



# Article NRC Designs—New Tools for Successful Agricultural Experiments

Agnieszka Łacka 匝

Department of Mathematical and Statistical Methods, Poznań University of Life Sciences, Wojska Polskiego 28, 60-637 Poznań, Poland; agnieszka.lacka@up.poznan.pl

**Abstract:** In a nested row–column design (NRC), the experimental units in each of n blocks are grouped into  $n_1$  rows and  $n_2$  columns. Due to its structure, this experimental design allows full control of the experimental material and a relatively simple feedback loop within the "statistical triangle". By applying such designs in agricultural experiments, we provide an insurance policy against future unexpected problems. Until now, the cost of this policy has been a complex statistical analysis of experimental data. This paper proposes a new "direct" approach to ANOVA based on the latest literature on the subject. The paper provides the theoretical foundations of this approach, indicates the possibility of applying it to factorial and near-factorial experiments, and supplements the theory with a familiar letter-based representation of all-pairwise comparisons, which has so far been lacking in the literature. The methodology is illustrated by the analysis of a field experiment carried out to improve the use of fungicides against late blight in tomato processing. The presented analytical tools are supplemented with code in R.

**Keywords:** experimental design; randomization; blocking; nested row–column design; analysis of variance—direct approach; letter-based representation



Citation: Łacka, A. NRC Designs—New Tools for Successful Agricultural Experiments. *Agronomy* 2021, *11*, 2406. https://doi.org/ 10.3390/agronomy11122406

Received: 9 September 2021 Accepted: 22 November 2021 Published: 25 November 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/).

## 1. Introduction

Although almost 100 years have passed since Fisher published the basic principles of experimental planning, it remains one of the most important problems faced by biologists and statisticians [1–3]. Only proper planning and adequate analysis of experimental data can be a guarantee of success—a guarantee that the expenditure incurred will enable the formulation of correct and satisfactory conclusions regarding the research problems addressed.

In agricultural comparative experiments we conventionally follow the general theory of scientific research, the heart of which is the "statistical triangle" described by Hinkelmann and Kempthorne [4] and extended by Casler [5]. The starting point in conducting experiments is to pose questions and formulate hypotheses, which should be transcribed to models based on a specific topic (e.g., tomatoes) and then translated into a statistical model and developed in conjunction with statistical design (see Figure 1). The data obtained in the experiment, after being subjected to an analysis determined by the researcher's previous assumptions, constitute a basis for formulating conclusions regarding the research problems addressed, usually to determine the influence of the studied treatments on the observed variable [6]. At this point, when the process seems complete, a feedback loop should appear, as a good experiment usually leads to more questions and hypotheses. The acquired knowledge should be used to answer new biological and statistical questions, which will increase the effectiveness of the experiments [5].

One of the prerequisites for success is to incorporate the structure of the experimental material into the experimental design so that the results are not distorted, for example by soil variability. Even seemingly homogeneous fields can give heterogeneous answers [7]. In the literature, therefore, it is proposed to use bidirectional blocking in field experiments as an insurance policy, especially if the statistical analysis is to be based on the analysis of variance. In this paper, it is proposed to use nested row–column designs, which provide a

powerful structural base for the two-dimensional control of experimental trends [8], and also give the possibility of a relatively simple feedback loop (Figure 1). Experiments with multiple blocking structures are usually carried out using a mixed-model specification, which in the classical approach leads to the analysis of variance in the strata [8,9]. Unfortunately, the inconvenience of inference (see Section 2) in this approach can demotivate the use of safe design solutions. Motivated by the latest publications on the direct approach to ANOVA [10–13], I suggest applying this modern approach to the analysis of experiments laid out as NRC designs. The available literature on this topic requires the reader to have an advanced mathematical apparatus. For that reason, this paper refers only to those elements of direct ANOVA that are relevant from the practical point of view.

The aim of the paper is to present and illustrate a complete set of analytical tools that make it possible to carry out direct inference in a nested row–column design under a randomization-derived mixed model of observations. Moreover, to the best of my knowledge, for these designs, there is no package or statistical program that can run the described procedure. For this reason, the article has been supplemented with R code and procedures (Appendix A).

The paper is organized as follows. Section 2 describes three basic principles of an insurance policy in agricultural experiments. Section 3 begins with introduction of the NRC designs and presents the randomization-derived mixed model of observations, then gives the theoretical background of the analysis under the direct approach. The section ends by showing how to proceed with the analysis for the very common treatment factorial experiments, and a novel letter-based procedure for all-pair comparisons is proposed. A detailed analysis of a field experiment carried out to improve the use of fungicides against late blight in tomato processing, using the proposed approach, is presented in Section 4. The paper ends with a discussion in Section 5.



**Figure 1.** Casler's [5] flow diagram of the logical steps in scientific experimentation, including a feedback loop that allows for new scientific hypotheses and experimental design modifications to future experiments. The three central boxes form the "statistical triangle" [4].

#### 2. Insurance Policy in Agricultural Experiments

It should be remembered that statistical design consists of equal measure of treatment design and experimental design. Treatment design is responsible for the arrangement of treatments on experimental units, and sets up a theoretical plan by which the treatment levels are arranged in the experiment. The choice of a plan is related to the definition of the experimental factors and control treatments as well as the technical possibilities related to the methodology of conducting the research and the manner of observation and data collection. The problem of choosing a plan for different types of experiments has been discussed in numerous publications. For instance, some methods of designing factorial experiments with a control treatment in a block design with nested rows and columns are widely considered by Bailey and Łacka [14] and Bose and Mukerjee [15]. An example of this type of experiments is also presented in Section 4. Note the different number of replicates for the control compared to the rest of the treatments in the experiment considered there, which is often found in theoretical plans of experiments with a control treatment and multiple blocking structures [16].

Experimental design should be based on the principles of replication, randomization and local control, i.e., blocking. When planning an experiment, all these principles must be considered from the point of view of its purpose and the conditions under which the experiment is to be carried out, due to their fundamental influence on the analysis of the experimental data and further inference [9].

The principle of replication, i.e., the use of a treatment on several experimental units, is commonly known. It is well known that replication provides the possibility of estimating the experimental error affecting the observation. It is thanks to replication that we obtain greater precision in the estimation of treatment effects or their comparisons. In general, the standard error of the estimator of effects decreases with an increase in the number of replications; however, the use of a large number of repetitions for each treatment can make it difficult to ensure the homogeneity of the experimental material. This fact should not always be viewed as a disadvantage, as such heterogeneity may reflect the natural variability of the population that we infer from the experiment. Thus, replication contributes to an increase in the representativeness of the experimental material.

In the accepted theory, blocking should be related to the earlier recognition of the directions of variability of the material used in the experiment, and lead to a grouping of the experimental units into a system of blocks such that the units inside the blocks are as homogeneous as possible. Local control is commonly used for field trials and plant protection trials carried out under found conditions; however, as research has shown, it should also be included in greenhouse experiments. Hartung et al. [17] indicate blocking as more efficient in improving the precision of greenhouse experiments than the re-arrangement of pots, and hence they recommend it for comparative greenhouse experiments. On the other hand, Casler [5] points to the "dark side" of blocking, noting that block designs in which blocks are linearly arranged without prior recognition of spatial variation can seriously reduce the likelihood of success of the experiment. Taking into account the fact that many experimental stations have visually homogeneous fields, the author suggests the use of two-way blocking as a kind of insurance policy against erroneous blocking. It has been proposed in the literature to use many different designs to control two sources of external variation, such as Latin squares, Youden squares, generalized Youden designs, or row-column designs. Unfortunately, most of these designs are characterized by several important constraints: Youden squares and generalized Youden designs exist only for a limited number of parameter combinations, hence their practical application is limited. On the other hand, row-column designs can have row-column interaction problems when multiple rows and columns are used. Moreover, neither of these designs are applicable to the analysis of a series of experiments carried out, for example, in different locations. An excellent answer to these constraints are nested row-column (NRC) designs, i.e., designs in which in each of *n* blocks,  $n_1n_2$  experimental units are grouped into  $n_1$  rows and  $n_2$ columns [8,12,18]. These experimental designs, which are a natural extension of the previously introduced row-column designs, enable not only full control of the experimental material, but also a relatively simple feedback loop and the making of modifications within the "statistical triangle" when the experimental setup turns out to be "overestimated" and the block insurance policy is exaggerated or other unexpected problems arise in the course

of the experiment. Moreover, as Chang and Notz [19] note, for a given number of experimental units, NRC designs have fewer rows and columns per block. The row–column interactions in the NRC designs are likely not as severe as in row–column designs. Therefore, NRC designs are particularly useful for eliminating heterogenity in two directions when there are row and column interactions.

