Journal of Agricultural Science, Cambridge (1998), **131**, 135–142. *Printed in United Kingdom*

Trojan square and incomplete Trojan square designs for crop research

R. N. EDMONDSON*

Biometrics Department, Horticulture Research International, Wellesbourne, Warwick, CV35 9EF, UK

(*Revised MS received 16 February 1998*)

SUMMARY

Latin square and near-Latin square designs are valuable row-and-column designs for crop research but the practical size range of such designs is severely limited. Semi-Latin square designs extend this range but not all semi-Latin squares are suitable for experimental designs. Trojan square designs are a special class of optimal semi-Latin squares that generalizes the class of Latin square designs. The construction of Trojan squares both for unstructured and for factorial treatment sets is discussed and the utility of Trojan square designs for practical crop research is demonstrated. The corpus of available designs is further extended by a discussion of incomplete Trojan square designs obtained by omitting one main row or one main column from a complete Trojan square design. Some advantages of Trojan square and incomplete Trojan square designs for crop research are discussed and some suggestions for further design research are made.

INTRODUCTION

In agricultural and horticultural research, row-and-column designs are of proven value and are widely used for the control of non-treatment variability in experiments both in the field and in the glasshouse. Such designs are particularly suitable where the experimental units are in a rectangular array and where two independent and mutually orthogonal sources of variation run across the rows and across the columns of the array. The simplest and perhaps the most effective row-and-column design for agricultural and horticultural research is the $n \times n$ Latin square in which n replicates of n treatments are arranged in n rows and n columns and in which every treatment occurs once in every row and once in every column. The treatment effects of a Latin square are orthogonal to both additive row effects and to additive column effects and, where a model with additive row and column effects can be assumed adequate, treatment effects can be estimated with full efficiency by simple averaging over replicates.

Latin squares are valuable for practical experiments where the number of treatments is small, but because the number of replicates must equal the number of treatments, the available range of designs is generally restricted to sizes from about 4×4 to about 7×7. The upper end of the size range can be extended by using incomplete Latin squares of size $(n-1) \times n$ or size $n \times (n-1)$ obtained by dropping one complete row or one complete column from a Latin square of size $n \times n$ (Yates 1936), whereas the lower end of the scale can be extended by using augmented Latin squares of size $(n+1) \times n$ or of size $n \times (n+1)$ obtained by repeating a complete row or a complete column of a Latin square of size $n \times n$ (Pearce 1952). By using such augmented and incomplete Latin squares, the useful practical range of designs of Latin square type can be extended to include from, say, three treatments in a 3×4 or 4×3 augmented design up to eight treatments in a 7×8 or 8×7 incomplete design. For many agricultural experiments, however, the number of treatments may be substantially larger than the number of replicates; then standard Latin square or near-orthogonal Latin square designs are not available and a more general class of row-and-column designs is required.

Semi-Latin squares are generalized $n \times n$ Latin squares for nk treatments where k is an integer greater than one. There are n rows and n columns and the intersection of each row and each column contains a block of k plots. Each row and each column contains a set of nk plots and every treatment occurs once in every row and once in every column. Semi-Latin squares are randomized by rows, by columns and by plots within blocks and are best regarded as doubly resolvable incomplete block designs.

In this paper, a special class of semi-Latin squares called Trojan squares will be discussed and it will be shown that, where available, Trojan squares are

* Email: Rodney.Edmondson@hri.ac.uk

normally the best choice of semi-Latin squares for crop research.

BACKGROUND

Historical

Semi-Latin squares have a long history in the statistical literature (see Preece & Freeman (1983) for an historical discussion) but they appear to have been little used in practice. One criticism of semi-Latin squares (Yates 1935) was that a simple row and column analysis does not take proper account of the multiple error structure of the design: certain treatment contrasts are split between two strata in the analysis, a blocks stratum and a plots-within-blocks stratum. The development of modern computer software has substantially removed the basis of this criticism by facilitating the proper stratified analysis of incomplete block designs, but a further difficulty is that not all semi-Latin squares are equally efficient for estimating treatment contrasts. Trojan squares are a special class of semi-Latin squares that have efficiency properties that make them particularly suitable for crop research experiments either in the field or in the glasshouse; they have been used at the Glasshouse Crops Research Institute, England, now part of Horticulture Research International, since 1958.

