the present paper for calculating the bias in the various treatment sums of squares and incidentally, the formulæ derived earlier by the least square principle for estimating the value of a missing plot were also verified. According to this method let $x = 0$ and $y =$ the actual yield for plots with known yield and $x = -1$ and $y = 0$ for the missing plot. Analyses of variance of x and covariance between x and y are then worked out and the expressions for Sx^2 and Sxy for the various items in the analysis of variance are obtained. The estimate of the yield of the missing plot is then simply the corresponding error regression coefficient $\frac{Sxy}{Sx^2}$. By using the estimated value of the missing plot, all sums of squares except that for the error which is minimised will be slightly overestimated. The bias thus introduced in any treatment sum of squares is:

$$\frac{(Sxy\, S'x^2 - S'xy\, Sx^2)^2}{(Sx^2)^2\, (Sx^2 + S'x^2)}$$

where Sx^2 and Sxy are sum of squares and sum of products for the error minimised and

$S'x^2$ and $S'xy$ are sum of squares and sum of products for a treatment.

Sx^2 and $S'x^2$ are really the respective degrees of freedom divided by total number of plots, *i.e.*, n in all cases.

It is seen that the bias is positive. This bias was calculated for the numerical examples described earlier with a split-plot design with two plot sizes and a strip treatment design with five plot sizes. The % bias is shown in Tables III and IV. Expressions for $S'xy$ and Sxy

* The present paper was first submitted for publication before Anderson's results were published.

are also given for the strip design in Table IV. Corresponding expressions for the split-plot design are not included in Table III as these have been given by Anderson (1946).

TABLE III

Split-plot trial with two plot sizes

(Plot value 82 in example 1 was replaced by 135 calculated as a missing plot)

Due to	% bias in the sum of squares
Main treatments	8·9%
Error (a)	5·5%
Sub-tretments	4·8%
Main treatments x sub-treatments	0·1%
Error (b)	..

TABLE IV

Strip trial with five plot sizes

(Plot value 48 in example 2 was replaced by 56·8 calculated as a missing plot)

Due to	Sums of products $S'xy$ and Sxy	% bias in the sum of squares
Row-strip treatments	$\frac{-pP+T}{n}$	5·3%
Error (a)	$\frac{-pqM+qB+pP-T}{n}$	7·5%
Column-strip treatments	$\frac{-q_1S+T}{n}$	1·2%
Error (b)	$\frac{-qq_1M'+q_1S+qB-T}{n}$	3·1%
Row-strip Tr. × Column-strip Tr.	$\frac{-pq_1R+pP+q_1S-T}{n}$	3·6%
Error (c)	$\frac{-pqq_1M''+pqM+qq_1M'+pq_1R-qB-pP-q_1S+T}{n}$	1·7%
Sub-row strip treatments	$\frac{-p_1Q'''+T}{n}$	5·3%
Sub-row strip Tr. × Row strip Tr.	$\frac{-pp_1Q+pP+p_1Q'''-T}{n}$	5·3%
Error (d)	$\frac{-pp_1qm+pqM+pp_1Q-pP}{n}$	0·3%
Sub-row stripTr. × Column-strip Tr.	$\frac{-p_1q_1Q'+p_1Q'''+q_1S-T}{n}$	0·4%
Sub-row strip Tr. × Row-strip Tr. × Column-strip Tr.	$\frac{-pp_1q_1Q''+pp_1Q+p_1q_1Q'+pq_1R-p_1Q'''-pP-q_1S+T}{n}$ ..	5·6%
Error (e)	$\frac{pqq_1M''+pp_1qm+pp_1q_1Q''-pqM-pp_1Q-pq_1R+pP}{n}$	..

The letters in the above expressions represent the same quantities as explained under formulæ V and VI.

Q' = the sum of known yields in the sub-row strip plots from column-strip treatment containing the missing sub-plot.

Q''' = the sum of known yields in the sub-row strip treatment containing the missing sub-plot.

It will be seen from Tables III and IV that the bias introduced in the various treatment sums of squares is small. Consequently the unbiased treatment sums of squares would not generally differ from the corresponding uncorrected sums of squares to such an extent as to upset the significance of difference between treatments. Only the differences which are just significant might be rendered just non-significant after removing the bias from the treatment sums of squares; but in such cases the result of the test of significance would be regarded as merely suggestive of a real difference both before as well as after the correction for bias. Secondly, the effect of bias will be even less in tests of significance involving errors other than the one which is minimised in estimating the missing plot value; because here both the treatment as well as the error sum of squares will be slightly overestimated. It may be concluded that removal of bias from the treatment sum of squares need not be considered as an essential routine before making tests of significance in the analysis of variance, although this might be done in borderline cases.

In calculating the standard errors for comparisons between two treatments, one of which includes a missing plot, a component of variance corresponding to the contribution of the missing plot value to the treatment mean which includes the missing plot has to be added to the ordinary variance of this treatment mean, which would have been appropriate had there been no missing plot. For example, in the split-plot trial with two plot sizes the variance of a sub-treatment mean without missing value would be $\frac{E_b}{qm}$, where

E_b = mean square for the sub-plot error

q = number of replicates or blocks

m = number of main treatments

If this treatment contained a missing value, the contribution of this value to the treatment mean would be the missing plot value divided by qm. The regression coefficient of y on x for the sub-plot error provides the estimate of the missing value and its variance is

$$\frac{E_{bmpq}}{m(q-1)(p-1)},$$

where E_b, m and q are the same quantities as explained above and p = the number of sub-treatments.