In field experiments, it is worth taking into account the possibility of loss of some of the biological material as a result of unforeseen weather conditions, such as flooding or hail, which may destroy some plots or the entire experiment. If, as a result of such events, only one block of research material is obtained, it will still be possible to analyze it in a row–column design.

If the experimental material turns out to be homogeneous within a range of rows or columns, ignoring the system of rows (columns) leaves the system of blocks and the system of columns (rows), forming a nested block design or—in a very specific situation for two-factorial experiments—a split-plot design [13].

Ignoring any two of the systems—blocks and rows or blocks and columns or rows and columns—still leaves a proper block design. Of course, any decision to ignore the previously considered blocking must have a biological and statistical basis [20].

Another insurance policy, no less important than blocking, is the two-level randomness principle. The first level concerns sampling from a previously defined population so as to ensure the representativeness of the sample. The second level concerns the arrangement of treatments on experimental units. Thus, randomization is a way to eliminate the bias of measurements performed due to systematic differences between experimental units. Unbias is perhaps the most important goal of randomizing plots before assigning them to treatments, ensuring that certain treatments are not continually favored or harmed by random sources of variation in the experimental material and the environment. Moreover, randomization introduces randomness into the existing variability of experimental units. With randomization, a properly performed statistical analysis of the experimental results becomes correct, assuming that in deriving this analysis, the randomization is fully taken into account [21]. In the case of NRC designs, the randomization scheme is performed by randomizing three times, first blocks, then rows within each block, and similarly for columns. It consists of assigning rows and columns of the theoretical plan to empirical blocks. The mixed observation model obtained in this way is presented in Section 3.

As for any policy, randomization comes at a cost. This cost is, on the one hand, the time devoted to carrying out this mathematical process, and on the other, what is sometimes considered a complicated statistical analysis under a randomization-derived mixed model [5,8,22]. According to the classical procedure, the analysis of variance (ANOVA) under such a model is first performed in strata and then combines the information obtained in them, as originally suggested by Yates [23,24] and thoroughly discussed by Kala [25]. ANOVA in a nested row-column design is related to four strata (apart from the grand mean): the between-blocks stratum, the between-rows-within-blocks stratum, the betweencolumns-within-blocks stratum, and the so-called bottom stratum or rows-by-columns stratum. Four strata mean also four ANOVA tables and the need to combine the information contained in them. Awareness of this may discourage researchers from planning experiments with multiple blocking structures. These doubts are answered in a series of publications on a new approach to the analysis of variance for experiments with an orthogonal block structure (OBS) [10-13]. The authors indicate that for designs with an OBS, the analysis of variance can be performed in a relatively simple way, directly and not by combining results from analyses based on some stratum submodels. It is this approach, which makes the analysis simple and conclusions intuitive, that will be applied in this paper.

## 3. Materials and Methods

#### 3.1. NRC Design

Consider an experiment carried out in a nested row–column design with v treatments arranged in b blocks, each grouped perpendicularly, into  $r_0$  rows of  $c_0$  units and  $c_0$  columns of  $r_0$  units, according to the scheme shown in Figure 2 [14,26].

The usual procedure of randomization [27] of blocks, as well as of rows and of columns within the blocks, makes it possible to present the following derived mixed model:

$$y = X_1 \tau + X_B \beta + X_{R(B)} \rho + X_{C(B)} \gamma + \eta + e, \qquad (1)$$

where  $y = [y'_1, y'_2, ..., y'_b]'$  is an  $n \times 1$  vector of yield data observed on  $n = br_0c_0$  plots of the experiment,  $y_g = [y_{1g}, y_{2g}, ..., y_{n_0g}]'$  representing the yields observed on  $n_0 = r_0c_0$ units (plots) of the block g (=1, 2, ..., b) of the experiment. Suppose also that each of the data vectors  $\{y_1, y_2, ..., y_b\}$  is ordered according to the design rows, and let  $I_x$  and  $\mathbf{1}_x$  denote the unit matrix of order x and the column vector of x ones, respectively. In the model (1),  $X_1 = [X'_{11} : X'_{12} : \cdots : X'_{1b}]'$ ,  $X_B = I_b \otimes \mathbf{1}_{n_0}$ ,  $X_{R(B)} = I_b \otimes I_{r_0} \otimes \mathbf{1}_{c_0}$ ,  $X_{C(B)} = I_b \otimes \mathbf{1}_{r_0} \otimes I_{c_0}$  are the known design matrices for treatments, blocks, rows and columns,  $\boldsymbol{\tau} = [\tau_1, \tau_2, ..., \tau_v]'$  represents the vector of fixed treatment effects,  $\beta$ ,  $\rho$ ,  $\gamma$  stand for the random effects of block, row and column, respectively, while the  $n \times 1$  vectors  $\boldsymbol{\eta}$  and  $\boldsymbol{e}$ stand for the unit error and technical error random variables, all of these random variables being unobservable [12,28,29].



**Figure 2.** Schematic diagram of a nested row–column design with *b* blocks (here b = 3), each grouped perpendicularly into  $r_0$  rows of  $c_0$  units and  $c_0$  columns of  $r_0$  units.

Because all blocks are of equal size, since the rows of the design are of equal size and its columns are also of equal size, not necessarily the same size as the rows, an experiment in such a nested row–column design has, under the randomization-derived model (1), the orthogonal block structure (OBS) property as defined by Houtman and Speed [30]. This means that the considered model may be resolved into five simple stratum submodels, in accordance with the stratification of the experimental units [12]. Using Nelder's [27] notation, this stratification ("block-structure") can be represented by the relation

Units (plots)  $\rightarrow$  (Rows  $\times$  Columns)  $\rightarrow$  Blocks  $\rightarrow$  Total exp. area

and the vector of observations may be written as

$$y = y_1 + y_2 + y_3 + y_4 + y_5$$

where

$$y_1 = \phi_1 y, \ y_2 = \phi_2 y, \ y_3 = \phi_3 y, \ y_4 = \phi_4 y, \ y_5 = \phi_5 y$$

and  $\phi_1 = I_n - c_0^{-1} X_{R(B)} X'_{R(B)} - r_0^{-1} X_{C(B)} X'_{C(B)} + n_0^{-1} X_B X'_B, \phi_2 = c_0^{-1} X_{R(B)} X'_{R(B)} - n_0^{-1} X_B X'_B, \phi_3 = r_0^{-1} X_{C(B)} X'_{C(B)} - n_0^{-1} X_B X'_B, \phi_4 = n_0^{-1} X_B X'_B - n^{-1} \mathbf{1}_n \mathbf{1}'_n$  and  $\phi_5 = n^{-1} \mathbf{1}_n \mathbf{1}'_n$  are symmetric, idempotent and pairwise orthogonal matrices, summing to the identity matrix  $I_n$ .

Thus, under model (1) and using the above representation of y, the expectation vector and covariance (dispersion) matrix of y can be written as

$$E(y) = \phi_1 X_1 \tau + \phi_2 X_1 \tau + \phi_3 X_1 \tau + \phi_4 X_1 \tau + \phi_5 X_1 \tau = X_1 \tau,$$
(2)

$$D(\boldsymbol{y}) \equiv \boldsymbol{V} = \sigma_1^2 \boldsymbol{\phi}_1 + \sigma_2^2 \boldsymbol{\phi}_2 + \sigma_3^2 \boldsymbol{\phi}_3 + \sigma_4^2 \boldsymbol{\phi}_4 + \sigma_5^2 \boldsymbol{\phi}_5,$$
(3)

where the scalars  $\sigma_1^2$ ,  $\sigma_2^2$ ,  $\sigma_3^2$ ,  $\sigma_4^2$  and  $\sigma_5^2$  represent the relevant unknown stratum variances [12,30].