Trojan square literature

Trojan squares were first discussed in the literature by Harshbarger & Davis (1952) but they called their designs Latinized Near Balance Rectangular Lattices and restricted discussion to those designs having $k = n-1$. Later, Darby & Gilbert (1958) discussed the general case for $k < n$ and introduced the name Trojan square for designs where $k > 2$. However, all designs of the Latinized Rectangular Lattice type are now commonly described as Trojan squares for any $1 < k < n$. Preece & Freeman (1983) discussed the combinatorial properties of semi-Latin squares and related designs and Bailey (1988) discussed further constructions for a range of semi-Latin and Trojan square designs. Bailey (1992) discussed the efficiency of semi-Latin squares and compared the efficiency of the general semi-Latin square with that of the Trojan square. She showed that where all treatment comparisons are of equal importance, the Trojan square is the optimal choice of semi-Latin square for pairwise comparison of treatment means.

TROJAN SQUARE CONSTRUCTION

Mutually orthogonal Latin squares

Trojan squares of size $(n \times n)/k$ can be constructed by taking k mutually orthogonal Latin squares, each square using a different set of n letters, and superimposing them so that all the letters in the same row and column of the superimposed design form a single block of k plots. Every row and every column contains every letter from every square once whilst every pair of combinations of letters from different squares occurs together in the same block once. Table 1 shows an example of a non-randomized Trojan square design of size $(4 \times 4)/2$ constructed by superimposing two mutually orthogonal Latin square designs of size 4×4, one with upper case letters and the other with lower case letters. This design could be extended to a design of size $(4 \times 4)/3$ by superimposing an additional mutually orthogonal Latin square of size 4×4 but no further Trojan extension is possible, there being only three mutually orthogonal Latin squares of size 4×4.

Table 1. *A $(4 \times 4)/2$ Trojan square for eight treatments arranged in a 4×8 spatial array of plots*

	Columns …	1		2		3		4	
Rows	1	A	a	B	b	C	c	D	d
	2	B	c	A	d	D	a	C	b
	3	C	d	D	c	A	b	B	a
	4	D	b	C	a	B	d	A	c

Pseudo-factor structure

The structure of Trojan square designs is best discussed by using pseudo-factor methods (see Monod & Bailey 1992). Suppose that the k superimposed squares of a Trojan square design of size $(n \times n)/k$ are labelled by a k-level *Squares* factor and that the n letters of each square are labelled by an n-level *Letters* factor. The *Squares* and *Letters* factors are pseudo-factors and each combination of pseudo-factor levels can be used to label an actual treatment. Thus in Table 1 the eight pseudo-factor level combinations A, B, C, D, a, b, c, d represent eight actual treatments. The actual treatments themselves can be completely unstructured or they can have a $k \times n$ crossed or k/n nested factorial treatment structure. Unstructured treatments can be randomly assigned to pseudo-factor levels but structured treatments must be assigned according to the pseudo-factor structure of the design.

Factorial treatment structure

Suppose that the design in Table 1 is required to examine a 2^3 factorial treatment set where the treatment factors are, say, F, G and H and where the levels of the three factors are, respectively, f_1 and f_2, g_1 and g_2 and h_1 and h_2. There are two squares each with four letters and Table 2 shows an allocation of treatment factor levels that confounds the levels of F

https://doi.org/10.1017/S002185969800567X Published online by Cambridge University Press

Table 2. *Allocation of three 2-level factors* F, G *and* H *to the* Squares *and* Letters *labels of the Trojan square shown in Table 1*

		Letters			
		g_1h_1	g_1h_2	g_2h_1	g_2h_2
Squares	f_1	A	B	C	D
	f_2	a	b	c	d

with the levels of *Squares* and the levels of *G* and *H* with the levels of *Letters*. This allocation could also be used for a 2/4 nested factorial treatment set provided that *F* was the two-level treatment factor. In general, any $(n \times n)/k$ Trojan square has orthogonal factorial treatment structure if the k levels of the *Squares* factor coincide with the k levels of a real treatment factor or with the k levels of a product of real treatment factors.

TROJAN SQUARE ANALYSIS

Stratified analysis

A fully stratified analysis of variance for a Trojan square design contains four strata: a rows stratum, a columns stratum, a blocks stratum and a plots-within-blocks stratum. The plots-within-blocks stratum contains full information on the *Squares* pseudo-factor contrasts but information on the *Letters* pseudo-factor contrasts and the *Letters.Squares* pseudo-factor interaction contrasts is split between the plots-within-blocks stratum and the blocks stratum with relative information in the proportions $(k-1)/k$ to $1/k$ (see, for example, Bailey (1992)). Usually k will be small and the blocks stratum will contain a substantial amount of treatment information. For example, if $k = 2$ and the blocks and the plots-within-blocks stratum variances are equal, 50% of the *Letters* and *Letters.Squares* contrast information will be contained in the blocks stratum.