This variance divided by $(qm)^2$ would be the component of the variance of the treatment mean due to the missing plot value so that,

Variance of sub-treatment mean with missing value

$$= \frac{E_b}{qm} + \frac{E_{bp}}{qm^2 (q-1)(p-1)}$$

and

Variance of sub-treatment mean without missing value

$$= \frac{E_b}{qm},$$

Therefore the variance of difference between the two sub-treatment means is

$$\frac{2E_b}{qm} \left\{ 1 + \frac{p}{2m (q-1)(p-1)} \right\} \qquad \text{(VIII)}$$

Similarly the variance of the difference between two main treatment means, one with and the other without a missing unit, would be given by

$$\frac{2}{pq} \left\{ E_a + \frac{E_b}{2 (q-1)(p-1)} \right\} \qquad \text{(IX)}$$

where E_a is the mean square for the main plot error.

The formula for the variance of difference between two sub-plot treatment means in the same main treatment, one with and the other without the missing unit can similarly be derived as

$$\frac{2E_b}{q} \left\{ 1 + \frac{p}{2 (q-1)(p-1)} \right\} \qquad \text{(X)}$$

Two other types of comparisons may occur in split-plot designs. One of them is the comparison between two sub-plot treatments consisting of the same sub-treatment but belonging to different main treatments, *e.g.*, $a_1b_1 - a_2b_1$ where a_1, a_2, etc., represent the main treatments and b_1, b_2, etc., represent sub-treatments. The other type is the comparison between two sub-plot treatments consisting of two different sub-treatments from two different main treatments, *e.g.*, $a_1b_1 - a_2b_2$. For both types of comparisons the formula for the variance of difference as given by Nair (1944), for the ordinary case without any missing plot is $\frac{2}{q} \left\{ \frac{E_a + (p-1) E_b}{p} \right\}$.

Adding to this formula the variance contributed by the missing plot value, we obtain the formula appropriate for the variance of difference between two treatments one of which contains a missing plot. This formula is

$$\frac{2E_a}{pq} + \frac{2E_b}{pq}\left\{(p-1) + \frac{p^2}{2(q-1)(p-1)}\right\} \tag{XI}$$

All these formulæ for the split-plot design involving a missing plot agree with those given by Anderson (1946).

Formulæ for the strip treatment designs can be similarly set down by adding to the ordinary formulæ for variance of difference between two treatments, the component corresponding to the contribution of the missing plot value to the treatment mean. For strip trials with three plot sizes, *i.e.*, where treatments in column and row-strips are laid across one another, the various formulæ are as follows:

V (Difference between row-strip treatment means)

$$= \frac{2}{q_1 q}\left\{E_a + \frac{E_c p}{2(q-1)(q_1-1)(p-1)}\right\} \tag{XII}$$

V (Difference between column-strip treatment means)

$$= \frac{2}{pq}\left\{E_b + \frac{E_c q_1}{2(q-1)(q_1-1)(p-1)}\right\} \tag{XIII}$$

$V(a_1b_1 - a_1b_2)$

$$= \frac{2E_b}{pq} + \frac{2E_c}{pq}\left\{(p-1) + \frac{p^2 q_1}{2(q-1)(q_1-1)(p-1)}\right. \tag{XIV}$$

$V(a_1b_1 - a_2b_1)$

$$= \frac{2E_a}{q_1 q} + \frac{2E_c}{q_1 q}\left\{(q_1-1) + \frac{p q_1^2}{2(q-1)(q_1-1)(p-1)}\right. \tag{XV}$$

$V(a_1b_1 - a_2b_2)$

$$= \frac{2E_a}{q_1 q} + \frac{2E_b}{pq} + \frac{2E_c}{pqq_1}\left\{pq - p - q + \frac{p^2 q_1^2}{2(q-1)(q_1-1)(p-1)}\right\} \tag{XVI}$$

where p = Number of row-strip treatments, $a_1, a_2 \ldots a_p$

q_1 = Number of column-strip treatments, $b_1, b_2 \ldots b_{q1}$

q = Number of replicates or blocks

E_a = Mean square for row-strip plot error

E_b = Mean square for column-strip plot error

E_c = Mean square for ultimate sub-plot error

The last three formulæ are based on those given by Nair (1944) for the ordinary case not involving any missing plot.

More complicated cases involving a larger number of plot sizes in split-plot and strip designs have not been examined, but in such cases also the appropriate error variances may presumably be obtained by setting down the formula for variance for the ordinary case without a missing plot and adding to it a component for the contribution of the missing plot value to the treatment mean. The use of ordinary formulæ for standard errors for comparisons involving a missing plot would underestimate the standard error; but where the number of plots in the experiment is not too small the bias in the standard errors is not likely to affect the significance of differences.

I wish to express my thanks to Dr. V. G. Panse for his guidance in preparing the paper.

SUMMARY

Formulæ for calculating missing values in various split-plot and strip treatment designs are derived in this paper by the procedure of minimising the appropriate error variances. The calculation of the missing values is illustrated by numerical examples. The covariance method is employed for calculating the bias in the various sums of squares due to the use of missing plot value in the analysis of variance. Formulæ for standard errors for the various types of treatment comparisons, when one treatment contains a missing plot are given for the split-plot and strip trials.

REFERENCES

Anderson, R. L. .. *Biometrics Bull.*, 1946, **2**, 41.
Allan, F. E. and Wishart, J. *J. Agri. Sci.*, 1930, **20**, 399.
Bartlett, M. S. .. *Supp. J. Roy. Stat. Soc.*, 1937, **4**, 137.
Hutchinson, J. B. and Panse, V. G. *Ind. J. Agri. Sci.*, 1937, **7**, 531.
Nair, K. R. .. *Ibid.*, 1944, **14**, 315.
Yates, F. .. *Empire J. Expt. Agric.*, 1933, **1**, 129.