Testing Interaction and Estimating Variance Components in Block Designs - Based on a Linear Model

Karl Moder(✉)

Institute for Applied Statistics, Department of Landscape, Spatial and Infrastructure Sciences, University of Natural Resources and Life Sciences, Vienna, Austria
karl.moder@boku.ac.at

Abstract. Randomized complete block designs (RCBD) introduced by [3] are probably the most widely used experimental designs. Despite many advantages, they suffer from one serious drawback. It is not possible to test interaction effects in analysis of variance (ANOVA) as there is only one observation for each combination of a block and factor level. Although there are some attempts to overcome this problem none of these methods are used in practice, especially as most of the underlying models are non-linear. A review on such tests is given by [6] and [1].

Here a new method is introduced which permits a test of interactions in block designs. The model for RCBDs is linear and identical to that of a two factorial design. The method as such is not restricted to simple block designs, but can also be applied to other designs like Split-Plot-design, Strip-Plot-design, ... and probably to incomplete block designs.

ANOVA based on this method is very simple. Any common statistical program packages like SAS, SPSS, R, ... can be used. Although a test on interaction in two- or multi- factorial designs makes sense only for fixed and a certain class of mixed models, the proposed method can also be used for estimating variance components in any kind of block models (fixed, random, mixed) if the sample size is not too small.

1 Introduction

A Randomized complete block design is a kind of two-factorial design which is based on the model:

$$y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + e_{ijk} \tag{1}$$

y_{ijk} ... observation k at level α_i of factor A and level β_j of factor B $(k = 1, \ldots, n)$

μ ... overall mean

α_i ... effect of level i of factor A $\quad i = 1, \ldots, a$

β_j ... effect of level j of factor B $\quad j = 1, \ldots, b$

$(\alpha\beta)_{ij}$ ... interaction effect at level α_i of factor A and at level β_j of factor B

e_{ijk} ... error term at level α_i of factor A, at level β_j of factor B and at replication k $\quad k = 1, \ldots, n$.

J. Pilz et al. (Eds.): ICRA8 2019, *Mindful Topics on Risk Analysis and Design of Experiments*, pp. 135–146, 2022.
https://doi.org/10.1007/978-3-031-06685-6_10

Depending on the chosen model, mean square errors include different kinds of variance components (Table 1).

Table 1. Variance components of mean square values in a two factorial design for different kind of models

Source of variation	Fixed effects model	Random effects model	Mixed effects model
Main A	$\sigma^2 + \frac{bn}{a-1}\sum_i \alpha_i^2$	$\sigma^2 + n\sigma_{ab}^2 + bn\sigma_a^2$	$\sigma^2 + n\sigma_{ab}^2 + \frac{bn}{a-1}\sum_i \alpha_i^2$
Main B	$\sigma^2 + \frac{an}{b-1}\sum_j \beta_i^2$	$\sigma^2 + n\sigma_{ab}^2 + an\sigma_b^2$	$\sigma^2 + \kappa n\sigma_{ab}^2 + an\sigma_b^2$
Interaction $A \times B$	$\sigma^2 + \frac{n}{(a-1)(b-1)}\sum_{i,j}(\alpha\beta)_{ij}^2$	$\sigma^2 + n\sigma_{ab}^2$	$\sigma^2 + n\sigma_{ab}^2$
Error	σ^2	σ^2	σ^2

κ depends on the side condition about interaction.

$$\kappa = \begin{cases} \mathrm{cov}(ab_{ij}, ab_{ij'}) = 0 & 1 \quad (j \neq j') \\ \sum_{i=1}^{a} ab_{ij} = 0 & 0 \quad \forall\, j \end{cases}$$

As one may see from Table 1 a test on main effects in a fixed model is based on the error term, in a random model on interaction mean squares and in the mixed model approach either on the error term or the interaction term (depending on the definition of κ).

In the follow this article focuses primarily on fixed models or mixed models where $\kappa = 0$.

Several authors tried to find a solution for testing interaction in block designs. A short selection of these models can be found below.