Pseudo-factor analysis

Trojan square designs have generalized pseudo-factor balance (see Payne & Tobias (1992) for a discussion of generalized balance), and all Trojan square designs can be analysed by making an orthogonal stratified analysis of variance of treatment contrasts into *Letters*, *Squares* and *Letters.Squares* pseudo-factor contrasts. For unstructured treatment sets, a pseudo-factor analysis is usually meaningless and is merely a device for ensuring a balanced partitioning of treatment effects. For factorial treatment sets, however, the analysis of the real treatment factor effects must coincide with the analysis of the pseudo-factor treatment effects. Table 3 shows an example of a factorial analysis of variance for the $(4 \times 4)/2$ Trojan square shown in Table 1 assuming that the treatment factor levels have been allocated according to Table 2. All factorial treatment effects, whether main effects or interaction effects, are estimated orthogonally and can be interpreted independently of each other. The main effect of factor *F* is estimated fully within the plots stratum but all other factorial effects are split between the blocks and the plots-within-blocks strata with relative information in the ratio 1:1.

Table 3. *Analysis of variance of pseudo-factorial treatment effects and actual factorial treatment effects for a* 2^3 *factorial treatment design in a* $(4 \times 4)/2$ *Trojan square*

Stratum	Pseudo-factor effects	Treatment factor effects	D.F.	Information
Rows	—	—	—	—
Columns	—	—	—	—
Blocks	Letters	G	1	0·5
		H	1	0·5
		G.H	1	0·5
	Letters.Squares	F.G	1	0·5
		F.H	1	0·5
		F.G.H.	1	0·5
Plots	Squares	F	1	1
	Letters	G	1	0·5
		H	1	0·5
		G.H	1	0·5
	Letters.Squares	F.G	1	0·5
		F.H	1	0·5
		F.G.H	1	0·5

Combination of treatment information

For efficient estimation of treatment effects, treatment information must be combined across strata. Appendix 1 shows a GENSTAT (1993) computer program suitable for the analysis of the design shown in Table 1. The pseudo-factor treatment structure is described by pseudo-factors labelled *Squares* and *Letters* whereas the real treatment structure is described by a factor labelled *Trtmnts*. The pseudo-factor operator // in the treatment structure formula

Trtmnts//(*Squares***Letters*)

suppresses the pseudo-factor analysis and provides an analysis of the *Trtmnts* factor only. A real factorial treatment structure could be analysed by omitting both the *Trtmnts* factor and the pseudo-factor operator from the treatment structure formula and by regarding the pseudo-factors as real factors. In practice, the *dummy* variate would be replaced by a real variate and the factor level values would be re-ordered according to the randomization of the design.

The [*print* = *aov*,*cbmeans*,*stratumvar*] option of ANOVA provides an analysis of variance table (*aov*), combined treatment means (*cbmeans*) and estimates of stratum variances (*stratumvar*). The combined treatment means are weighted combinations of treatment estimates from the different strata of the analysis with weights inversely proportional to stratum variances. The *cbmeans* option also provides combined standard errors of differences of means (S.E.D.S) for comparing combined treatment means. For unstructured treatment sets there are two S.E.D.S, one for comparing treatment means with the same level of the *Squares* factor and one for comparing treatment means with different levels of the *Squares* factor. Exact residual error degrees of freedom for combined S.E.D.S are not available but a working estimate can be based on the residual error degrees of freedom in the plots-within-blocks stratum. Appendix 2 shows general equations for S.E.D.S combined across strata and specific equations for S.E.D.S from the plots-within-blocks stratum only. For the example shown in Table 1, the ratio of the S.E.D. for comparing treatments with the same level of *Squares* to the S.E.D. for comparing treatments with different levels of *Squares* in the plots-within-blocks stratum is $1:\surd(7/8)$.

