Listy Biometryczne – Biometrical Letters
Vol. 31 (1994), No. 1, 3-16

Some split-plot designs in series of experiments

Iwona Mejza

Agricultural University of Poznań, Wojska Polskiego 28,
PL-60-637 Poznań, Poland

Summary

The paper deals with the planning and analysis of a series of experiments carried out in some incomplete split-plot designs.

There are experiments in which treatments occur on many levels. Usually, on account of the limited structure of experimental material and because of costs, it is not possible to use a complete design. In such a case some incomplete split-plot design may be very useful.

The aim of the paper is to present the problems connected with the planning and analysis of a series of experiments (repeated over many environments), where every single experiment is of a split-plot type.

The considered single designs can be incomplete with regard to the whole-plot treatments or with regard to the sub-plot treatments.

In the paper we adapt the linear model called the randomization model. Our model is based on the three step randomization performed in every environment, i.e. randomization of blocks, randomization of whole-plots within each block, and randomization of sub-plots within each whole-plot of each block. We assume that the environments are not randomized. Additionally, some assumptions concerning additivity and statistical properties of the so called technical errors are also adopted.

The incomplete split-plot designs are assumed to have an orthogonal block structure. Hence, the appropriate analysis of multistrata experiments is adapted.

In the last part of the paper, the statistical properties of some incomplete split-plot designs generated by certain incomplete block designs are examined.

1. Introduction

Let us consider a two-factor experiment of the split-plot type in which the first factor, A, (whole-unit treatments) occurs on S levels $A_1, A_2, \ldots, A_S$, and the second factor, B, (sub-unit treatments) occurs on T levels $B_1, B_2, \ldots, B_T$.

Key words: split-plot design, series of experiments, incompleteness, factorial structure, BIBD, PBIBD(2), efficiency factors

Let a population of units (set of potential units) in an environment be divided into b blocks and let each block be additionally divided into k whole-units while each whole-unit is divided into t sub-units.

We assume that in each environment P_g, the experimental material structure is the same, $g = 1,2,\ldots,d$, where d denotes the number of environments. It means that in each environment we have $n_0 = bkt$ units.

The proposed design finds many applications in agricultural field experiments, then it is convenient to call units as plots. Such a convention will be used in the paper.

2. Linear model and its analysis

Let $D = \{D_1, D_2,\ldots, D_d\}$ be the theoretical design of whole experiment. The sub designs $D_1, D_2,\ldots, D_d$ may be the same or different in environments but they have to use the same structure of an experimental material.

In the paper we use linear model called randomization model (cf. Nelder, 1965; Mejza, 1987). Our model is based on the three step randomization, i.e., randomization of blocks, randomization of whole-plots within each block, randomization of sub-plots within each whole-plot of each block, performed independently in all environments. We assume that the environments are not randomized. It means that all the treatment effects, including the environment effect, are considered to be fixed.

The linear model of the yield obtained in the split-plot design in the g-th environment can be written in the form (cf. Mejza and Mejza, 1994):

$$\mathbf{Y}_g = \mathbf{\Delta}'_g\boldsymbol{\tau}_g + \mathbf{D}'_g\boldsymbol{\rho}_g + \mathbf{G}'_g\boldsymbol{\eta}_g + \varepsilon_g + \mathbf{e}_g\,, \tag{2.1}$$

where Y_g is an $n_0\times 1$ vector of yields, $\mathbf{\Delta}'_g$ is an $n_0\times ST$ design matrix for treatment combinations, $\boldsymbol{\tau}_g$ is an $ST\times 1$ vector of treatment combinations effects, $\mathbf{D}'_g$ is an $n_0\times b$ design matrix for blocks, $\boldsymbol{\rho}_g$ is a $b\times 1$ vector of block effects, $\mathbf{G}'_g$ is an $n_0\times bk$ design matrix for whole-plots, $\boldsymbol{\eta}_g$ is a $bk\times 1$ vector of whole-plot errors, ε_g and $\mathbf{e}_g$ are $n_0\times 1$ vectors of the sub-plot and technical errors, respectively, $g = 1,2,\ldots,d$.

The structure of the dispersion matrix of model (2.1) resulting from the randomization described above, has the form:

$$\mathrm{Cov}(\mathbf{Y}_g) = \mathbf{V}_g = \mathbf{D}'_g\mathbf{V}_{g\rho}\,\mathbf{D}_g + \mathbf{G}'_g\mathbf{V}_{g\eta}\,\mathbf{G}_g + \mathbf{V}_{g\varepsilon} + \sigma^2_{ge}\mathbf{I}_{n_0}\,, \tag{2.2}$$

where $\mathbf{V}_{g\rho} = \mathrm{Cov}(\boldsymbol{\rho}_g) = \sigma^2_{g\rho}(\mathbf{I}_b - b^{-1}\mathbf{J}_b)$, $\mathbf{V}_{g\eta} = \mathrm{Cov}(\boldsymbol{\eta}_g) = \sigma^2_{g\eta}\mathbf{I}_b\otimes(\mathbf{I}_k - k^{-1}\mathbf{J}_k)$, $\mathbf{V}_{g\varepsilon} = \mathrm{Cov}(\varepsilon_g) = \sigma^2_{g\varepsilon}\mathbf{I}_{bk}\otimes(\mathbf{I}_t - t^{-1}\mathbf{J}_t)$, $\sigma^2_{g\rho}$ and $\sigma^2_{g\eta}$ denote the block variance and the whole-plot error variance, respectively, $\sigma^2_{g\varepsilon}$, σ^2_{ge} denote the error (sub-plot) variance and the technical error variance, respectively, for $g = 1,2,\ldots,d$, $\otimes$ stands for the Kronecker product of matrices, $\mathbf{I}_x$ denotes the identity matrix of order x and $\mathbf{J}_x = \mathbf{1}_x\mathbf{1}'_x$, with $\mathbf{1}_x$ denoting the vector of ones. All covariances are equal to zero. It results from the independent randomization and from assumptions concerning technical error term.