#### 3.2. ANOVA—Direct Approach

The classic approach to data analysis under the model (1) involves applying so-called stratum analysis, which for NRC designs is related to four strata (apart from the grand mean). In a series of four articles published in 2017–2020, the first two by Caliński and Siatkowski [10,11] and the others by Caliński et al. [12,13], a different approach was presented. It turns out that thanks to the OBS property, the analysis of variance can be performed directly, not by combining results from analyses based on stratum submodels. This approach is based on the decomposition of the data vector y into two uncorrelated parts, as

$$y = P_{X_1(V^{-1})}y + (I_n - P_{X_1(V^{-1})})y$$
(4)

using the  $V^{-1}$ -orthogonal projector  $P_{X_1(V^{-1})} = X_1(X'_1V^{-1}X_1)^{-1}X'_1V^{-1}$ , where  $V^{-1}$  is the inverse of the covariance matrix V given in (3) [31]. The roles of these two parts are as follows: the first term provides the best linear unbiased estimator (BLUE) of  $X_1\tau$  in (2), which can be expressed as

$$\widetilde{\boldsymbol{X}}_{1}\boldsymbol{\tau} = \boldsymbol{P}_{\boldsymbol{X}_{1}(\boldsymbol{V}^{-1})}\boldsymbol{y},\tag{5}$$

and the second term can be seen as the residual vector, giving the residual sum of squares in the form

$$\|(I_n - P_{X_1(V^{-1})})y\|_{V^{-1}}^2 = y'V^{-1}(I_n - P_{X_1(V^{-1})})y,$$
(6)

with the residual degrees of freedom given by  $\operatorname{rank}(V : X_1) - \operatorname{rank}(X_1) = n - v$ . See Rao [32] and Caliński et al. [12] for details. It is worth noting that when the projector  $P_{X_1(V^{-1})}$  is used in the way described, the variance  $\sigma_5^2$  in the matrix *V* can be replaced by 1.

Therefore, consider the following hypothesis concerning the treatment's main effects:

$$H_0: \boldsymbol{\tau}_* = \boldsymbol{0},\tag{7}$$

where  $\tau_* = (I_v - n^{-1}\mathbf{1}_v \mathbf{r}')\tau$  and  $\mathbf{r}$  stands for the treatment replication vector in the whole design. To verify this hypothesis—for the analysis of variance—usually the unknown stratum variances in (3) have to be estimated. Because  $\mathbb{E}\{\|\boldsymbol{\phi}_i(\mathbf{I}_n - \mathbf{P}_{X_1(V^{-1})})\mathbf{y}\|^2\} = \sigma_i^2 d'_i$  where  $d'_i = \operatorname{tr}[\boldsymbol{\phi}_i(\mathbf{I}_n - \mathbf{P}_{X_1(V^{-1})})]$ , it is suggested by many authors to use as estimators of  $\sigma_i^2$  the solutions of the equations

$$\|\boldsymbol{\phi}_{i}(\boldsymbol{I}_{n} - \boldsymbol{P}_{X_{1}(V^{-1})})\boldsymbol{y}\|^{2} = \sigma_{i}^{2}d_{i'}^{\prime}$$
(8)

for i = 1, 2, 3, 4 [11,30,33]. Caliński et al. [12] note that the component  $\sigma_5^{-2} \phi_5$  in  $V^{-1} = \sigma_1^{-2} \phi_1 + \sigma_2^{-2} \phi_2 + \sigma_3^{-2} \phi_3 + \sigma_4^{-2} \phi_4 + \sigma_5^{-2} \phi_5$  seems to play no role in the formulae applicable

in the considered analysis of experimental data. They propose a reformulation in the methodology, which would simplify the analysis without causing any changes to its results, obtained by replacing the matrix (3) with

$$V_* = \sigma_1^2 \phi_1 + \sigma_2^2 \phi_2 + \sigma_3^2 \phi_3 + \sigma_4^2 (\phi_4 + \phi_5),$$

where  $\phi_4 + \phi_5 = n_0^{-1} X_B X'_B = I_b \otimes n_0^{-1} \mathbf{1}_{n_0} \mathbf{1}'_{n_0}$ . Furthermore, hence  $V_*^{-1}$  is now obtainable as  $V_*^{-1} = \sigma_1^{-2} \phi_1 + \sigma_2^{-2} \phi_2 + \sigma_3^{-2} \phi_3 + \sigma_4^{-2} (\phi_4 + \phi_5)$ . The relations between the matrices V and  $V_*$ , and their inverses, can be written as

$$\mathbf{V} = \mathbf{V}_* + (\sigma_5^2 - \sigma_4^2)n^{-1}\mathbf{1}_n\mathbf{1}_n' \text{ and } \mathbf{V}^{-1} = \mathbf{V}_*^{-1} + (\sigma_5^{-2} - \sigma_4^{-2})n^{-1}\mathbf{1}_n\mathbf{1}_n'$$

The authors also indicate that the BLUE of  $au_*$  can be written as

$$\hat{\boldsymbol{\tau}}_* = (\boldsymbol{I}_v - n^{-1} \boldsymbol{1}_v \boldsymbol{r}') (\boldsymbol{X}_1' \boldsymbol{V}_*^{-1} \boldsymbol{X}_1)^{-1} \boldsymbol{X}_1' \boldsymbol{V}_*^{-1} \boldsymbol{y}_*,$$
(9)

where  $y_* = (I_n - n^{-1} \mathbf{1}_n \mathbf{1}'_n) y$ . Now, assuming that  $y \sim N_n(X_1 \tau, V)$ , with variances estimated by solving Equation (8), verification of the hypothesis (7) will be based on the formulae presented in Table 1, which correspond to the statistics

$$\widehat{F} = \frac{n-v}{v-1} \frac{\widehat{SS}_V}{n-v} = \frac{\widehat{SS}_V}{v-1'},$$
(10)

where the estimated mean square has, under  $H_0$ , an approximated  $\chi^2(v-1,0)/(v-1)$  distribution.

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Square
Treatments	v-1	$\widehat{\mathrm{SS}}_V = \hat{\tau}'_* X_1' \hat{\boldsymbol{V}}_*^{-1} X_1 \hat{\tau}_*$	$\widehat{\mathrm{MS}}_V = \frac{\widehat{\mathrm{SS}}_V}{(v-1)}$
Residuals	n - v	$\widehat{SS}_{R} = y'_{*} [\hat{V}_{*}^{-1} - \hat{V}_{*}^{-1} X_{1} (X'_{1} \hat{V}_{*}^{-1} X_{1})^{-1} X'_{1} \hat{V}_{*}^{-1}] y_{*} =$	1
		= n - v	
Total	n-1	$\widehat{\mathrm{SS}}_T = \boldsymbol{y}'_* \boldsymbol{\hat{V}}_*^{-1} \boldsymbol{y}_*$	_

Table 1. Analysis of variance for an experiment in an NRC design.

### 3.3. Hypothesis for a Set of Contrasts—Two Ways

Rejecting a general hypothesis (7) usually generates more questions and thus more hypotheses. To begin with, consider one hypothesis concerning a set of contrasts (or a single contrast) among treatment parameters

$$H_{0,\mathrm{L}}: \boldsymbol{U}_{\mathrm{L}}^{\prime}\boldsymbol{\tau}_{*} = \boldsymbol{0}, \quad \text{where} \quad \boldsymbol{U}_{\mathrm{L}}^{\prime}\boldsymbol{1}_{v} = \boldsymbol{0}. \tag{11}$$

Caliński et al. [13] indicate that the BLUE of  $U'_{\rm L} \tau_*$  is of the form

$$\boldsymbol{U}_{\mathrm{L}}'\hat{\boldsymbol{\tau}}_{*} = \boldsymbol{U}_{\mathrm{L}}'\hat{\boldsymbol{\tau}} = \boldsymbol{U}_{\mathrm{L}}'(\boldsymbol{X}_{1}'\boldsymbol{V}_{*}^{-1}\boldsymbol{X}_{1})^{-1}\boldsymbol{X}_{1}'\boldsymbol{V}_{*}^{-1}\boldsymbol{y}_{*}.$$
(12)

and the relevant sum of squares can then be obtained in the form

$$SS(\boldsymbol{U}_{L}) = \hat{\boldsymbol{\tau}}_{*}^{\prime} \boldsymbol{U}_{L} [\boldsymbol{U}_{L}^{\prime} (\boldsymbol{X}_{1}^{\prime} \boldsymbol{V}_{*}^{-1} \boldsymbol{X}_{1})^{-1} \boldsymbol{U}_{L}]^{-} \boldsymbol{U}_{L}^{\prime} \hat{\boldsymbol{\tau}}_{*}, \qquad (13)$$

with the d.f. equal to rank( $U_L$ ).