1.1 Tukey's Test

The first one who proposed a block model which includes interaction was [12].

$$y_{ij} = \mu + \alpha_i + \beta_j + \lambda \times \alpha_i \times \beta_j + e_{ij}$$

λ ... interaction parameter

Here the interaction term is bound to the height of the factor levels which is a very strong restriction. [11] showed that Tukey's test can be derived as a test of $H_0 : \lambda = 0$.

1.2 Johnson and Graybill's Test

Another solution to the problem of interactions in block designs is given by the following model [5]:

$$y_{ij} = \mu + \alpha_i + \beta_j + \Phi \times \xi_i \times \eta_j + e_{ij}$$

Limitations to interaction are not that strict as with Tukey's model, but some additional assumptions have to be made ($\sum_{i=1}^{a} \xi_i = 0,\ \sum_{j=1}^{b} \eta_j = 0,\ \sum_{i=1}^{a} \xi_i^2 = 1,\ \sum_{j=1}^{b} \eta_j^2 = 1$).

1.3 Mandel's Test

[7] assumes a systematic type of row column interaction. The underlying "Mandel-row" model is given by

$$y_{ij} = \mu + \alpha_i + \beta_j + \gamma_i \times \beta_j + e_{ij}$$

γ_j ... interaction parameter depending on row i

Other models including non-linear interaction effects can be found by [2,4,8,9] and [10].

2 Deriving Sum of Squares for Error Term and Interaction in Block Designs

Deriving sum of squares and mean squares in a block model is based on a common two-factorial design.

2.1 Sum of Squares in Two Factorial Designs

The sum of squares value for interaction in case of a balanced design can be calculated as

$$SS_{AB} = n \sum_{i=1}^{a} \sum_{j=1}^{b} (\bar{x}_{ij.} - \bar{x}_{i..} - \bar{x}_{.j.} + \bar{x}_{...})^2.$$

with $(a-1)(b-1)$ number of degrees of freedom.

The sum of squares value for the error term is given by

$$SS_E = \sum_{i=1}^{a} \sum_{j=1}^{b} \sum_{k=1}^{n} (x_{ijk} - \bar{x}_{ij.})^2.$$

with $ab(n-1)$ number of degrees of freedom.

In block designs the number of observations (n) for each combination of a factor level and a block is equal to 1. It follows, that:

- $\bar{x}_{ij.}$ for each combination of factor and block levels is identical to x_{ij1}.
- degree of freedom for the error sum of squares becomes zero.
- $MS_E = \frac{SS_E}{df_E}$ can not be estimated.

In the follow we the mean square value for interaction ($MS_{AB} = \frac{SS_{AB}}{df_{AB}}$) serves as an error term – assuming, that no interaction effects exist. That's a very daring assumption, since almost all tow-factorial experiments show interactions (no matter whether significant or not).

2.2 Separating Error from Interaction in Block Designs

In a two factorial design the following restrictions are usually assumed in regard to interaction:

- The sum of all interaction effects within a block is equal to zero ($\sum_{i=1}^{a}(\alpha\beta)_{ij} = 0 \quad \forall j \quad j = 1, \ldots, b$).
- The sum of all interaction effects within a certain factor level is equal to zero ($\sum_{j=1}^{b}(\alpha\beta)_{ij} = 0 \quad \forall i \quad i = 1, \ldots, a$).

To separate the error term from interaction we may look at a Latin square like block design (Fig. 1)

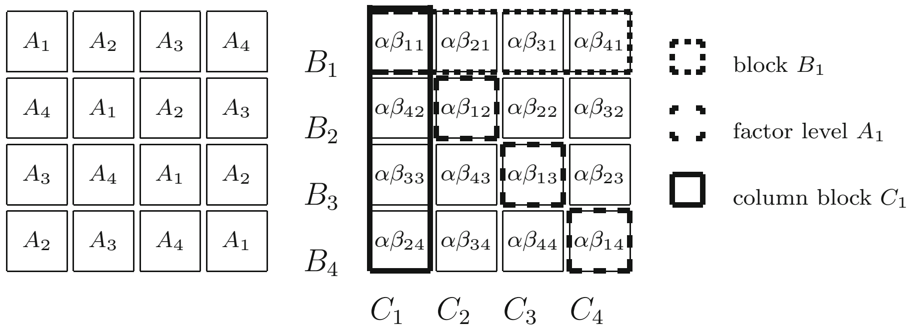

Fig. 1. Block plan and corresponding interaction terms in a Latin square like block design