The problem is how to extend model (2.1) to the case of experiments conducted in d environments. We need some additional assumptions. Let us note that the structure of experimental material is assumed to be the same in the environments and the same scheme of randomization is assumed to be performed at every environment. Hence, we may assume in the paper that the dispersion structure of random terms in model (2.1) is the same in all environments, i.e. $\sigma^2_{g\rho} = \sigma^2_{\rho}$, $\sigma^2_{g\eta} = \sigma^2_{\eta}$, $\sigma^2_{g\varepsilon} = \sigma^2_{\varepsilon}$, $\sigma^2_{ge} = \sigma^2_{e}$, for every $g = 1,2,\ldots,d$. Moreover, on account of the same structure of experimental material in each environment, we have $\mathbf{D}'_g = \mathbf{D}'$, $\mathbf{G}'_g = \mathbf{G}'$, $g=1,2,\ldots,d$. It is assumed that the environments influence only the yield expected value.

In the paper, by the treatment we will mean the treatment combinations $P_g A_w B_j$, $g = 1,2,\ldots,d$, $w = 1,2,\ldots,S$, $j = 1,2,\ldots,T$, while by the effect of the treatment we will mean:

$$\tau_{gwj} = \mu + \pi_g + \alpha_w + \beta_j + (\pi\alpha)_{gw} + (\pi\beta)_{gj} + (\alpha\beta)_{wj} + (\pi\alpha\beta)_{gwj}\,, \tag{2.3}$$

$$g = 1,2,\ldots,d,\ w = 1,2,\ldots,S,\ j = 1,2,\ldots,T\,,$$

where μ denotes the general parameter, π_g denotes the effect of the g-th environment P_g, α_w denotes the effect of the w-th level of factor A, β_j denotes the effect of the j-th level of factor B and $(\pi\alpha)_{gw}$, $(\pi\beta)_{gj}$, $(\alpha\beta)_{wj}$, $(\pi\alpha\beta)_{gwj}$ stand for interaction effects.

Let $v=dST$ denote the number of treatments (three-way treatment combinations: environments $\times$ whole-plot treatments $\times$ sub-plot treatments).

Using all assumptions described above, the final linear model for observed yield can be written as:

$$\mathbf{y} = \mathbf{\Delta}'\boldsymbol{\tau} + \mathbf{D}^{*\prime}\boldsymbol{\rho}^* + \mathbf{G}^{*\prime}\boldsymbol{\eta}^* + \boldsymbol{\varepsilon}^* + \mathbf{e}^*\,, \tag{2.4}$$

where $\mathbf{y}$ is an $n\times 1$, $n = dn_0$, vector of observations, $\mathbf{y} = (y'_1, y'_2,\ldots,y'_d)'$, $\mathbf{\Delta}'$ is an $n\times v$ design matrix for treatments, $\boldsymbol{\tau}$ is a $v\times 1$ vector of treatment parameters, $\mathbf{D}^{*\prime}$ is an $n\times bd$ design matrix for blocks, $\boldsymbol{\rho}^*$ is a $db\times 1$ vector of block effects, $\mathbf{G}^{*\prime}$ is an $n\times bkd$ design matrix for whole-plots, $\boldsymbol{\eta}^*$ is a $dbk\times 1$ vector of whole-plot errors, $\boldsymbol{\varepsilon}^*$ and $\mathbf{e}^*$ are $n\times 1$ vectors of the sub-plot and technical errors, respectively.

The dispersion matrix of model (2.4) has the form

$$\mathrm{Cov}(\mathbf{y}) = \mathbf{V} = \mathbf{D}^{*\prime}\mathbf{V}_\rho^*\mathbf{D}^* + \mathbf{G}^{*\prime}\mathbf{V}_\eta^*\mathbf{G}^* + \mathbf{V}_\varepsilon^* + \sigma_e^2\mathbf{I}\ , \tag{2.5}$$

where $\mathbf{V}_\rho^* = \mathbf{I}_d \otimes \mathbf{V}_\rho$, $\mathbf{V}_\eta^* = \mathbf{I}_d \otimes \mathbf{V}_\eta$, $\mathbf{V}_\varepsilon^* = \mathbf{I}_d \otimes \mathbf{V}_e$, $\mathbf{D}^{*\prime} = \mathbf{I}_d \otimes \mathbf{D}'$, $\mathbf{G}^{*\prime} = \mathbf{I}_d \otimes \mathbf{G}'$.

The approach proposed by Nelder (1965) to the analysis of multistrata experiments possessing orthogonal block structure will be adopted. In the considered case (cf. Mejza, 1987) the set of matrices (projectors):

$$\mathbf{P}_0^* = \mathbf{I}_d \otimes \mathbf{P}_0,\quad \mathbf{P}_0 = n_0^{-1}\mathbf{J}_{n_0}, \qquad r(\mathbf{P}_0^*) = d,$$

$$\mathbf{P}_1^* = \mathbf{I}_d \otimes \mathbf{P}_1,\quad \mathbf{P}_1 = (kt)^{-1}\mathbf{D}'\mathbf{D} - n_0^{-1}\mathbf{J}_{n_0}, \qquad r(\mathbf{P}_1^*) = d(b-1),$$

$$\mathbf{P}_2^* = \mathbf{I}_d \otimes \mathbf{P}_2,\quad \mathbf{P}_2 = t^{-1}\mathbf{G}'\mathbf{G} - (kt)^{-1}\mathbf{D}'\mathbf{D}, \qquad r(\mathbf{P}_2^*) = db(k-1),$$

$$\mathbf{P}_3^* = \mathbf{I}_d \otimes \mathbf{P}_3,\quad \mathbf{P}_3 = \mathbf{I}_n - t^{-1}\mathbf{G}'\mathbf{G}, \qquad r(\mathbf{P}_3^*) = dbk(t-1),$$

plays an important role. These matrices are idempotent, pairwise orthogonal and their sum is the identity matrix. Moreover, $\mathbf{P}_f\mathbf{1}_n = 0$, $f = 1,2,3$.