For example, if a factorial experiment is analyzed, say with two factors:  $F_1$  with  $F_1$  levels and  $F_2$  with  $F_2$  levels with a cross structure, then we will usually be interested in a partition for ANOVA treatment sums of squares for the sum of squares for the factor  $F_1$ ,

the sum of squares for the factor  $F_2$  and the sum of squares related to the interaction  $F_1F_2$  such that the following relation holds:

$$SS(\boldsymbol{u}_{F_1}) + SS(\boldsymbol{u}_{F_2}) + SS(\boldsymbol{u}_{F_1F_2}) = SS_V.$$
 (14)

In this case, we can use matrices  $U'_L$  for  $L \in \{F_1, F_2, F_1F_2\}$  of the form  $U'_{F_1} = (I_{f_1} - \frac{1}{f_1}\mathbf{1}_{f_1}\mathbf{1}'_{f_1}) \otimes \frac{1}{f_2}\mathbf{1}_{f_2}\mathbf{1}'_{f_2}$ ,  $U'_{F_2} = \frac{1}{f_1}\mathbf{1}_{f_1}\mathbf{1}'_{f_1} \otimes (I_{f_2} - \frac{1}{f_2}\mathbf{1}_{f_2}\mathbf{1}'_{f_2})$ ,  $U'_{F_1F_2} = (I_{f_1} - \frac{1}{f_1}\mathbf{1}_{f_1}\mathbf{1}'_{f_1}) \otimes (I_{f_2} - \frac{1}{f_2}\mathbf{1}_{f_2}\mathbf{1}'_{f_2})$ . A similar partition of the treatment sum of squares for the near-factorial experiment is shown in Section 4. As can be checked, it satisfies the condition  $U'_L(X'_1V^{-1}_*X_1)^{-1}U_{L^*} = \mathbf{0}$  for  $L \neq L^*$  necessary for the partition SS<sub>V</sub>, as in (14) [12].

Sometimes we will be interested in even more detailed conclusions. Currently, a letterbased representation of all-pairwise comparisons is a universal complement of the ANOVA and a standard in the presentation of results. For the presented analytical procedure, no convenient solution to this issue has yet been presented. The answer to this problem may be based on the analysis of all (v - 1)v/2 contrasts between pairs of treatments. However, it should be remembered that this is not a set of orthogonal contrasts, therefore the comparisons should not be treated as simultaneous. Letter representation is obtained using the Piepho [34] algorithm and the multcompView package in R [35], as also suggested by Lenth [36]. This package has been implemented for the analytical methodology presented above. Based on all-pairwise significance statements (*P* values) for *v* treatments obtained during the verification of the hypotheses (11) for all simple contrasts, a familiar letter display was obtained, where as usual treatments that do not differ significantly share a common letter. A detailed analytical procedure that enables the direct application of the proposed methodology for any data set for an NRC design is given in Appendix A.4.

#### 4. Application for a Near-Factorial Experiment

#### 4.1. Experimental Setup

Ratajkiewicz et al. [37] analyzed a field experiment carried out in Poznań, Poland in 2011. Studies were conducted to improve the use of fungicides against potato late blight (Phytophthora infestans (Mont.) De Bary) (PLB) in tomato processing. Based on the infestation of the plants, a determination was made of the influence of the constant (300 L/ha, SV300) and variable (PSV) spray volume and adjuvants when alternating azoxystrobin and chlorothalonil for coarse spraying with the IDKT12003 twin-jet induction nozzle. The variable spray volume (PSV) was calculated from the number of leaves per plant according to the model presented in the paper [38]. In accordance with the methodology described in the article, two fungicides were used in the study alternatively: azoxystrobin (methyl (E)-2-[2-[6-(2-cyanophenoxy)pyrimidine-4-yl]oxyphenyl]-3- methoxyprop-2-enoate, Amistar 250 SC; Syngenta Ltd., Guildford, UK) at a dose of 125 g/ha (50% of the recommended dose) or chlorothalonil (2,4,5,6-tetrachlorobenzene-1,3-dicarbonitrile, Gwarant 500 SC; Arysta LifeScience S.A.S., Noguères, France) at a rate of 625 g/ha (50% of the recommended dose). The fungicide treatment began with azoxystrobin. The first spraying was carried out five and four days after the small pale green change characteristic of PLB was observed on the first leaf.

One of two adjuvants, Slippa (Interagro Ltd., Great Notley, UK) containing poly(alkylene oxide)-modified heptamethyltrisiloxane (PMH) (655 g/L) or Torpedo II (De Sangosse Ltd., Swaffham Bulbeck, UK) was added to the fungicide slurry. The last adjuvant comprises alkoxylated tallowamine 210 g/kg, alcohol alkoxylate 380 g/kg, natural fatty acids 75 g/kg and polyalkylene glycol 210 g/kg; it is therefore defined as a multi-component adjuvant. Both substances were used at a concentration of 1 mL/L. The study considered the case when no adjuvant was added to the herbicides. Furthermore, included is the standard control treatment, i.e., no protective procedure, called (by Yates [1937]) a dummy treatment, also known as "witness" (in French) [39].

Therefore, the experiment investigated v = 7 treatments: the control treatment and 6 experimental combinations of two spray systems (SV300 and PSV)–factor F<sub>1</sub>–and adjuvant

additive (NO, MULTI or PMH)–factor  $F_2$ . This type of arrangement gives a treatment structure called a near-factorial experiment [14,15]. Numerous construction methods for this type of experimental situation in the case of NRC designs are given by Bailey and Lacka [14] and Bose and Mukerjee [15], and it is from the first of these papers from which the plan for the distribution of treatments on experimental units comes—see Figure 3.

						_							TREATMEN	г sv	Adjuvant
6	1	3	0	5	2		0	2	3	6	4	1	Т	factor F <sub>1</sub>	factor F <sub>2</sub>
5	4	0	1	6	3		6	3	2	5	0	4	0 (control)	_	_
3	0	1	5	2	4		1	6	5	4	2	0	1	SV300	NO
2	5	4	6	0	1	1	3	1	4	0	5	6	2	SV300	MULTI
4	2	6	2	1	0	1	2	0	1	2	6	E	3	SV300	PMH
4	3	0			0		2	0	1	3	0	5	4	PSV	NO
0	2	5	3	4	6		4	5	0	1	3	2	5	PSV	MULTI
													6	PSV	PMH

**Figure 3.** Scheme of distribution of treatments 0–6 on experimental units of the NRC design with b = 2 blocks. Each block has  $r_0 = 6$  rows and  $c_0 = 6$  columns. The number at the intersection of a row and column indicates the treatment used in that plot. Treatment numbers are consistent with the legend on the right side of the diagram.

#### 4.2. Statistical Analysis and Results

Based on the infestation results obtained in 2011, data were generated (Table 2) and analyzed according to the methodology described in Section 3.2 (see Appendix A). The experimental field was divided into b = 2 blocks with  $c_0 = r_0 = 6$  perpendicular rows and columns in each. The individual plot observations obtained for the combinations of the levels of factors  $F_1$  and  $F_2$  or the control treatment, taking into account their location in the rows and columns of each block, are shown in Table 2. Analyzing this experiment, we are interested in the verification of the general hypothesis (7) as well as determining the effectiveness of the applied protective procedures ( $H_{0,C} : \mathbf{U}'_C \boldsymbol{\tau}_* = \mathbf{0}$ ) and the significance of the main effects of the levels of factor  $F_1$  ( $H_{0,F_1} : \mathbf{U}'_{F_1} \boldsymbol{\tau}_* = \mathbf{0}$ ), the main effects of the levels of factors  $F_2$  ( $H_{0,F_2} : \mathbf{U}'_{F_2} \boldsymbol{\tau}_* = \mathbf{0}$ ), and the interaction effects of these two factors ( $H_{0,F_1F_2} : \mathbf{U}'_{F_1F_2} \boldsymbol{\tau}_* = \mathbf{0}$ ). These questions will be answered by estimating and testing certain sets of treatment parametric functions, which can be defined as follows:

$$\begin{bmatrix} 0 & \mathbf{0}_{6}' \\ \mathbf{0}_{6} & \left(I_{2} - \frac{1}{2}\mathbf{1}_{2}\mathbf{1}_{2}'\right) \otimes \frac{1}{3}\mathbf{1}_{3}\mathbf{1}_{3}' \end{bmatrix} \boldsymbol{\tau} = \boldsymbol{U}_{F_{1}}'\boldsymbol{\tau} \equiv \boldsymbol{U}_{F_{1}}'\boldsymbol{\tau}_{*},$$