Looking at Fig. 1, there is obviously no reason why one should not restrict the sum of interaction effects within a column block to zero, too

$$\sum_{i=1,j=1}^{a,b} (\alpha\beta)_{ij^k} = 0 \qquad \forall k \quad k = 1, \ldots, c. \tag{2}$$

Using this restriction has some important implications:

- Within a column block all levels of a factor are included exactly one time. As a common restriction $\sum_{i=1}^{a} \alpha_i = 0$. So the mean value of a column block does not include any factor effects.
- In Latin Square like block designs a column block comprises all levels of blocks. Usually it is assumed, that these effects sum up to zero ($\sum_{j=1}^{b} \beta_j = 0$). So the mean value of a column block does not include any block effects.

As a consequence of these two last restrictions and of restriction 2 the mean value of a column block encloses only some error effects besides μ. This enables us to separate error variance and interaction variance by calculating the sum of squares for column blocks.

$$SS_{column\ block} = SS_E = a \sum_{k=1}^{c} (\bar{x}_{c_k} - \bar{x}_{..})^2 \tag{3}$$

The error sum of squares for the error term (SS_{E^*}) in a common block design actually is the sum of squares value for interaction. So the difference zu SS_E obviously gives the sum of squares for interaction.

$$SS_{AB} = SS_{E^*} - SS_E \tag{4}$$

or alternatively $SS_{AB} = SS_T - SS_A - SS_B - SS_E$.

Degrees of freedom are defined as:

$$df_E = c - 1 \ = b - 1$$
$$df_{AB} = df_T - df_A - df_B - df_E \ = \ (a-2)(b-1)$$

Mean squares are calculated the usual way:

$$MS_E = SS_E/df_E, \qquad MS_{AB} = SS_{AB}/df_{AB}$$

Until now we restricted or model to a Latin Square design. The method as such can be applied to any kind of block design. Figure 2 depicts interaction terms in a 5×2 block design.

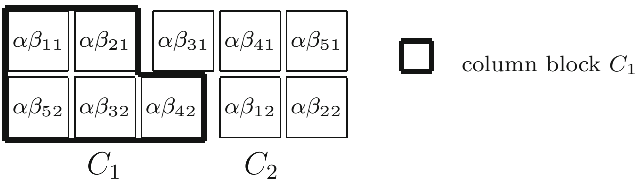

Fig. 2. Interaction terms in an arbitrary block design

Within a column block all levels of the interesting factor are included. By definition, the sum of all interaction effects within a column block is zero. The number of block effects however is different. For example, the first column block of Fig. 2 contains 2 times the effect of the first and 3 times the effect of the second block. Since we can estimate these block effects using the block means, we are able to correct for this biased column block.

3 Illustrative Example

The example in Fig. 3 shows a block design [A] with 3 blocks and 4 factor levels.

Based on the experimental design 3 column blocks were defined [B]. Each column block comprises every level of factor A. So their effects sum up to zero. Interaction effects within each column block sum up to zero too, but block effects do not, as in every column block a different number of block effects is included. So we have to correct for these effects. To do this, we need to calculate block means, factor level means, uncorrected column block means and the overall mean.

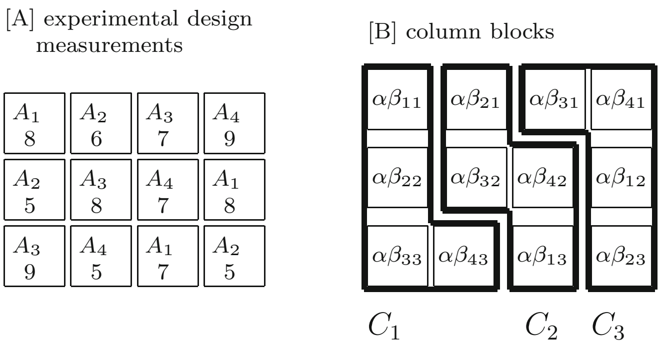

Fig. 3. Illustrative example including experimental design and corresponding measurements [A] and defined column blocks with interaction effects [B].