Dispersion matrix (2.5) can be written as: $\mathbf{V} = \gamma_0\mathbf{P}_0^* + \gamma_1\mathbf{P}_1^* + \gamma_2\mathbf{P}_2^* + \gamma_3\mathbf{P}_3^*$, where $\gamma_0 = \sigma_e^2$, $\gamma_1 = tk\sigma_\rho^2 + \sigma_e^2$, $\gamma_2 = t\sigma_\eta^2 + \sigma_e^2$, $\gamma_3 = \sigma_\varepsilon^2 + \sigma_e^2$.

The overall analysis of the model (2.4) can be divided into the so called strata analyses, i.e. the analyses which are based on strata's models:

$$\mathbf{y}_f = \mathbf{P}_f^*\mathbf{y},\quad \mathrm{E}(\mathbf{y}_f) = \mathbf{P}_f^*\boldsymbol{\Delta}'\mathbf{y},\quad \mathrm{Cov}(\mathbf{y}_f) = \gamma_f\mathbf{P}_f^*,\quad f = 0,1,2,3\ . \tag{2.6}$$

(For details of the above procedure see Nelder, 1965; Houtman and Speed, 1983).

Applying the least squares method to models (2.3), the normal equations for estimation of τ are of the form:

$$\mathbf{C}_f\boldsymbol{\tau}_f^o = \mathbf{Q}_f\ , \tag{2.7}$$

where $\mathbf{C}_f = \boldsymbol{\Delta}\mathbf{P}_f^*\boldsymbol{\Delta}'$, $\mathbf{Q}_f = \boldsymbol{\Delta}\mathbf{P}_f^*\mathbf{y}$, $f = 0,1,2,3$.

Let $\mathbf{N}_1 = \boldsymbol{\Delta}\mathbf{D}^{*\prime}$ and $\mathbf{N}_2 = \boldsymbol{\Delta}\mathbf{G}^{*\prime}$ be treatments *vs.* blocks and treatments *vs.* whole-plots incidence matrices. Let us note that $\boldsymbol{\Delta}' = \mathrm{diag}(\boldsymbol{\Delta}_1', \boldsymbol{\Delta}_2', \ldots, \boldsymbol{\Delta}_d')$, where $\boldsymbol{\Delta}_g'$ is the design matrix for treatments in the g-th environment, $g = 1,2,\ldots,d$.

Let us consider the treatment structure of the design with respect to estimability property of some linear treatment functions in each stratum.

In the zero stratum we have:

$$\mathbf{C}_0 = n_0^{-1}\mathrm{diag}(\mathbf{r}_1\mathbf{r}_1', \mathbf{r}_2\mathbf{r}_2', \ldots, \mathbf{r}_d\mathbf{r}_d'),\quad \mathbf{r}_g = \boldsymbol{\Delta}_g\mathbf{1}_{n_0},\quad \mathbf{C}_0^- = n_0^{-1}\mathbf{I}_d \otimes \mathbf{J}_{ST}\ ,$$

$$\boldsymbol{\tau}_0^o = \mathbf{C}_0^-\boldsymbol{\Delta}\mathbf{P}_0\mathbf{y} = [\bar{y}_1\mathbf{1}', \bar{y}_2\mathbf{1}', \ldots, \bar{y}_d\mathbf{1}']',\quad \bar{y}_g = n_0^{-1}\mathbf{1}'\mathbf{y}_g,\qquad g = 1,2,\ldots,d.$$

In the first stratum we have:

$$\mathbf{C}_1 = \mathrm{diag}(\mathbf{C}_{11},\mathbf{C}_{12},\ldots,\mathbf{C}_{1d}), \quad \text{where} \quad \mathbf{C}_{1g} = (kt)^{-1}\mathbf{N}_{1g}\mathbf{N}'_{1g} - n_0^{-1}\mathbf{r}_g\mathbf{r}'_g, \tag{2.8}$$

and

$$\mathbf{N}_{1g} = \boldsymbol{\Delta}_g\mathbf{D}', \ \mathbf{C}_{1g}\mathbf{1} = \mathbf{0}, \qquad g = 1,2,\ldots,d.$$

In the second stratum we have:

$$\mathbf{C}_2 = \mathrm{diag}(\mathbf{C}_{21},\mathbf{C}_{22},\ldots,\mathbf{C}_{2d}), \quad \text{where} \quad \mathbf{C}_{2g} = t^{-1}\mathbf{N}_{2g}\mathbf{N}'_{2g} - (kt)^{-1}\mathbf{N}_{1g}\mathbf{N}'_{1g}, \tag{2.9}$$

$$\mathbf{N}_{2g} = \boldsymbol{\Delta}_g\mathbf{G}', \quad \mathbf{C}_{2g}\mathbf{1} = \mathbf{0}, \qquad g = 1,2,\ldots,d.$$

In the third stratum we have:

$$\mathbf{C}_3 = \mathrm{diag}(\mathbf{C}_{31},\mathbf{C}_{32},\ldots,\mathbf{C}_{3d}), \ \text{where}\ \mathbf{C}_{3g} = \mathbf{r}_g^{\delta} - t^{-1}\mathbf{N}_{2g}\mathbf{N}'_{2g}, \tag{2.10}$$

and

$$\mathbf{C}_{3g} = \mathrm{diag}(\mathbf{C}_{3g1},\mathbf{C}_{3g2},\ldots,\mathbf{C}_{3gS}), \quad \mathbf{C}_{3gw} = \mathbf{r}_{gw}^{\delta} - t^{-1}\mathbf{N}_{2gw}\mathbf{N}'_{2gw},$$

$$g = 1,2,\ldots,d, \ w = 1,2,\ldots,S, \ \mathbf{r}^{\delta} = \mathrm{diag}(r_1,r_2,\ldots,r_v).$$

The estimability of a linear treatment function $\mathbf{c}'\boldsymbol{\tau}$ within the f-th stratum can be verified by the criterion $\mathbf{c}'\mathbf{C}_f^{-}\mathbf{C}_f = \mathbf{c}'$ (cf. Rao and Mitra, 1971, Theorem 7.2.1).