TROJAN SQUARE EXAMPLE

Block structure

Table 4 shows an analysis of variance of a $(4\times4)/3$ Trojan square experiment designed to compare 12 pest control treatments on apple trees. For cultural reasons, four long replicate rows, each one tree wide, were used with 12 plots per row. Each row was subdivided into four blocks of three plots and the four adjacent blocks at any one position along the four rows formed a replicate column of 12 plots. The 12 treatment comparisons were all equally important and a Trojan square was the most appropriate choice of resolvable row-and-column design for eliminating positional effects both between rows and between columns.

Table 4. *Analysis of insect pest count data from a $(4\times4)/3$ apple orchard trial using a square root transformation of the count data*

Stratum	Source	D.F.	Mean squares	E.D.F.	Stratum variances
Rows		3	0·0369	3	0·0369
Columns		3	0·9489	3	0·9489
Blocks				7·19	0·2066
	Standards	3	0·1448		
	Dursban	3	0·1602		
	Novosol	3	0·0398		
Plots				22·81	0·1040
	Squares	2	2·2484		
	Standards	3	0·9897		
	Dursban	3	1·0086		
	Novosol	3	0·0827		
	Residual	21	0·1040		

D.F. are conventional degrees of freedom.
E.D.F. are effective degrees of freedom.

Treatment analysis

The 12 treatments comprised four rates of each of two insecticide sprays, Dursban and Novosol, and four different types of standard treatments. The design was constructed by superimposing an additional mutually orthogonal Latin square on the design shown in Table 1 and by using the letters of the first square for the rates of Dursban, the letters of the second square for the rates of Novosol and the letters of the third square for the four standard treatments. Separate comparisons were required for the rates of Dursban, the rates of Novosol and the differences between the four standard treatments so a nested analysis of variance was made using a modified version of the program shown in Appendix 1. The *Rows*, *Columns* and *Squares* factors of the program were extended to accommodate the additional square of the design and three new factors, *Dursban*, *Novosol* and *Standards* were defined. The new factors each had five levels; the first four levels representing the actual levels of a particular factor and the fifth level representing all the remaining treatments. For example, the first four levels of *Dursban* represented the four Dursban treatments whereas the fifth level represented all the non-Dursban treatments. The nested analysis of variance was then defined by using the nested treatment structure formula

$$Squares/(Dursban+Novosol+Standards).$$

Interpretation

Table 4 shows an analysis of the mean squares and the stratum variances of the square root of the total insect pest count per plot. The stratum variances are the residual variances of each stratum after eliminating combined treatment effects and have effective numbers of error degrees of freedom that depend both upon the design and upon the stratum variances. For this reason, effective error degrees of freedom are real numbers not integers.

There was no evidence of significant differences between the rows of the design but there was significant evidence of column differences. These were due to a heavier infestation of pests occurring at one end of the trial than at the other end and the column blocks were successful in eliminating this source of variability from the blocks stratum analysis. The ratio of blocks to plots-within-blocks stratum variances was 1·986 whereas the ratio of blocks to plots-within-blocks stratum information for the split treatment contrasts was 0·5. Since the variance of a stratum treatment contrast is directly proportional to the stratum variance and inversely proportional to the stratum information, the blocks stratum treatment contrasts were about four times as variable as the corresponding plots-within-blocks stratum treatment contrasts. The plots-within-blocks stratum contained about 75% of the useful information on the *Dursban*, *Novosol* and *Standards* contrasts and, for this reason, only this stratum was used for significance testing. For treatment estimation, however, treatment information was combined across strata by using the combine means option of ANOVA and treatment estimates were compared by making pairwise comparisons using combined S.E.D.S. The combined S.E.D.S were assumed to have 21 residual error degrees of freedom.

INCOMPLETE TROJAN SQUARES

Construction

Complete Trojan squares of size $(n \times n)/k$ have n^2 blocks of size k and require n replicates of nk treatments. Sometimes design or cost constraints make complete Trojan squares impossible and then incomplete Trojan squares of size $((n-1) \times n)/k$ or of size $(n \times (n-1))/k$ can be useful. Such incomplete Trojan squares can be constructed by omitting any complete row or any complete column from any Trojan design of size $(n \times n)/k$. For a balanced pseudo-factor analysis of *Letters* and *Squares* effects, however, each omitted block must contain only a single level of the *Letters* factor. By using standard Trojan squares constructed from standard orthogonal Latin squares in which the arrangement of the *Letters* factor is the same for the first row of each square, incomplete Trojan squares with pseudo-factor balance can be obtained by omitting the first row of any standard Trojan square. For example, omitting the first row of the $(4 \times 4)/2$ standard Trojan square shown in Table 1 gives the incomplete Trojan square of size (3x4)/2 shown in Table 5.