$$\begin{bmatrix} 0 & \mathbf{0}_{6}' \\ \mathbf{0}_{6} & \frac{1}{2}\mathbf{1}_{2}\mathbf{1}_{2}'\otimes \left(I_{3} - \frac{1}{3}\mathbf{1}_{3}\mathbf{1}_{3}'\right) \end{bmatrix} \boldsymbol{\tau} = \boldsymbol{U}_{F_{2}}'\boldsymbol{\tau} \equiv \boldsymbol{U}_{F_{2}}'\boldsymbol{\tau}_{*},$$

$$\begin{bmatrix} 0 & \mathbf{0}_{6}' \\ \mathbf{0}_{6} & \left(I_{2} - \frac{1}{2}\mathbf{1}_{2}\mathbf{1}_{2}'\right) \otimes \left(I_{3} - \frac{1}{3}\mathbf{1}_{3}\mathbf{1}_{3}'\right) \end{bmatrix} \boldsymbol{\tau} = \boldsymbol{U}_{F_{1}F_{2}}'\boldsymbol{\tau} \equiv \boldsymbol{U}_{F_{1}F_{2}}'\boldsymbol{\tau}_{*},$$

$$\frac{\sqrt{10}}{2} \begin{bmatrix} 6 & -\mathbf{1}_{1}' \end{bmatrix} \boldsymbol{\tau} = \boldsymbol{U}_{1}' \circ \boldsymbol{\tau} = \boldsymbol{U}_{1}' \circ \boldsymbol{\tau}_{*},$$

$$\frac{1}{6} \begin{bmatrix} 6 & -\mathbf{1}_6' \end{bmatrix} \boldsymbol{\tau} = \boldsymbol{U}_C' \boldsymbol{\tau} \equiv \boldsymbol{U}_C' \boldsymbol{\tau}_*,$$

All these linear functions can be seen as contrasts of treatment parameters. For each of these four sets of contrasts, say  $U_L$ , the BLUE is obtainable according to Formula (12), and the relevant sum of squares, SS( $U_L$ ), follows from (13). It can also be checked that

$$SS(\boldsymbol{U}_{C}) + SS(\boldsymbol{U}_{F_{1}}) + SS(\boldsymbol{U}_{F_{2}}) + SS(\boldsymbol{U}_{F_{1}F_{2}}) = SS_{V},$$
(15)

Now one can proceed to the general ANOVA (see Appendices A.1–A.3) and its partition into four components related to the sets of contrasts as in Table 3.

**Table 2.** Experimental data for a field experiment concerning the infestation of tomato plants by potato late blight (*Phytophthora infestans* (Mont.) De Bary). The experiment was carried out according to the scheme in Figure 3.

Т	Block	Row <sub>NR</sub>	Col. <sub>NR</sub>	Obs.	Т	Block	Row <sub>NR</sub>	Col. <sub>NR</sub>	Obs.
6	1	1	1	71.960	0	2	7	7	91.812
1	1	1	2	77.501	2	2	7	8	73.984
3	1	1	3	67.575	3	2	7	9	54.820
0	1	1	4	94.480	6	2	7	10	61.700
5	1	1	5	75.321	4	2	7	11	70.860
2	1	1	6	86.494	1	2	7	12	71.447
5	1	2	1	71.806	6	2	8	7	65.981
4	1	2	2	74.680	3	2	8	8	49.203
0	1	2	3	95.943	2	2	8	9	67.858
1	1	2	4	67.492	5	2	8	10	60.500
6	1	2	5	60.864	0	2	8	11	91.317
3	1	2	6	77.532	4	2	8	12	73.392
3	1	3	1	69.535	1	2	9	7	73.379
0	1	3	2	91.750	6	2	9	8	59.850
1	1	3	3	73.470	5	2	9	9	59.127
5	1	3	4	68.000	4	2	9	10	62.750
2	1	3	5	78.925	2	2	9	11	70.702
4	1	3	6	78.283	0	2	9	12	91.160
2	1	4	1	88.079	3	2	10	7	65.208
5	1	4	2	65.770	1	2	10	8	67.356
4	1	4	3	78.756	4	2	10	9	57.450
6	1	4	4	70.900	0	2	10	10	90.479
0	1	4	5	94.250	5	2	10	11	59.904
1	1	4	6	83.996	6	2	10	12	62.756
4	1	5	1	86.975	2	2	11	7	76.406
3	1	5	2	65.698	0	2	11	8	87.267
6	1	5	3	69.600	1	2	11	9	61.321
2	1	5	4	74.124	3	2	11	10	64.752
1	1	5	5	75.596	6	2	11	11	58.880
0	1	5	6	100.000	5	2	11	12	62.284
0	1	6	1	100.000	4	2	12	7	62.861
2	1	6	2	78.192	5	2	12	8	58.007
5	1	6	3	64.362	0	2	12	9	89.048
3	1	6	4	59.416	1	2	12	10	65.904
4	1	6	5	66.512	3	2	12	11	63.815
6	1	6	6	82.500	2	2	12	12	78.917

Table 3. Analysis of variance for the sets of contrasts.

Source	Degrees of Freedom	Sum of Squares	Mean Square	Ê	p Value
Treatments	6	450.024	75.004	75.004	< 0.0001
F <sub>1</sub>	1	14.3922	14.3922	14.3922	0.00015
F <sub>2</sub>	2	35.9117	17.9558	17.9558	< 0.0001
$F_1F_2$	2	35.3518	17.6759	17.6759	< 0.0001
С	1	364.368	364.368	364.368	< 0.0001
Residuals	65	65	1		
Total	71	515.0241			

The critical values, at the 52.099 for 6 d.f., 3.841 for 1 d.f. and 2.996 for 2 d.f. The results presented in Table 3 were obtained with the use of the empirical estimates  $\tilde{\tau}$  and  $\tilde{\tau}_*$ , i.e., based on  $\hat{\sigma}_1^2 = 15.726$ ,  $\hat{\sigma}_2^2 = 9.487$ ,  $\hat{\sigma}_3^2 = 93.042$  and  $\hat{\sigma}_4^2 = 1282.51$ . The elements of vectors  $\tilde{\tau}$  and  $\tilde{\tau}_*$  are presented in Table 4.

The estimates  $\hat{\sigma}_i^2$  for i = 1, ..., 4 were obtained by solving Equation (8) using the iterative method (after five iteration steps, out of 20 performed, convergence was obtained) with the matrix  $V_*^{-1}$  as suggested in Section 3.2.

The analysis of variance presented in Table 3 enabled rejection of the general hypothesis (7). The hypotheses concerning the factor  $F_1$ , the factor  $F_2$  and their interaction were also rejected. The verification of the hypothesis  $H_{0,C}$  enabled determination of the effectiveness of the applied protective procedures. In this situation, the inference should be supplemented with detailed results, that is, a letter-based representation of all-pairwise comparisons (see Section 3.3). The contrasts used were normalized with respect to the treatment replication matrix, and the results are presented in Table 4.

**Table 4.** Empirical estimates of  $\tau$  and  $\tau_*$ . Treatments that do not differ significantly share a common letter.

TREATMENT T	SV Factor F <sub>1</sub>	Adjuvant Factor F <sub>2</sub>	Treatment Effect $ ilde{ au}$	Main Effect $ ilde{ au}_*$	
0 (control)	_	_	93.125	(19.948)	а
1	SV300	NO	72.328	(-0.850)	С
2	SV300	MULTI	77.398	(4.221)	b
3	SV300	PMH	63.682	(-9.496)	d
4	PSV	NO	70.527	(-2.651)	с
5	PSV	MULTI	65.201	(-7.977)	d
6	PSV	PMH	65.993	(-7.185)	d

#### 5. Discussion and Conclusions

The high costs of conducting experiments should encourage the search for solutions enabling the effort put in to lead to success, understood by many researchers as the detection of significant differences between the studied treatments. When planning experiments, we make a number of decisions to solve the problems of which we are aware and those that we can only expect. These decisions will always influence the method of analyzing the results of the experiment. This paper recalls the principles that should be followed when planning an experiment. It is proposed to use a nested row–column design, which has wide practical application in agricultural experiments, and may be treated as an insurance policy against missed blocking. Thanks to the new, direct approach to the analysis of variance, inference for this type of experiment becomes simple and intuitive.