From the property $\mathbf{C}_f\mathbf{1} = \mathbf{0}$, $f = 1,2,3$, it results that if the function $\mathbf{c}'\boldsymbol{\tau}$ is estimable it must be a contrast, i.e. $\mathbf{c}'\mathbf{1} = 0$. Hence, the treatment contrasts will be considered mainly in the paper.

If the contrast $\mathbf{c}'\boldsymbol{\tau}$ is estimable within the f-th stratum, then its BLUE within that stratum has the form $(\widehat{\mathbf{c}'\boldsymbol{\tau}})_f = \mathbf{c}'\boldsymbol{\tau}_f^o$, where $\boldsymbol{\tau}_f^o = \mathbf{C}_f^{-}\mathbf{Q}_f$ is a solution of the normal equation in stratum f. The within stratum variance of $(\widehat{\mathbf{c}'\boldsymbol{\tau}})_f$ is equal to $\gamma_f\mathbf{c}'\mathbf{C}_f^{-}\mathbf{c}$, i.e. $\mathrm{Var}[(\widehat{\mathbf{c}'\boldsymbol{\tau}})_f] = \gamma_f\mathbf{c}'\mathbf{C}_f^{-}\mathbf{c}$ for $f = 1,2,3$.

The statistical analysis of experimental data usually consists of testing general and particular hypotheses. The tests can be obtained from the within stratum analysis of variance as given in Table 1.

The symbols occurring in Table 1 denote: $SSY_f = \mathbf{y}'\mathbf{P}_f^*\mathbf{y}$ – the total sum of squares, $SST_f = \mathbf{Q}'\mathbf{C}_f^{-}\mathbf{Q}$ – the treatment sum of squares, $SSE_f = SSY_f - SST_f$ – the error sum of squares, $\nu_{Tf} = r(\mathbf{C}_f)$, $\nu_f = r(\mathbf{P}_f^*)$, $\nu_{Ef} = \nu_f - \nu_{Tf}$, , for the f-th stratum, f=0,1,2,3.

If the normal distribution of the random terms of the linear model (2.5) is assumed, then it is easy to obtain an exact test (F-test) of the hypothesis

Table 1

Analysis of variance in the stratum f

Source of variation	d.f.	S.S.	E(S.S.)
Treatments (in f)	ν_{Tf}	SST_f	$\nu_{Tf}\,\gamma_f + \tau' C_f \tau$
Error (in f)	ν_{Ef}	SSE_f	$\nu_{Ef}\,\gamma_f$
Total (in f)	ν_f	SSY_f	$\nu_f \gamma_f + \tau' C_f \tau$

$H_{0f} : \boldsymbol{\tau}' \mathbf{C}_f \boldsymbol{\tau} = 0$, $f = 1,2,3$, or the test of some subhypothesis defined by a contrast $\mathbf{c}'\boldsymbol{\tau}$ estimable in that stratum, i.e., $H^*_{0f} : \mathbf{c}'\boldsymbol{\tau} = 0$, where $\mathbf{c}'\mathbf{1} = 0$ and $\mathbf{c}'\mathbf{C}_f^- \mathbf{C}_f = \mathbf{c}'$.

The hypothesis H^*_{0f} can be tested in every stratum in which the contrast is estimable. Also, a combined test can be used to improve the statistical properties of the strata tests.

Because of (2.3), the particular vector defining contrasts among the environments, whole-plot and sub-plot treatments can be expressed as: $\mathbf{c} = z\mathbf{q}_g \otimes \mathbf{p}_w \otimes \mathbf{h}_j$, $g = 1,2,\ldots,d$, $w = 1,2,\ldots,S$, $j = 1,2,\ldots,T$, where z is a normalizing constant, $\mathbf{q}_g$, $\mathbf{p}_w$, and $\mathbf{h}_j$ are contrast vectors or vectors of ones of the length d, S and T respectively. This structure is very useful in checking contrasts' estimability in a stratum.

Applying the criterion $\mathbf{c}'\mathbf{C}_0^- \mathbf{C}_0 = \mathbf{c}'$ of the estimability of a linear function $\mathbf{c}'\boldsymbol{\tau}$ and $\mathbf{C}_0^- \mathbf{C}_0 = n_0^{-1} \mathrm{diag}(\mathbf{1r}'_1, \mathbf{1r}'_2, \ldots, \mathbf{1r}'_d)$ we can see that some contrasts among the environments are estimable in this stratum, for example, $\mathbf{c}' = (\mathbf{r}', -\mathbf{r}', 0, \ldots, 0)$. In the paper we will not test the hypotheses concerning environments because the number of degrees of freedom for error is equal to zero. Some methods of testing hypotheses in such situation are known. However, they are beyond the scope of this paper.

From the structure of $\mathbf{C}_1^- \mathbf{C}_1$, it follows that in the first stratum no contrast among the environments is estimable.

From the structure of $\mathbf{C}_3^- \mathbf{C}_3$, it results that in the third stratum no contrast among environments and no contrast among levels of factor A (whole-plot treatments) and none of their interaction contrasts are estimable. No more general statements concerning contrasts' estimability can be concluded from the structure of matrices $\mathbf{C}_1$, $\mathbf{C}_2$ and $\mathbf{C}_3$.

The particular complete cases are considered for example by Utz (1971), Carmer *et al.* (1989), McIntosh (1983) and in some monographs, e.g. Gomez and Gomez (1984). The linear models adopted to the series of split-plot experiments are different from that considered in this paper.