overall	factor $\bar{x}_{A_i}$				block $\bar{x}_{B_j}$			uncorrected col. block $\bar{x}_{C_k}$		
$\bar{\bar{x}}$	1	2	3	4	1	2	3	1	2	3
7.0	$7.\overline{6}$	$5.\overline{3}$	8.0	7.0	7.5	7.0	6.50	6.75	7.00	7.25

3.1 Calculation of Column Block Means

Each uncorrected column block mean includes a specific block effect several times (e.g. The first column block C_1 includes 2 times block effect 3, C_2 includes 2× block effect 2, C_3 includes 2× block effect 1). So we have to correct this additional block effects. The effect of a block can be calculated as the difference between the block mean and the overall mean.

If we want to calculate the mean of C_1 we have reduce the sum for this column block by the effect of block 3, to get an unbiased estimator. The effect of block 3 is calculated es the difference between $\bar{x}_{B_3} - \bar{\bar{x}}$. So an unbiased estimation of column block 1 for the example given above can be calculated based on the sum of this column block and the effect size of block 3 (ef_{B_3}) as this is included in the sum for this column block.

uncorrected sum: $\sum_{j=1}^{n} x_{1j} \;=\; 8+5+9+5 \;=\; 27$
effect size block 3: $ef_{B_3} \;=\; \bar{x}_{B_3} - \bar{\bar{x}} \;=\; 6.5 - 7.0 \;=\; -0.5$
corrected sum: $\sum_{j=1}^{n} x_{1j} - ef_{B_3} \;=\; 27 - (-0.5) \;=\; 27.5$
corrected mean: $\bar{x}_{C_1} \;=\; \frac{\sum_{j=1}^{n_f} x_{1j} - ef_{B_3}}{n_f} \;=\; \frac{27.5}{4} \;=\; 6.875$

Similarly, we can derive all other correction factors and column block means.

$\bar{x}_B - \bar{\bar{x}}$		
1	2	3
0.5	0.000	-0.5

corrected column block means $\bar{x}_C$		
1	2	3
6.875	7.000	7.125

3.2 Calculation of Sum of Squares and Degrees of Freedom

Sum of squares for blocks, factors, total and the (uncorrected) error is calculated the usual way and can be done with any statistical package for a common block design.

Sum of squares for the error term is calculated by the corrected column block means:

$$SS_E \;=\; n_f \sum_{k=1}^{n_c b} (\bar{x}_{cb_k} - \bar{\bar{x}})^2 \;=\; 4((6.875-7.0)^2 + (7.000-7.0)^2 + (7.125-7.0)^2) \;=\; 0.125$$

From evaluating a common block design we get $SS_E^* = 9.\overline{3}$. As mentioned above this actually includes possible interaction effects. By subtracting SS_E from SS_{E^*} we find the sum of squares value for interaction.

$$SS_{IA} = SS_E^* - SS_E = 9.\overline{3} - 0.125 \;=\; 9.208\overline{3}$$

Degrees of freedom for factor, block and total are those of a common block design. Degrees of freedom for main effects and interaction effects are calculated as follows:

$$\begin{aligned} df_E &= c-1 = b-1 & 3-1 &= 2 \\ df_{AB} &= df_E^* - df_E & 6-2 &= 4 \\ &= (a-2)(b-1) & (4-2)(3-1) &= 4 \end{aligned}$$

3.3 ANOVA Table

Based on calculations for a common block design and those of 3.2 we find results as shown in an ANOVA Table 2.

As can be seen from Table 2, separating interaction from the error of the common block model dramatically changes the result. If interaction is possible or can be expected, it should be included in the model regardless of whether it is significant or not.