The methodology used here results from the use in the estimation and hypotheses testing procedures, of dispersion matrix not in the form

$$\mathbf{V} = \sigma_1^2 [\boldsymbol{\phi}_1 + (\sigma_2^2 / \sigma_1^2) \boldsymbol{\phi}_2 + (\sigma_3^2 / \sigma_1^2) \boldsymbol{\phi}_3 + (\sigma_4^2 / \sigma_1^2) \boldsymbol{\phi}_4 + (\sigma_5^2 / \sigma_1^2) \boldsymbol{\phi}_5] = \sigma_1^2 \mathbf{F}_5$$

usually applied in the literature, as Kala [25] mentions, but in its original form (3) which ensures that  $E(SS_R) = n - v$ .

Another feature of the proposed approach concerns the simplification of the analytical procedures. One of the resulting advantages is the reduction of the number of stratum variances involved from 5 to 4. This widely simplifies the computations.

As Kala [25] notes, Equation (8) coincide-as it was observed by Patterson and Thompson [40]—with the equations following from their maximum likelihood approach under the assumption that *y* has a multivariate normal distribution, with such an advantage, that here iteration is distribution free as it uses only geometrical arguments. The proposed approach ignores within-stratum analyses based on the stratum submodels and involved combined analysis. It is shown how to perform the analysis also in the case of factorial and near-factorial structures of treatments. The main advantage of the proposed approach is that the ANOVA results can be obtained directly rather than by first performing stratum analyses and then combining their results, as in the classic approach [23,24].

As shown in the publications [10–13] and presented in Table 1, the proposed procedures lead to the reduction of the residual sum of squares to n - v, which in turn leads to the form of the *F* statistic expressed by the Formula (10). This can be seen as an advantage for the approximation of the relevant distribution.

In the applicative approach, the estimated mean square has (under  $H_0$ ) approximately the distribution of  $\chi^2(v-1,0)/(v-1)$ . The greater is *n*, i.e., the size of the experiment, the closer this approximation will be. Thus, the test of hypothesis  $H_0$  and the relevant critical level of significance are regarded as approximate. These conclusions correspond to those presented by Volaufova [41] and Johnson et al. [42].

The paper shows how to obtain a letter-based representation of all-pairwise comparisons using simple contrasts. The letter display, by far the most popular method for reporting the result of mean comparisons in all areas of research where ANOVA procedures are used, is now available after the analysis of variance under a derived mixed model in an NRC design is performed directly. All of this gives a complete set of tools that make it easy to analyze any NRC experiment. Appendix A provides all the necessary tools in R [43], and the example of a field experiment shows how the proposed procedure is performed in practice. An example of the data including a file structure compliant with the R code (Appendix A) is available in Table S1 in Supplementary Materials.

Finally, this approach can be applied to various classes of designs with the OBS property, in particular those shown in Table 5.

As previously noted, sometimes it turns out following an experiment that the applied approach needs to be reviewed. If the estimated row- (or column-) related variance components are small, close to zero, then the analysis reverts to the model for a simpler block structure—a nested block design. If both the row and column variances are close to zero, then the experiment should be run as for a proper block design [8]. Finally, if the researcher has data only from one block, the analysis will follow the model for a row–column design.

Since the starting point for the analysis is the NRC design, each of the possible modified analyses will be carried out for a design with an orthogonal block structure, which makes it possible to apply the direct approach [10–13]. As the practical application of this analytical approach depends on the structure of the experimental design, Table 5 shows how the model of the NRC design will change when some block structures are omitted. Of course, for new designs, the structure of the dispersion matrix *V* changes and therefore a different matrix  $V_*$  should be used for each of them. Next, the analytical procedure will be analogous to that in Section 3.2. The notation used in Table 5 is based on the notation of the output NRC design, so only some of the  $\sigma_i^2 \phi_i$  components appear in the matrix *V*. It should be noted here that since the roles of rows and columns in the NRC design can be changed, the nested block design is presented only for the case of the omission of the NRC's column system. Usually the system we call "blocks" (by analogy with the NRC design) is called "superblocks" for a nested block design, while our "rows" are called "blocks".

**Table 5.** Parameters and models of experimental designs obtained if certain block structures of the NRC design are omitted. The starting point is the design shown in Figure 2. The omitted block structures correspond to the deleted elements of the base model (1).



**Supplementary Materials:** The following are available online at https://www.mdpi.com/2073-4 395/11/12/2406/s1, Table S1. TOMATOES\_DATA.csv—experimental data for the field experiment analyzed in Section 4. The file structure follows the R code available in Appendix A.

Funding: This research received no external funding.

**Data Availability Statement:** The data used in the statistical analysis are included as supplementary material, and the R code is included as part of the manuscript.

Conflicts of Interest: The author declares no conflict of interest.

### 14 of 17

## Abbreviations

The following abbreviations are used in this manuscript:

NRC design	Nested row–column design
ANOVA	Analysis of variance
d.f.	Degrees of freedom
OBS	Orthogonal block structure
BLUE	Best linear unbiased estimator
SV	Spray volume

#### Appendix A

The following procedure is used to carry out analyses in R in accordance with the methodology presented in the paper. The packages psych, Matrix, matlib, multcompView and utilsIPEA are used here [35,44–47]. To analyze data other than those available in the file "TOMATOES \_DATA.csv", the structure of this data set should be maintained. Note that the control treatment in this file is numbered 7, while the remaining treatments are numbered as shown in Figure 3. All parts of the code should be treated as a whole, because subsequent pieces of the code are based on the calculations from the previous parts.

Appendix A.1. Basic: Parameters, Elements of the Design Structure and Auxiliary Markings

```
data<-read.csv2("TOMATOES_DATA.csv",h=T)</pre>
 dat <- as. data.frame(data)
 dat <- within (dat.
       BLOCK <- factor(BLOCK)
        ROW_NR <- factor (ROW_NR)
        COL_NR<- factor (COL_NR)
        TREATMENT<- factor (TREATMENT)
 3)
v<-nlevels (dat$TREATMENT) # NUMBER OF TREATMENTS
b<-nlevels(dat$BLOCK) # NUMBER OF BLOCKS
r<-nlevels (dat$ROW_NR) # NUMBER OF ROWS
                                            # w NUMBER OF ROWS PER BLOCK
r0 < -r/b
c<-nlevels(dat$COL_NR) # NUMBER OF COLUMNS
c0 < -c/b
                                            # NUMBER OF COLUMNS IN THE BLOCK
n<-b*r0*c0 # NUMBER OF EXPERIMENTAL UNITS
n0<-n/b # NUMBER OF EXPERIMENTAL UNITS PER BLOCK
 Y<-as.vector(dat$OBSERVATION) # VECTOR OF OBSERVATIONS
F1 <- with(dat, outer(BLOCK, levels(BLOCK), '==')*1)
colnames(F1) <- paste("BLOCK", sep="=", levels(dat$BLOCK)) # DESIGN MATRIX X_B
F2 <- with(dat, outer(ROW_NR, levels(ROW_NR), '==')*1)</pre>
                                                                                                                                                                                             ('==')*1)
colnames(F2) <- paste("ROW", sep="=", levels(dat$ROW_NR)) # DESIGN MATRIX X_R(B)
 F3 <- with (dat, outer (COL_NR, levels (COL_NR), '==')*1)
colnames(F3) <- paste("COL", sep ="=", levels(dat$COL_NR)) # DESIGN MATRIX X_C(B)
 F4 <- with (dat, outer (TREATMENT, levels (TREATMENT), '==')*1)
colnames (F4) <- paste ("TREATMENT", see ="=", levels (dat$TREATMENT)) # DESIGN MATRIX X_1
wr<-t(F4)%*%as.vector(matrix(1,nrow=n)) # VECTOR OF TREATMENT REPLICATIONS
 wr_minus<-(wr)^(-1) # VECTOR OF INVERSE TREATMENT REPLICATIONS
 rDelta <- t (F4)%*%F4 # TREATMENT REPLICATION MATRIX
 rDelta_minus <- solve (rDelta)
 In <- diag(nrow=n)
Iv <- diag(nrow=v)
w1<-as.vector(matrix(1,nrow=n))
 w1v<-as.vector(matrix(1,nrow=v))
 Fi1 < -In - (c0)^{(-1)} + F2\%*(t(F2) - (r0)^{(-1)} + F3\%*(t(F3) + (n0)^{(-1)} + F1\%*(t(F1) # Phi1 MATRX)^{(-1)} + F1\%*(t(F1) + F1)^{(-1)} + F1)^{(-1)} + F1\%*(t(F1) + F1)^{(
 \begin{array}{l} Fi2 < -(c0)^{(-1)*F2} \\ \# Phi2 \\ MATRIX \\ (see Equation 3) \\ Fi3 < -(r0)^{(-1)*F3} \\ \# Phi2 \\ MATRIX \\ (see Equation 3) \\ \# Phi3 \\ MATRIX \\ (see Equation 3) \\ \# Phi3 \\ \# Phi3 \\ MATRIX \\ (see Equation 3) \\ \# Phi3 \\ \# Phi3
Fi4 < -(n0)^{(-1)} F1\% t(F1) - (n)^{(-1)} w1\% t(W1) # Phi4 MATRIX (see Equation 3)
Fi5 < -(n)^{(-1)*w1\%*}(w1) # Phi5 MATRIX (see Equation 3)
```