3. Planning of experiments

The allocation of treatments to the blocks and to the whole-plots in the g-th environment (plan D_g) is described by two incidence matrices $\mathbf{N}_{1g}$ and $\mathbf{N}_{2g}$, $g = 1,2,\ldots,d$. They play an important role in the planning of considered experiments. So, they will be presented in details. Let us note that possible plans D_g, $g = 1,2,\ldots,d$, may be the same or different. If the chosen plans D_g are different then, without computer help, no useful methods of proper experiment planning can be given. At the beginning, let us consider some property of designs called general balance (cf. Houtman and Speed, 1983). It is connected with the relationship between the treatments structure and block structure of the design. The generally balanced design has many desirable statistical properties. One of them is connected with the eigenvalues (ε) and eigenvectors ($\mathbf{w}$) of the matrices $\mathbf{C}_f$, calculated with respect to $\mathbf{r}^{\delta}$, i.e., $\mathbf{C}_f\mathbf{w}_i = \varepsilon_{fi}\mathbf{r}^{\delta}\mathbf{w}_i$, f=0,1,2,3; i=1,2,...,v. The linear functions $\mathbf{c}_i = \mathbf{r}^{\delta}\mathbf{w}_i$ are linearly independent and all orthogonal to $\mathbf{1}$. It means that they span the subspace of vectors defining contrasts, often called basic contrasts (cf. Pearce *et al.*, 1974). In the generally balanced design the space of vectors defining contrasts is the same in all strata. Because of that fact, we can treat the eigenvalues ε_{fi} as the stratum efficiency factors of the design. This measure is very useful in the optimal choice of design D_g in the g-th environment. The generally balanced design will be considered in the paper only. The simple way of checking the general balance property of a design was given by Mejza (1992). Shortly, the design is generally balanced iff the matrices $\mathbf{C}_f$, $f = 1,2,3$, mutually commute with respect to $\mathbf{r}^{-\delta}$, i.e., $\mathbf{C}_f\mathbf{r}^{-\delta}\mathbf{C}_{f'} = \mathbf{C}_{f'}\mathbf{r}^{-\delta}\mathbf{C}_f$, $f \neq f'$, $f, f' = 1,2,3$.

As was mentioned earlier, the considered design can be incomplete with respect to whole-plot treatments only, with respect to sub-plot treatments only and finally, with respect to both kinds of treatments. For the last case, which is the most general, it is difficult to give general methods of construction which do not use computer. The first two cases are simpler and there are some useful cases of designs worth taking into account. The two cases, called A and B, will be considered separately.

In case A, we will have incompleteness with respect to whole-plot treatments in the blocks and completeness with respect to sub-plot treatments in the whole-plots.

In case B, the design considered is complete with respect to whole-plot treatments and incomplete with respect to sub-plot treatments.

3.1. Case A

The assumption concerning completeness of sub-plot treatments, $t=T$, leads to some simplifications. We have $\mathbf{N}_{1g} = \mathbf{N}_{Ag} \otimes \mathbf{1}_T$, where $\mathbf{N}_{Ag}$ are $S \times b$ incidence matrices of whole-plot treatments *vs.* blocks, and let us note, that $\mathbf{N}_{Ag}\mathbf{1} = \mathbf{r}_{Ag}$, where $\mathbf{r}_{Ag}$ denotes the vector of whole-plot replications in the g-th environment, $g = 1,2,\ldots,d$. Hence, we have:

$$\mathbf{C}_{1g} = (kT)^{-1}(\mathbf{N}_{Ag}\mathbf{N}'_{Ag} - b^{-1}\mathbf{r}_{Ag}\mathbf{r}'_{Ag}) \otimes \mathbf{J}_T \ , \quad \mathbf{C}_{2g} = T^{-1}(\mathbf{r}^{\delta}_{Ag} - k^{-1}\mathbf{N}_{Ag}\mathbf{N}'_{Ag}) \otimes \mathbf{J}_T \ ,$$

$$\mathbf{C}_{3g} = \mathbf{r}^{\delta}_{Ag} \otimes (\mathbf{I}_T - T^{-1}\mathbf{J}_T)\,, \quad g = 1,2,\ldots,d \ . \tag{3.1}$$

Let us note that generally, planning of considered series of experiments is equivalent to a proper choice of the incidence matrices for whole-plot treatments in environments, i.e. $\mathbf{N}_{Ag}$, $g=1,2,\ldots,d$.

Now, to simplify the calculations, we consider the series, where the same incidence matrix for whole-plot treatments in all environments is applied, i.e. $\mathbf{N}_{Ag} = \mathbf{N}_A$, $g = 1,2,\ldots,d$. It leads to following simplifications: $\mathbf{C}_{1g} = \mathbf{C}^{\#}_1$, $\mathbf{C}_{2g} = \mathbf{C}^{\#}_2$, $\mathbf{C}_{3g} = \mathbf{C}^{\#}_3$, for $g = 1,2,\ldots,d$.

Let ε_w denote the eigenvalue of the matrix $\mathbf{C}_A = \mathbf{r}^{\delta}_A - k^{-1}\mathbf{N}_A\mathbf{N}'_A$ with respect to $\mathbf{r}^{\delta}_A$, $w=1,2,\ldots,S$.

From the structures of the matrices $\mathbf{C}_1$, $\mathbf{C}_2$ and $\mathbf{C}_3$ (see (2.8)–(2.10)) and from (3.1) we have:

1) the contrasts among whole-plot treatments (A) and interaction environments × whole-plot treatments contrasts ($P \times A$) are estimated in the first and second stratum with efficiency equal to $1 - \varepsilon_w$ and ε_w, respectively,

2) the contrasts among the sub-plot treatments (B), two-way interaction contrasts: whole-plot treatments × sub-plot treatments ($A \times B$), environments × sub-plot treatments ($P \times B$), and three-way interaction contrasts: environments × whole-plot treatments × sub-plot treatments ($P \times A \times B$), are estimated in the third stratum with full efficiency.

As the particular case, let us consider the series of incomplete split-plot designs where in all environments the same balanced incomplete block design (BIB) for the whole-plot treatments is applied. Then all the occurrence matrices $\mathbf{N}_A\mathbf{N}'_A = (r - \lambda)\mathbf{I} + \lambda\mathbf{J}$ are equal, where $\lambda = r(k-1)/(S-1)$ and r denotes the number of whole-plot treatment replication in each environment. The efficiency factors $1-\varepsilon_w$ and ε_w are equal to $(S-k)/k(S-1)$ and $S(k-1)/k(S-1)$, respectively, $w = 1,2,\ldots,S-1$.