Table 5. *A $(3 \times 4)/2$ incomplete Trojan square for eight treatments arranged in a 3×8 spatial array of plots*

	Columns …	1		2		3		4	
Rows	1	B	c	A	d	D	a	C	b
	2	C	d	D	c	A	b	B	a
	3	D	b	C	a	B	d	A	c

Analysis

Incomplete Trojan squares with pseudo-factor balance have an orthogonal partitioning of treatment contrasts into *Letters*, *Squares* and *Letters.Squares* pseudo-factor contrasts. Treatment information is split between three strata but, unlike complete Trojan squares, the *Letters* and *Letters.Squares* pseudo-factor contrasts have different efficiency factors. Appendix 3 gives the efficiency factors for each set of pseudo-factor contrasts in each stratum of an incomplete Trojan square analysis. For factorial treatment sets, *Letters* and *Squares* represent real treatment factors and all factorial treatment effects and interaction effects are estimated orthogonally and can be interpreted independently of each other. For unstructured treatment sets, treatments are best compared by making pairwise comparisons between treatment means using appropriate S.E.D.S.

Incomplete Trojan squares have three S.E.D.S, one for comparing means with the same level of *Letters*, one for comparing means with the same level of *Squares*, and one for comparing means with different levels of both *Letters* and *Squares*. Appendix 2 shows general equations for the three S.E.D.S combined across strata and specific equations for the three S.E.D.S estimated from the plots-within-blocks stratum only. For the example shown in Table 5, the ratios of the S.E.D.S for comparing treatments with the same level of *Squares*, treatments with the same level of *Letters* and treatments with different levels of both *Letters* and *Squares* in the plots-within-blocks stratum of the analysis are (3/2):(5/3):(7/6). Normally, recovery of inter-block information will reduce the inequality of the S.E.D.S; only when the blocks stratum variance is very large relative to the plots-within-blocks stratum variance will the unreduced S.E.D.S be used. Then, however, the overall increase in precision of the intra-block analysis due to the block design should more than compensate for the inequality of the S.E.D.S. As with complete Trojan squares, exact residual error degrees of freedom are not available for combined S.E.D.S but a working estimate can be based on the residual error degrees of freedom in the plots-within-blocks stratum of the analysis.

Prime-power level factorials

For certain factorial treatment structures where the treatments include combinations of factors with the same prime number of levels, alternative confounding schemes are possible. Suppose, for example, that a $(3 \times 4)/2$ incomplete Trojan square is required for the three 2-level factors F, G and H shown in Table 2. From Table 3, the treatment effects confounded with the *Letters* contrasts are the main effects of G and H and the $G.H$ interaction effect. With this factor allocation, the main effects of G and H, as well as the

https://doi.org/10.1017/S002185969800567X Published online by Cambridge University Press

$G.H$ interaction, will be partially confounded with columns. A better alternative confounding scheme might be to partially confound only the two-factor interactions $F.G$, $F.H$ and $G.H$ with columns. For the square shown in Table 1, this could be done by allocating treatment factor levels according to Table 2 but with the bottom line of Table 2 re-ordered to read *d*, *c*, *b*, *a* instead of *a*, *b*, *c*, *d*. Kempthorne (1952) gives a full discussion of confounding schemes available for prime-power level factorials.

INCOMPLETE TROJAN SQUARE EXAMPLE

Block structure

Table 6 shows an analysis of variance of a $(3 \times 4)/2$ incomplete Trojan square experiment designed to investigate eight nutrient regimes for glasshouse tomato. There were eight beds of plants and these were divided into four pairs of adjacent beds to form the four columns of the design. Each bed was subdivided into three treatment plots and the three sets of eight adjacent plots running across the beds of the design were used as the rows of the design. The treatment contrasts in the blocks stratum of the design were orthogonal to row effects and were near-orthogonal to column effects. A complete Trojan square of size $(4 \times 4)/2$ would have given complete orthogonality but, for practical reasons, it was impossible to subdivide each bed into more than three plots.