This is where the first code check is proposed. In accordance with Section 3, the relation  $\phi_1 + \phi_2 + \phi_3 + \phi_4 + \phi_5 = I_n$  holds, which can be checked like this:

all.equal(as.matrix(round(Fi1+Fi2+Fi3+Fi4+Fi5,digits=3)), In)

#### Appendix A.2. Iteration Procedure to Determine Estimates of Stratum Variances

In the iterative procedure, 20 steps are declared. Usually this is a sufficient value. You should of course check if the iteration procedure converges—the test below then gives "TRUE". If not, the number of iteration steps should be increased.

```
sigma2M1<-1
sigma2M2<-1
sigma2M3<-1
sigma2M4<-1
sigma2M5<-1
ITERATIONS<-list()
for(i in 1:20){
   VstarMinus <- sigma2M1 * Fi1 + sigma2M2 * Fi2 + sigma2M3 * Fi3 + sigma2M4 * (Fi4 + Fi5)
   V_Minus < -VstarMinus + (sigma2M5 - sigma2M4) * (n)^{(-1)*w1\%*}t(w1)
   Dodatkiwa<-V_Minus-VstarMinus
   all<-Dodatkiwa[1,1]
   sigma2M5<-n*a11+sigma2M4
   sigma2_5 <-1/(sigma2M5)</pre>
   P<-F4%*%solve(t(F4)%*%V_Minus%*%F4)%*%t(F4)%*%V_Minus
   d1 < -tr(Fi1\%((In-P)))
   d2<-tr (Fi2%*%(In-P))
   d3 < -tr(Fi3\%*(In-P))
   d4<-tr(Fi4%*%(In-P))
   sigma2_1 < -(d1)^{(-1)*t} ((Fi1\%*((In-P)))\%*(Y))\%*((Fi1\%*((In-P)))\%*(Y))

sigma2_2 < -(d2)^{(-1)*t} ((Fi2\%*((In-P)))\%*(Y))\%*((Fi2\%*((In-P)))\%*(Y))
   \begin{split} & \text{sigma2_3} < -(\text{d3}) \land (-1) * \mathsf{t} ((\text{Fi3}\%*\%(\text{In}-P))\%*\%Y) \%*\%((\text{Fi3}\%*\%(\text{In}-P))\%*\%Y) \\ & \text{sigma2_4} < -(\text{d4}) \land (-1) * \mathsf{t} ((\text{Fi4}\%*\%(\text{In}-P))\%*\%Y) \%*\%((\text{Fi4}\%*\%(\text{In}-P))\%*\%Y) \\ & \text{sigma2M1} < -((\text{sigma2_1}) \land (-1))[1,1] \end{split} 
  sigma2M2<-((sigma2_2)^(-1))[1,1]
sigma2M3<-((sigma2_3)^(-1))[1,1]
   sigma2M4<-((sigma2_4)^(-1))[1,1]
   ITERATIONS[[i]]<-c(sigma2_1,sigma2_2,sigma2_3,sigma2_4,sigma2_5)
ITER<-matrix (as.numeric (unlist (ITERATIONS)), 20, 5, byrow=T)
```

all.equal((as.matrix(round(ITER[20,]-ITER[19,],digits=10))), matrix(0,nrow=5)) # TEST

*Appendix A.3. General Solutions and ANOVA Table* 

```
tau <- solve(t(F4)%*%V_Minus%*%F4)%*%t(F4)%*%V_Minus%*%Y # tau empirical estimates
tau_star <- (Iv - (n)^{(-1)*w1v}(wr))^{*}t(wr))^{*}tau # tau* empirical estimates
SSv<-t(tau_star)%*%t(F4)%*%V_Minus%*%F4%*%tau_star
SSr <- t (Y)%*%V_Minus%*%(In -P)%*%Y
y_star < -(In - (1/n) * w1\% * (w1))\% * \%Y
SSt <- t (y_star)%*%VstarMinus%*%y_star
F_general < -(n-v) * SSv / ((v-1) * SSr)
p_value<-pchisq(F_general, df=v-1, lower.tail=FALSE)/(v-1)
ANOVA_TABLE <- matrix (c(v-1,n-v,n-1,round(SSv,digits=4),round(SSr),
                            round(SSv+SSr, digits=4),
                            round (SSv/(v-1), digits = 4), round (SSr/(n-v)), '
round (F_general, digits = 4), ' ', ' '), ncol=4)
colnames (ANOVA_TABLE)<- c("df", "Sum_of_squares", "Mean_square", "F")</pre>
rownames(ANOVA_TABLE)<- c('Treatments', 'Residuals','Total')
ANOVA_TABLE<-as.table(ANOVA_TABLE)
ANOVA_TABLE
p_value
```

Appendix A.4. Letter-Based Representation of All-Pairwise Comparisons

```
combin<-combin(c(1,2,3,4,5,6,7), 2, simplify = TRUE)
zo<- matrix(0,v,v*(v-1)/2)
for(i in 1:(v*(v-1)/2)){
    k<-combin[[1,i]]
    l<-combin[[2,i]]
    zo[k,i] <- 1
    zo[l,i] <- -1
}
wspl<- matrix(0,v*(v-1)/2,v*(v-1)/2)
for(i in 1:(v*(v-1)/2)){
    k<-combin[[1,i]]</pre>
```

```
l<-combin[[2,i]]</pre>
  wsp1[i,i]<-(1/sqrt(1/rDelta[k,k]+1/rDelta[l,l]))
col_names < -matrix(0,1,v*(v-1)/2)
for (i in 1:(v*(v-1)/2)){
  col_names[1,i]<-paste(combin[1,i], combin[2,i], sep = "-")</pre>
col_names<-as.vector(col_names)
simple_contrasts <- zo‰ wsp1 # DEFINING SIMPLE CONTRASTS</pre>
rownames(simple_contrasts)<- paste("tr", sep ="=", levels(dat$TREATMENT))</pre>
p_values_simple <-as.vector(matrix(0,1,v*(v-1)/2))
for (i in 1:(v*(v-1)/2)){
  ci <- simple_contrasts [, i]</pre>
  SS_ci<-t(tau_star)%*%ci%*%solve(t(ci)%*%solve(t(F4)%*%VstarMinus%*%F4)%*%ci)%*%
  t(ci)%*%tau star
  p_values_simple[i]<-pf(q=SS_ci, df1=1, df2=Inf, lower.tail=FALSE)</pre>
names(p_values_simple)<-col_names</pre>
Letters <- multcompLetters (p_values_simple)
Letters
           #### Letter-based representation
```