It will be convenient to introduce abbreviations to describe the properties of balance of a design. Let $E_f(e,\varepsilon)$ denote the property that e contrasts among the levels of factor E (or interaction contrasts) are estimated in the f-th stratum with efficiency of ε or, in other words, we say that the design is $E_f(e,\varepsilon)$-balanced (if e includes all contrasts then in some cases it is omitted).

Summarizing, the design considered above is $A_1(S-1, (S-k)/k(S-1))$-balanced, AP_1 $((S-1)(d-1),(S-k)/k(S-1))$-balanced, A_2 $(S-1, S(k-1)/k(S-1))$-balanced, AP_2 $((S-1)(d-1), S(k-1)/k(S-1))$-balanced. All other contrasts are estimated with full efficiency.

Partially Balanced Incomplete Block Designs with Two Associate Classes (PBIBD(2)) are other types of block designs useful to generate some series of the considered incomplete split-plot designs.

Let us note that sometimes the levels of factor A are in fact the treatment combinations of levels of two factors C and D, say, with m and q levels, respectively. Then the treatment combinations can be expressed as: $A = \{C_hD_p,\ \ h = 1,2,\ldots,m,\ p = 1,2,\ldots,q\}$. In this case we can use the subclasses of the PBIBD(2) called the Group Divisible Designs (GDPBIBD(2)). The details of planning and of the analysis of incomplete split-plot design generated by GDPBIBD(2) are given by Mejza (1991).

Let us consider the GDPBIBD(2) with parameters $(mq, r^*, k^*, b^*, \lambda_1, \lambda_2)$ for mq treatments divided into m groups of q treatments each. The design is such that all pairs of the treatments belonging to the same group occur together in λ_1 blocks while pairs of treatments from different groups occur together in λ_2 blocks. The parameters k^*, r^*, b^* denote the block size, number of treatment replications and number of blocks, respectively.

The levels of factor C can be treated as groups. In such an association scheme the occurrence matrix of GDPBIBD(2) can be expressed as:

$$\mathbf{NN}' = \omega_0\mathbf{L}_0 + \omega_1\mathbf{L}_1 + \omega_2\mathbf{L}_2\,, \tag{3.2}$$

where ω_0, ω_1, ω_2 are eigenvalues of matrix $\mathbf{NN}'$ so that $\omega_0 = r^*k^*$, with multiplicity 1, $\omega_1 = r^* - \lambda_1$, with multiplicity $m(q-1)$, $\omega_2 = r^*k^* - v^*\lambda_2$, with multiplicity $m-1$ (cf. Raghavarao, 1971), while the matrices $\mathbf{L}_0$, $\mathbf{L}_1$, $\mathbf{L}_2$ have the form

$$\begin{aligned}
\mathbf{L}_0 &= v^*\mathbf{J}_{v^*},\\
\mathbf{L}_1 &= m^{-1}\mathbf{J}_m \otimes (\mathbf{I}_q - q^{-1}\mathbf{J}_q) + (\mathbf{I}_m - m^{-1}\mathbf{J}_m) \otimes (\mathbf{I}_q - q^{-1}\mathbf{J}_q),\\
\mathbf{L}_2 &= q^{-1}(\mathbf{I}_m - m^{-1}\mathbf{J}_m) \otimes \mathbf{J}_q\,.
\end{aligned} \tag{3.3}$$

Using the occurrence matrix (3.2) instead of the $\mathbf{N}_A\mathbf{N}'_A$ in (3.1) we have that above design is:

$C_1(m-1),\ \omega_2/rk)$- and $C_2(m-1),\ 1-\omega_2/rk)$-balanced,
$D_1(q-1,\ \omega_1/rk)$- and $D_2(q-1,\ 1-\omega_1/rk)$-balanced,
$CD_1((m-1)(q-1), \omega_1/rk)$- and $CD_2((m-1)(q-1),\ 1-\omega_1\ /\ rk)$-balanced,
$PC_1(m-1),\ \omega_2/rk)$- and $PC_2(m-1),\ 1-\omega_2/rk)$-balanced,
$PD_1(q-1,\ \omega_1/rk)$- and $PD_2(q-1,\ 1-\omega_1/rk)$-balanced,
$PCD_1((m-1)(q-1),\ \omega_1/rk)$- and $PCD_2((m-1)(q-1),\ 1-\omega_1\ /\ rk)$-balanced.

All other contrasts are estimated with full efficiency in the second stratum.

3.2. Case B

Let the levels of factor B (sub-plot treatments) be arranged in some incomplete block design where whole-plots are treated as blocks. Moreover, it will be assumed that all levels of the factor A (whole-plot treatments) occur within blocks, i.e., k=S. In general, within the factor A levels different designs for the factor B levels may be used. It follows that $\mathbf{N}_{1g} = [\mathbf{N}'_{B1g},\ \mathbf{N}'_{B2g},\ldots,\ \mathbf{N}'_{BSg}]'$, where $\mathbf{N}_{Bwg}$ is $T\times b$ incidence matrix for sub-plot treatments within w-th whole-plot treatment (w=1,2,...,S) in the g-th environment (g=1,2,...,d). Nevertheless, to simplify the calculations, we assume that occurrence matrices $\mathbf{N}_{Bwg}\mathbf{N}'_{Bw'g}$ are the same in each environment, i.e. $\mathbf{N}_{Bwg}\mathbf{N}'_{Bw'g}= \mathbf{N}_{Bg}\mathbf{N}'_{Bg}$ for w,w'=1,2,...,S and g=1,2,...,d.