Treatment structure and analysis

The eight treatments had a crossed factorial treatment structure with two rates of aeration and four rates of nutrient feed. A design based on Table 5 was used with upper and lower case letters representing high and low rates of aeration respectively. Letters represented different rates of nutrient feed but the same letter, whether upper or lower case, always represented the same rate of nutrient feed. There were three complete replicates of each treatment combination and each pair of combinations of different rates of aeration with different rates of nutrition occurred together in the same block once. The design was analysed by the program in Appendix 1 with the *Rows*, *Columns*, *Squares* and *Letters* factors modified to account for the omitted row. The *Squares* and *Letters* factors were re-named *Aeration* and *Feed*, respectively, and the *Trtmnts* factor and the pseudo-factor operator // were omitted from the program. The resulting analysis is shown in Table 6.

Table 6. *Analysis of total tomato yields in* kg/m^2 *from a* $(3 \times 4)/2$ *glasshouse tomato experiment*

Stratum	Source	D.F.	Mean squares	E.D.F.	Stratum variances
Rows		2	118·47	2	118·47
Columns				2·86	16·23
	Feed	3	13·04		
Blocks				2·65	4·46
	Feed	3	4·70		
	Aer.Feed	3	9·28		
Plots				8·49	7·59
	Aeration	1	0·61		
	Feed	3	14·93		
	Aer.Feed	3	4·16		
	Residual	5	5·95		

D.F. are conventional degrees of freedom.
E.D.F. are effective degrees of freedom.

Interpretation

Table 6 shows an analysis of variance of the mean squares and the stratum variances of the total tomato yield in kg/m^2. There were very large differences between rows and there was some evidence of smaller differences between columns so both the row and the column classification improved the precision of estimation in the blocks stratum. The blocks stratum variance estimate was smaller than the plots-within-blocks stratum variance estimate, but prior experience from other experiments indicated that this was unlikely to be a real effect. There was substantial treatment information in the blocks stratum and, as there was no evidence that the blocks stratum variance was inflated relative to the plots stratum variance, treatment information was combined across strata by merging the bottom two strata into a single stratum. This was done by substituting *Rows* + *Columns* for *Rows*∗*Columns* in the block structure formula in Appendix 1. Treatment estimates and significance tests were then based on the bottom stratum of the analysis only. The nutrient feed information contained in the columns stratum was relatively small and since it was likely that differences between columns were confounded with systematic positional trend effects, treatment information in the columns stratum was ignored.

DISCUSSION

Total inherent variability in crop experiments can be substantial and proper blocking and replication are essential for reliable and efficient experimentation. Trojan squares generalize Latin squares and provide designs with two mutually orthogonal blocking systems for experiments where the number of treatments is a multiple of the number of replicates. Trojan squares are valuable for field and glasshouse crop experiments where mutually orthogonal row and column block effects are anticipated but they can also be valuable in place of complete randomized blocks

Table 7. *The partition of pseudo-factor treatment information from an incomplete Trojan square of size* $(n\times(n-1))/k$ *or size* $((n-1)\times n)/k$

Stratum	Source	Information
Rows (or columns)	Letters	$1/(n-1)^2$
Blocks	Letters	$n(n-k-1)/(k(n-1)^2)$
	Letters. Squares	$n/(k(n-1))$
Plots	Squares	1
	Letters	$n(k-1)/(k(n-1))$
	Letters. Squares	$(kn-n-k)/(k(n-1))$

designs. Often the best choice of blocking system is unknown prior to an experiment and then a double blocking system can give extra protection against a wrong choice of blocks in the layout of the experiment.

The examples show that routine design and analysis of Trojan and incomplete Trojan square experiments is now feasible. However, the underlying assumptions of a design and analysis must always be checked for validity. For example, in the glasshouse experiment, prior knowledge indicated that the plots-within-blocks stratum variance was unlikely to be less than the blocks stratum variance whereas the experiment itself indicated that the blocks stratum variance was not greater than the plots-within-blocks stratum variance. These two pieces of information, taken together, justified merging the bottom two strata into a single stratum. In other circumstances, it might have been necessary to investigate and, perhaps, model the causes of variability in the different strata of the analysis. The assumption of additivity of row and column effects in a row-and-column design should also be checked. Sometimes there may be significant low-order polynomial interaction effects between rows and columns and then a polynomial response surface analysis may be appropriate. Edmondson (1993) has discussed systematic Graeco-Latin squares balanced for low-order polynomial row and column interaction effects and these designs are suitable for systematic $(n\times n)/2$ Trojan square designs whenever smooth low-order trend effects are anticipated. Further work on systematic $(n\times n)/k$ Trojan squares for $k>2$ is needed and work on systematic incomplete Trojan designs for $k\geqslant 2$ would also be of value.