### References

- 1. Fisher, R.A. Statistical Methods for Research Workers, 11th ed. rev.; Oliver and Boyd: Edinburgh, UK, 1925.
- 2. Fisher, R.A. The arrangement of field experiments. J. Minist. Agric. 1926, 33, 503–515.
- Yates, F. A fresh look at the basic principles of the design and analysis of experiments. In Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Statistical Laboratory, University of California, Berkeley, CA, USA, 21 June–18 July 1965 and 27 December 1965–7 January 1966; University of California Press: Berkeley, CA, USA, 1967; pp. 777–790.
- 4. Hinkelmann, K.; Kempthorne, O. Design and Analysis of Experiments. Vol. 1. Introduction to Experimental Design, 2nd ed.; Wiley-Interscience: New York, NY, USA, 2008.
- Casler, M.D. Fundamentals of Experimental Design: Guidelines for Designing Successful Experiments. Agron. J. 2015, 107, 692–705. [CrossRef]
- 6. LeClerg, E.L.; Leonard, W.H.; Clark, A.G. Field Plot Technique, 2nd ed.; Burgess Publishing Company: Minneapolis, MN, USA, 1966.
- Grzebisz, W.; Łukowiak, R. Nitrogen Gap Amelioration Is a Core for Sustainable Intensification of Agriculture—A Concept. Agronomy 2021, 11, 419. [CrossRef]
- 8. Bailey, R.A.; Williams, E.R. Optimal nested row-column designs with specified components. *Biometrika* 2007, 94, 459–468. [CrossRef]
- Caliński, T.; Kageyama, S. Block Designs: A Randomization Approach, Vol. I: Analysis; Lecture Notes in Statistics (Volume 150); Springer: New York, NY, USA, 2000; pp. 154–196.
- 10. Caliński, T.; Siatkowski, I. On a new approach to the analysis of variance for experiments with orthogonal block structure. I. Experiments in proper block designs. *Biom. Lett.* **2017**, *54*, 91–122. [CrossRef]
- 11. Caliński, T.; Siatkowski, I. On a new approach to the analysis of variance for experiments with orthogonal block structure. II. Experiments in nested block designs. *Biom. Lett.* **2018**, *55*, 147–178. [CrossRef]
- 12. Caliński, T.; Łacka, A.; Siatkowski, I. On a new approach to the analysis of variance for experiments with orthogonal block structure. III. Experiments in row-column designs. *Biom. Lett.* **2019**, *56*, 183–213. [CrossRef]
- 13. Caliński, T.; Łacka, A.; Siatkowski, I. On a new approach to the analysis of variance for experiments with orthogonal block structure. IV. Experiments in split-plot designs. *Biom. Lett.* **2020**, *57*, 183–213. [CrossRef]
- 14. Bailey, R.A.; Łacka, A. Nested row-column designs for near-factorial experiments with two treatment factors and one control treatment. *J. Stat. Plan. Inference* **2015**, *165*, 63–77. [CrossRef]
- 15. Bose, M.; Mukerjee, R. Near-factorial experiments in nested row-column designs regulating efficiencies. *J. Stat. Plan. Inference* **2017**, *193*, 109–116. [CrossRef]
- 16. Gupta, S.; Kageyama, S. TypeS designs with nested rows and columns. Metrika 1991, 38, 195–202. [CrossRef]
- 17. Hartung, J.; Wagener, J.; Reiner, R.; Piepho, H.-P. Blocking and re-arrangement of pots in greenhouse experiments: Which approach is more effective? *Plant Methods* **2019**, *15*, 143. [CrossRef]
- 18. Kozłowska, M.; Łacka, A.; Skorupska, A. Block designs with nested rows and columns for research on food acceptability limitation. *Commun. Stat. A-Theory* **2012**, *41*, 2456–2464. [CrossRef]
- 19. Chang, J.Y.; Notz, W.I. Some optimal nested row-column designs. Stat. Sin. 1994, 4, 249–263.
- Łacka, A.; Kozłowska, M. Planning of factorial experiments in a block design with nested rows and columns for environmental research. *Environmetrics* 2009, 20, 730–742. [CrossRef]
- 21. Pearce, S. The factorial field experiment. Exp. Agric. 2005, 41, 109–120. [CrossRef]

- 22. Caliński, T.; Łacka, A. On Combining Information in Generally Balanced Nested Block Designs. *Commun. Stat. A-Theory* **2014**, *43*, 954–974. [CrossRef]
- 23. Yates, F. The recovery of inter-block information in variety trials arranged in three-dimensional lattices. *Ann. Eugen.* **1939**, *9*, 136–156. [CrossRef]
- 24. Yates, F. The recovery of inter-block information in balanced incompleteblock designs. Ann. Eugen. 1940, 10, 317–325. [CrossRef]
- 25. Kala, R. A new look at combining information from stratum submodels. In *Matrices, Statistics and Big Data, Selected Contributions from IWMS 2016;* Ahmed, S.E., Carvalho, F., Puntanen, S., Eds.; Springer: Berlin/Heidelberg, Germany, 2019; pp. 35–49.
- 26. Singh, M.; Dey, A. Block designs with nested rows and columns. Biometrika 1979, 66, 321–326. [CrossRef]
- 27. Nelder, J.A. The analysis of randomized experiments with orthogonal block structure. Proc. R. Soc. Lond. A 1965, 283, 147–178.
- 28. Kozłowska, M.; Łacka, A.; Krawczyk, R.; Kozłowski, R.J. Some block designs with nested rows and columns for research on pesticide dose limitation. *Environmetrics* **2011**, *22*, 781–788. [CrossRef]
- 29. Mejza, I.; Mejza, S. Model Building and Analysis for Block designs with Nested Rows and Columns. *Biom. J.* **1994**, *36*, 327–340. [CrossRef]
- 30. Houtman, A.M.; Speed, T.P. Balance in designed experiments with orthogonal block structure. *Ann. Stat.* **1983**, *11*, 1069–1085. [CrossRef]
- 31. Rao, C.R.; Kleffe, J. Estimation of Variance Components and Applications; North-Holland: Amsterdam, The Netherlands, 1988.
- 32. Rao, C.R. Projectors, generalized inverses and the BLUEs. J. R. Stat. Soc. Ser. B 1974, 36, 442–448.
- 33. Nelder, J.A. The combination of information in generally balanced designs. J. R. Stat. Soc. Ser. B 1968, 30, 303–311. [CrossRef]
- 34. Piepho, H.-P. An Algorithm for a Letter-Based Representation of All-Pairwise Comparisons. *J. Comput. Graph. Stat.* **2004**, *13*, 456–466. [CrossRef]
- 35. Graves, S.; Piepho, H.-P.; Selzer, L. With Help from Dorai-Raj S. multcompView: Visualizations of Paired Comparisons. R Package Version 0.1-8. 2019. Available online: https://CRAN.R-project.org/package=multcompView (accessed on 19 December 2019).
- 36. Lenth, R.V. Least-Squares Means: The R Package Ismeans. J. Stat. Softw. 2016, 69, 1–33. [CrossRef]
- Ratajkiewicz, H.; Kierzek, R.; Raczkowski, M.; Hołodyńska-Kulas, A.; Łacka, A.; Szulc, T. The effect of coarse-droplet spraying with double flat fan air induction nozzle and spray volume adjustment model on the efficiency of fungicides and residues in processing tomato. *Span. J. Agric. Res.* 2018, *16*, e1001. [CrossRef]
- Ratajkiewicz, H.; Kierzek, R.; Raczkowski, M.; Hołodyńska-Kulas, A.; Łacka, A.; Wójtowicz, A.; Wachowiak, M. Effect of the spray volume adjustment model on the efficiency of fungicides and residues in processing tomato. *Span. J. Agric. Res.* 2016, 14, e1007. [CrossRef]
- 39. Bailey, R.A. Hasse diagrams as a visual aid for linear models and analysis of variance. *Commun. Stat. A-Theory* **2020**, *50*, 5034–5067. [CrossRef]
- 40. Patterson, H.D.; Thompson, R. Recovery of inter-block information when the block sizes are unequal. *Biometrika* **1971**, *58*, 545–554. [CrossRef]
- 41. Volaufova, J. Heteroscedastic ANOVA: old p values, new views. Stat. Pap. 2009, 50, 943–962. [CrossRef]
- 42. Johnson, N.L.; Kotz, S.; Balakrishnan, N. Continuous Univariate Distributions, 2nd ed.; Wiley: New York, NY, USA, 1995; Volume 2.
- 43. R Core Team. *R: A Language and Environment for Statistical Computing;* R Foundation for Statistical Computing: Vienna, Austria, 2020. Available online: https://www.R-project.org/ (accessed on 15 February 2021).
- 44. Bates, D.; Maechler, M. Matrix: Sparse and Dense Matrix Classes and Methods. R Package Version 1.3-4. 2021. Available online: https://CRAN.R-project.org/package=Matrix (accessed on 1 June 2021).
- Coelho, G.; Mation, L. utilsIPEA: IPEA Common Functions. R Package Version 0.0.6. 2018. Available online: https://CRAN.R-project.org/package=utilsIPEA (accessed on 20 April 2018).
- Friendly, M.; Fox J.; Chalmers P. matlib: Matrix Functions for Teaching and Learning Linear Algebra and Multivariate Statistics. R Package Version 0.9.4. 2020. Available online: https://CRAN.R-project.org/package=matlib (accessed on 25 October 2020).
- Revelle, W. psych: Procedures for Personality and Psychological Research; Northwestern University: Evanston, IL, USA, 2021. Available online: https://CRAN.R-project.org/package=psychVersion=2.1.6. (accessed on 20 June 2021).