From (2.8)–(2.10) the structures of matrices $\mathbf{C}_{1g}$, $\mathbf{C}_{2g}$ and $\mathbf{C}_{3g}$ are as follows:

$$\mathbf{C}_{1g} = (St)^{-1}\mathbf{J}_S \otimes (\mathbf{N}_{Bg}\mathbf{N}'_{Bg} - b^{-1}\mathbf{r}_{Bg}\mathbf{r}'_{Bg}),$$

$$\mathbf{C}_{2g} = t^{-1}(\mathbf{I}_S - S^{-1}\mathbf{J}_S) \otimes \mathbf{N}_{Bg}\mathbf{N}'_{Bg} \tag{3.4}$$

$$\mathbf{C}_{3g} = \mathbf{I}_S \otimes (\mathbf{r}^{\delta}_{Bg} - t^{-1}\mathbf{N}_{Bg}\mathbf{N}'_{Bg}), \qquad g = 1,2,\ldots,d\ ,$$

where $\mathbf{r}_{Bg} = \mathbf{N}_{Bg}\mathbf{1}$ denotes the vector of sub-plot replications, the same in each whole-plot treatment in the g-th environment.

Similarly as in case A, let us assume that all $\mathbf{N}_{Bg}\mathbf{N}'_{Bg}$ are the same in all environments, i.e. $\mathbf{N}_{Bg}\mathbf{N}'_{Bg} = \mathbf{N}_B\mathbf{N}'_B$ for $g = 1,2,\ldots,d$. This leads to $\mathbf{C}_{fg} = \mathbf{C}^{\#}_f$, for $f = 1,2,3;\ g = 1,2,\ldots,d$.

Let ε_j denote the eigenvalue of the matrix $\mathbf{C}_B = \mathbf{r}^{\delta}_B - t^{-1}\mathbf{N}_B\mathbf{N}'_B$ with respect to $\mathbf{r}^{\delta}_B$, j=1,2,...,T.

From the structure of matrices (3.4) we have that in this type of designs the contrasts among the levels of factor B are estimated with efficiency equal to $1-\varepsilon_j$ in the first stratum and with efficiency ε_j in the third stratum. Similarly, the two-way interaction contrasts, environments × sub-plot treatments ($P\times B$), are estimated with efficiencies $1-\varepsilon_j$ and ε_j in the first and third strata, respectively, while whole-plot treatments × sub-plot treatments ($A\times B$), and three-way interaction contrasts, environment × whole-plot treatments × sub-plot treatments

($P \times A \times B$), are estimated in the second and third strata with efficiencies $1-\varepsilon_j$ and ε_j, respectively. All other contrasts, i.e. among whole-plot treatments (A) and two-way interaction contrasts, environments × whole-plot treatments ($P \times A$), are estimated with full efficiency in the second stratum.

Let us consider two particular cases of block designs, the same as previously, i.e. BIB design and GDPBIBD(2), to generate the incomplete split-plot designs.

The series of incomplete split-plot designs generated by BIB design, i.e. when $\mathbf{N}_B\mathbf{N}'_B = (r-\lambda)\mathbf{I}_T + \lambda\mathbf{J}_T$, are:

$B_1(T-1, (T-t)/t(T-1))$-balanced,
$PB_1((T-1)(d-1), (T-t) / t(T-1))$-balanced,
$B_3(T-1, T(t-1)/t(T-1))$-balanced,
$PB_3((T-1)(d-1), T(t-1)/t(T-1))$-balanced,
$AB_2((S-1)(T-1), (T-t) / t(T-1))$-balanced,
$AB_3((S-1)(T-1), T(t-1) / t(T-1)$-balanced,
$PAB_2((d-1)(S-1)(T-1), (T-t)/t(T-1))$-balanced,
$PAB_3((d-1)(S-1)(T-1), T(t-1) / t(T-1))$-balanced.

All other contrasts are estimated with full efficiency in the second stratum.

Similarly, the series of incomplete split-plot designs generated by GDPBIBD(2) are:

$C_1(m-1, \omega_2/rt$- and $C_3(m-1, 1-\omega_2/rt)$-balanced,
$D_1(q-1, \omega_1/rt)$- and $D_3(q-1, 1-\omega_1/rt)$-balanced,
$CD_1((m-1)(q-1), \omega_1/rt)$- and $CD_3((m-1)(q-1), 1-\omega_1/rt)$-balanced,
$AC_2((S-1)(m-1), \omega_2/rt)$- and $AC_3((S-1)(m-1), 1-\omega_2/rt)$-balanced,
$AD_2((S-1)(q-1), \omega_1/rt)$- and $AD_3((S-1)(q-1), 1-\omega_1/rt)$-balanced,
$ACD_2((S-1)(m-1)(q-1), \omega_1/rt)$- and
$ACD_3((S-1)(m-1)(q-1), 1-\omega_1/rt)$-balanced,
$PC_1((d-1)(m-1), \omega_2/rt)$- and $PC_3((d-1)(m-1), 1-\omega_2/rt)$-balanced,
$PD_1((d-1)(q-1), \omega_1/rt)$- and $PD_3((d-1)(q-1), 1-\omega_1/rt)$-balanced,
$PCD_1((d-1)(m-1)(q-1), \omega_1/rt)$- and $PCD_3((d-1)(m-1)(q-1), 1-\omega_1/rt)$-balanced,
$PAC_2((d-1)(S-1)(m-1), \omega_2/rt)$- and
$PAC_3((d-1)(S-1)(m-1), 1-\omega_2/rt)$- balanced,
$PAD_2((d-1)(S-1)(q-1), \omega_1/rt)$- and $PAd_3((d-1)(S-1)(q-1), 1-\omega_1/rt)$- balanced,
$PACD_2((d-1)(S-1)(m-1)(q-1), \omega_1/rt)$- and
$PACD_3((d-1)(S-1)(m-1)q-1), 1-\omega_1/rt)$-balanced.

Let us note that there are many different efficiencies connected with treatment contrasts estimation. These efficiencies allow us to choose the proper design D_g, g=1,2,...,d. Moreover, for some of the GDPBIBD(2) the efficiency factor ε_j is equal to one. In this case the contrasts connected with this eigenvalue are estimated with full efficiency. The plans of PBIBD(2) are given by Clatworthy (1973).