Finally, more work is needed to extend the concept of Trojan and incomplete Trojan squares to include a wider range of design sizes. For example, Preece (1966) has discussed balanced row-and-column designs for sets of superimposed treatments where each set of treatments is arranged in a Youden square. These designs generalize incomplete Trojan designs of size $((n-1)\times n)/k$ to designs of size $(m\times n)/k$ where m and n are the dimensions of a suitable Youden square and $k>1$. Augmented Trojan squares of size $((n+1)\times n)/k$ can be useful when n is small and can be constructed by repeating the first row of any standard Trojan square of size $(n\times n)/k$.

The author is grateful to S. Hammond for help with an earlier version of this paper, to J. Cross and the Apple and Pear Research Council for permission to use the apple orchard data and to I. Burns and MAFF for permission to use the glasshouse tomato data. The author is also grateful to D. A. Preece and H. Monod for many constructive comments.

REFERENCES

Bailey, R. A. (1988). Semi-Latin squares. *Journal of Statistical Planning and Inference* **18**, 299–312.

Bailey, R. A. (1992). Efficient semi-Latin squares. *Statistica Sinica* **2**, 413–437.

Darby, L. A. & Gilbert, N. (1958). The Trojan square. *Euphytica* **7**, 183–188.

Edmondson, R. N. (1993). Systematic row-and-column designs balanced for low-order polynomial interactions between rows and columns. *Journal of the Royal Statistical Society B* **55**, 707–723.

Genstat 5 Committee (1993). *Genstat 5 Release 3 Reference Manual*. Oxford: Clarendon Press.

Harshbarger, B. & Davis, L. L. (1952). Latinized Rectangular Lattices. *Biometrics* **8**, 73–84.

Kempthorne, O. (1952). *The Design and Analysis of Experiments*. New York: Wiley

Monod, H. & Bailey, R. A. (1992). Pseudofactors: normal use to improve design and facilitate analysis. *Applied Statistics* **41**, 317–336

Payne, R. W. & Tobias, R. D. (1992). General balance, combination of information and the analysis of covariance. *Scandinavian Journal of Statistics* **19**, 3–23.

Pearce, S. C. (1952). Some new designs of Latin square type. *Journal of the Royal Statistical Society Series B* **14**, 101–106.

Preece, D. A. (1966). Some row and column designs for two sets of treatments. *Biometrics* **22**, 1–25.

Preece, D. A. & Freeman, G. H. (1983). Semi-Latin squares and related designs. *Journal of the Royal Statistical Society Series B* **45**, 267–277.

Yates, F. (1935). Complex experiments (with Discussion). *Supplement of the Journal of the Royal Statistical Society* **2**, 181–247.

Yates, F. (1936). Incomplete Latin squares. *Journal of Agricultural Science* **26**, 301–315.

APPENDIX 1

```
units[32]
factor[levels = 4;values = 8(1 ... 4)] Rows
factor[levels = 4;values = 2(1 ... 4) 4] Columns
factor[levels = 2;values = (1,2)16] Squares
factor[levels = 4;values = 1, 1, 2, 2, 3, 3, 4, 4,\
                           2, 3, 1, 4, 4, 1, 3, 2,\
                           3, 4, 4, 3, 1, 2, 2, 1,\
                           4, 2, 3, 1, 2, 4, 1, 3] Letters
factor[levels = 8;values = 1, 5, 2, 6, 3, 7, 4, 8,\
                           2, 7, 1, 8, 4, 5, 3, 6,\
                           3, 8, 4, 7, 1, 6, 2, 5,\
                           4, 6, 3, 5, 2, 8, 1, 7] Trtmnts
block Rows*Columns
treat Trtmnts//(Squares*Letters)
anova[print = aov,cbmeans,stratumvar] dummy
stop
```