Table 2. Example of analysis of variance for the interaction model in compare to that of a common block model

	block design with Interaction					common block design				
Effect	*df*	*SS*	*MS*	*F*	Prob	*df*	*SS*	*MS*	*F*	Prob
Factor	3	$12.66\overline{6}$	$4.222\overline{2}$	$67.5\overline{5}$	0.0146	3	$12.\overline{6}$	$4.\overline{2}$	2.71	0.1377
Block	2	2.000	1.0000	16.00	0.0588	2	2.0	1.0	0.64	0.5585
Int.act.	4	9.208	2.3021	$36.8\overline{3}$	0.0266					
Error	2	0.125	0.0625			6	$9.\overline{3}$	$1.\overline{5}$		
Total	11	24.000				11	24.0			

4 Simulations

To get an idea about the power of the interaction model several simulations were performed. A Fortran program (as well as a SAS macro and a R script) was developed to calculate ANOVA results for both the interaction model as well as for the common block model. Simulation were based on 100000 runs each.

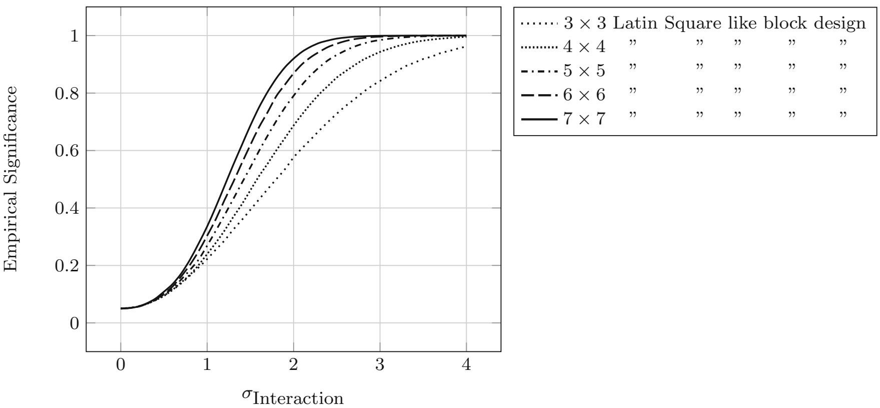

Fig. 4. Empirical power of the test on interaction as a function of the standard deviation of interaction for different sized Latin Square like block designs ($\sigma_E = 1$).

Figure 4 illustrates the empirical power of a test on interaction in 3×3 to 7×7 Latin Square like block designs. The standard deviation for this effect varied from 0 to 4 in steps of 0.5. The standard deviation for the error term (σ_E) was held constant and set to 1. As one would expect the power increases if sample size increases too. Thus, for example the power for the test on interaction in 7×7 design is about 92% if the standard deviation of the interaction is twice as high as that of the error (99.9% if it is three times as high). Whereas power in a 3×3 design is about 58% respectively 84%.

In block analysis one is primarily interested in the test regarding the main effect. Figure 5 displays the power for a 3×3 and 7×7 Latin Square like block

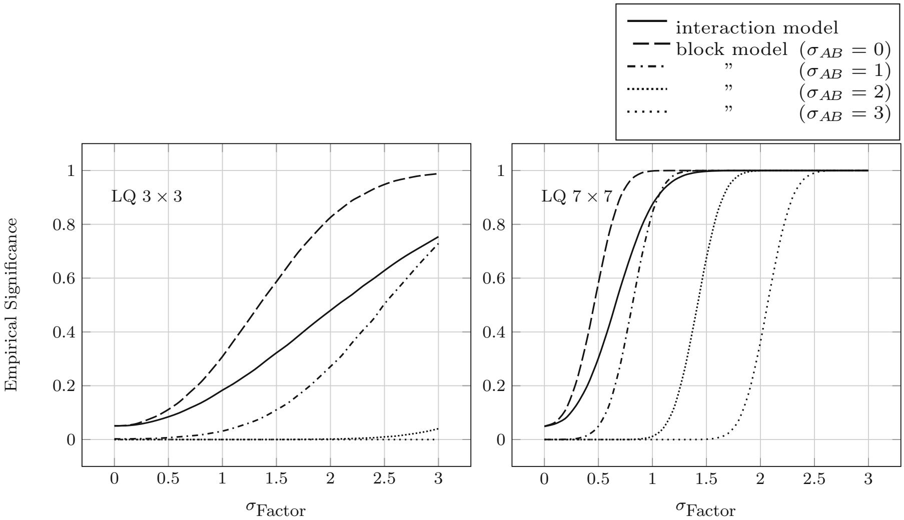

Fig. 5. Empirical power of the test on the main effect as a function of the standard deviation of this effect for 3×3 and 7×7 different Latin Square like block designs ($\sigma_E = 1$).