4. Example

Let us consider series of experiments carried out in incomplete split-plot design, incomplete with respect to sub-plot treatments (Case B). The example is artificial and serves to illustrate the statistical properties of the series.

Let us consider a design where three whole-plot treatments (S=3) and six sub-plot treatments (T=6) have to be observed in three locations (d=3). For each location, the experimental material consists of b=10 blocks. Each block has three (k=3) whole-plots and each whole-plot has three (t=3) sub-plots. We assume that the same (with precision to a permutation of rows and/or columns of the incidence matrix) incomplete split-plot design will be used in all locations.

According to the structure of the experimental material we can use the design described by the incidence matrix $\mathbf{N}_1$ of the form $\mathbf{N}_1 = \mathbf{1}_S \otimes \mathbf{N}_B$, where $\mathbf{N}_B$ is the incidence matrix of the BIB design (Plan 11.4, Cochran and Cox, 1957) for the sub-plot treatments. Exactly we have

$$\mathbf{N}_B = \begin{pmatrix} 1&1&1&1&1&0&0&0&0&0 \\ 1&1&0&0&0&1&1&1&0&0 \\ 0&0&1&1&0&1&1&0&1&0 \\ 0&0&1&0&1&1&0&1&0&1 \\ 1&0&0&0&1&0&1&0&1&1 \\ 0&1&0&1&0&0&0&1&1&1 \end{pmatrix}.$$

The treatment combinations are replicated 5 times at each location. The matrix $\mathbf{C}_1^{\#}$ has the form:

$$\mathbf{C}_1^{\#} = (1/18)\begin{bmatrix} \mathbf{X}&\mathbf{X}&\mathbf{X} \\ \mathbf{X}&\mathbf{X}&\mathbf{X} \\ \mathbf{X}&\mathbf{X}&\mathbf{X} \end{bmatrix},$$

where $\mathbf{X} = 6\mathbf{I}_6 - \mathbf{J}_6$. The matrix $\mathbf{C}_2^{\#}$ is of the form:

$$\mathbf{C}_2^{\#} = \begin{bmatrix} \mathbf{Z}&\mathbf{M}&\mathbf{M} \\ \mathbf{M}&\mathbf{Z}&\mathbf{M} \\ \mathbf{M}&\mathbf{M}&\mathbf{Z} \end{bmatrix},$$

where $\mathbf{Z} = (2/9)(3\mathbf{I}_6 + 2\mathbf{J}_6)$, $\mathbf{M} = (-1/9)(3\mathbf{I}_6 + 2\mathbf{J}_6)$. The matrix $\mathbf{C}_3^{\#}$ has the form $\mathbf{C}_3^{\#} = (2/3)\text{diag}(\mathbf{X}, \mathbf{X}, \mathbf{X})$, where $\mathbf{X}$ is as above.

Table 2
Stratum efficiency factors

Type of contrast	Stratum		
	first	second	third
A	---	1	---
B	0.2	---	0.8
A×*B*	---	0.2	0.8
P×*A*	---	1	---
P×*B*	0.2	---	0.8
P×*A*×*B*	---	0.2	0.8

From the structure of these matrices we can obtain the efficiency factors, which are given in Table 2.

The contrasts between locations are estimated with full efficiency (=1) in the zero stratum.

Nearly all information concerning sub-plot treatment effects and different interaction effects with these treatments is included in the third stratum. In this stratum the precision of comparisons is the highest. It means that in the overall analysis we can either use whole information from all strata (for example by combining the information) or to limit our inference to the third stratum (accepting some loss in precision).

LITERATURE

Carmer, S.G., Nyquist, W.E., Walker, W.M. (1989). Least significance for combined analysis of experiments with two- or three- factor treatment designs. *Agron. J.* **81**, 665-672.

Clatworthy, W.H. (1973). Tables of two-associate-class partially balanced designs. U.S. Department of Commerce, National Bureau of Standards, Series 63. Washington, D.C.

Cochran, W.G., Cox, G.M. (1957). *Experimental designs*. Wiley, New York.

Gomez, K.A., and Gomez, A.A. (1984). *Statistical Procedures for Agricultural Research*. Wiley, New York.

Houtman, A.M., Speed, T.P. (1983). Balance in designed experiments with orthogonal block structure. *Ann. Statist.* **11**, 1069-1085.

McIntosh, M.S. (1983). Analysis of combined experiments. *Agron.J.* **75**, 153-155.

Mejza, I., Mejza, S. (1994). Series of experiments carried out in incomplete split-plot designs. *SoftStat'93, Advances in Statistical Software* 4, F. Faulbaum (ed). Gustav Fisher Verlag, New York, 635-642.

Mejza, I. (1991). Układ niekompletny o jednostkach pojedynczo rozszczepionych, w którym poziomy czynnika drugiego są rozmieszczone w GDPBIBD(2). *Dwudzieste pierwsze Colloquium Metodologiczne z Agro-Biometrii,* PAN, Warszawa, 108-119.

Mejza, S. (1987). Incomplete split-plot designs. In: *Proc. Second International Tampere Conference in Statistics,* T.Pukkila and S.Puntanen eds., 575-584.

Mejza, S. (1992). On some aspects of general balance in designed experiments. *Statistica* LII.

Nelder, J.A. (1965). The analysis of randomized experiments with orthogonal structure. *Proc. Royal Soc. London, Ser.* **A**, 147-178.

Pearce, S.C., Caliński, T., Marshall, T.F. de C. (1974). The basic contrasts of experimental design with special reference to the analysis of data. *Biometrika* **61**, 449-460.

Raghavarao, D. (1971). *Construction and combinatorial problems in design of experiments*. Wiley, New York.

Rao, C.R. and Mitra, S.K. (1971). *Generalized inverse of matrices and its applications*. Wiley, New York.

Utz, H.F. (1971). Die zusammennfassende Analyse einer Serie von Spaltanlagen. *EDV in Medizin und Biologie* **2**, 50-55.

Received 12 September 1993; revised 15 March 1994