APPENDIX 2

The variance matrix for a k-level *Squares* factor crossed with an n-level *Letters* factor in a design with orthogonal factorial structure is:

$$\mathbf{V} = (\mathrm{v_{SS}}(k\mathbf{I}-\mathbf{J})\otimes\mathbf{J}+\mathrm{v_{LL}}\mathbf{J}\otimes(n\mathbf{I}-\mathbf{J}) + \mathrm{v_{SL}}(k\mathbf{I}-\mathbf{J})\otimes(n\mathbf{I}-\mathbf{J}))/(kn)$$

where $\mathrm{v_{SS}}$, $\mathrm{v_{LL}}$ and $\mathrm{v_{SL}}$ are the variances of the *Squares*, *Letters* and *Squares*.*Letters* contrasts respectively, **I** is an identity matrix, **J** is a matrix of unit elements and the Kroenecker product symbol $\otimes$ has $k\times k$ matrices on the left and $n\times n$ matrices on the right. General equations for S.E.D.s for differences between pairs of means with the same levels of *Squares*, pairs of means with the same levels of *Letters*, and pairs of means with different levels of both *Squares* and *Letters*, are, respectively

$$\text{S.E.D.}_{\text{SS}} = \sqrt{(2(\mathrm{v_{LL}}+(k-1)\mathrm{v_{SL}})/k)}$$
$$\text{S.E.D.}_{\text{LL}} = \sqrt{(2(\mathrm{v_{SS}}+(n-1)\ \mathrm{v_{SL}})/n)}$$
$$\text{S.E.D.}_{\text{SL}} = \sqrt{(2(k\mathrm{v_{SS}}+n\mathrm{v_{LL}}+(kn-k-n)\ \mathrm{v_{SL}})/(kn))}$$

Combined treatment estimates from Trojan and incomplete Trojan designs are functions of the stratum variances and the relative precision of the various treatment comparisons will depend upon the estimated stratum variances. The precision of the plots-within-blocks treatment comparisons, however, depend only on the plots-within-blocks stratum variance, σ^2, and the relative precision of these comparisons is constant. The variances $\mathrm{v_{LL}}$, $\mathrm{v_{SS}}$ and $\mathrm{v_{SL}}$ in any particular stratum are inversely proportional to the corresponding stratum efficiency factors therefore, using efficiency factors of 1 for $\mathrm{v_{SS}}$ and $(k-1)/k$ for $\mathrm{v_{LL}}$ and $\mathrm{v_{SL}}$, the plots-within-blocks S.E.D.s for an $(n\times n)/k$ Trojan square are:

$$\text{S.E.D.}_{\text{SS}} = \sigma\sqrt{(2\text{k}/(n(k-1)))}$$
$$\text{S.E.D.}_{\text{LL}} = \sigma\sqrt{(2(nk-1)/(n^2(k-1)))}$$
$$\text{S.E.D.}_{\text{SL}} = \sigma\sqrt{(2(nk-1)/(n^2(k-1)))}$$

Similarly, using appropriate plots-within-blocks treatment efficiency factors from Table 7, the plots-within-blocks S.E.D.s for an $((n-1)\times n)/k$ or an $(n\times(n-1))/k$ incomplete Trojan square are:

$$\text{S.E.D.}_{\text{SS}} = \sigma\sqrt{(2k(nk-n-1)/(n(k-1)(nk-n-k)))}$$
$$\text{S.E.D.}_{\text{LL}} = \sigma\sqrt{(2(nk-k-1)/((n-1)(nk-n-\text{k})))}$$
$$\text{S.E.D.}_{\text{SL}} = \sigma\sqrt{(2(nk-1)/(n(n-1)(k-1)))}$$

APPENDIX 3

For an incomplete Trojan square one complete set of *Letters* contrasts is confounded between incomplete rows or columns. Since there are $k(n-1)$ plots in each incomplete row or column and k plots in each omitted block the efficiency of estimation of *Letters* contrasts in the incomplete rows or columns stratum relative to a complete blocks design is $1/(n-1)^2$. The *Letters* contrasts in the plots-within-blocks stratum and the *Letters*.*Squares* contrasts in the blocks stratum are identical with the corresponding contrasts in the corresponding strata of a complete Trojan square. Assuming corresponding stratum variances are unchanged, the variances of the contrasts are also unchanged. However, the size of the incomplete Trojan design is reduced by the factor $(n-1)/n$ therefore, relative to the corresponding complete Trojan square, the efficiencies of estimation of *Letters* contrasts in the plots-within-blocks stratum and *Letters*.*Squares* contrasts in the blocks stratum are increased by the factor $n/(n-1)$. In orthogonal designs efficiencies must sum to unity therefore all remaining efficiencies can be found by subtraction. Table 7 summarizes the full set of pseudo-factor efficiency factors for each stratum of an incomplete Trojan square with pseudo-factor balance.