design. Four different standard deviations for interaction were taken into account ($\sigma_{AB} = 1, 2, 3, 4$). If there is no interaction the power for the common block analysis is highest (dashed curve). In this situation the interaction model is over parameterized and thus its power decreases. Whereas in those situations where the standard deviation of the interaction corresponds at least to that of the error, the interaction model (solid line) is best in most cases.

In addition, Fig. 5 shows another unwanted effect for the common block model. Depending on the level of interaction, there is a certain range in which it is impossible for the common block model to find any significant result for the main effect (although one awaits at least α% of significant cases even in the absence of any influences).

Results for 5×3 and 7×4 block designs in regard to the main effect are presented in Fig. 6. Again common block analysis is best in those situations were no interaction exists. Depending on the size of interaction the power of the interaction model gets more and more superior. Contrary to statistical theory, type I error rate is zero (although it should be α) in many situations where interaction exists and the main effect is low.

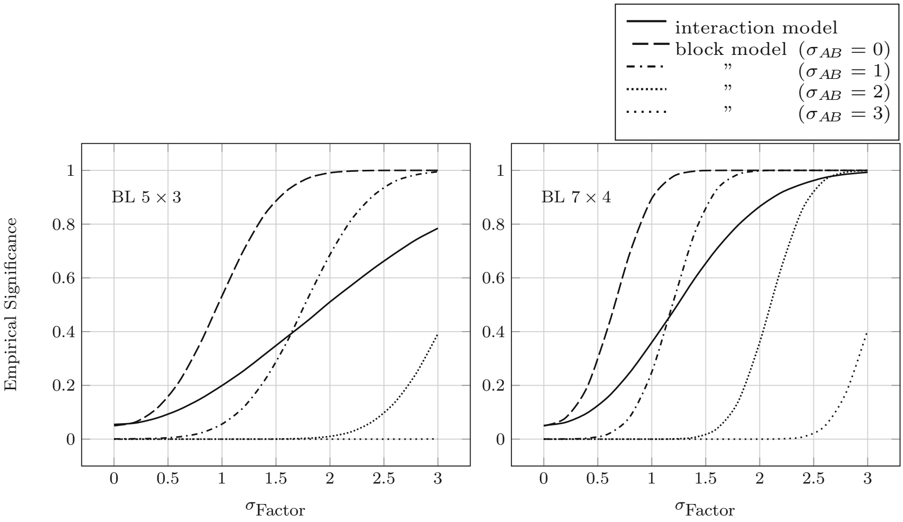

Fig. 6. Empirical power of the test on the main effect as a function of the standard deviation of this effect for different block designs ($\sigma_E = 1$).

5 Estimating Variance Components

The method described here is based on fixed effect model reps. a model with a special definition of interaction (Table 1, $\kappa = 0$). In this case we can estimate the Least Squares (LSQ) variance component for interaction as

$$MS_{AB} = \sigma^2 + n\sigma^2_{ab} \qquad \sigma^2 \equiv MS_E$$

$$\sigma^2_{ab} = \frac{MS_{AB} - MS_E}{n}$$

In block-designs:

$$n = 1 \implies \sigma^2_{ab} = MS_{AB} - MS_E$$

A usual assumption in estimating variance components of interaction based on restricted maximum likelihood (REML) estimation is $\text{cov}(ab_{ij}, ab_{ij'}) = 0, (j \neq j')$ In Fig. 7 both methods are compared for different heights of variance components of interaction (σ^2_{ab}) in compare to the error variance (σ^2).

As long as the variance component for interaction is similar to the of the error variance, there is almost no difference between different estimation methods. In 7 the height of the variance component for interaction is 9 time as high as that of the error variance. Even in this rather extreme situation the differences between LSQ and REML estimation is negligible with at least 6 blocks and 6 levels of the factor. This means the proposed method can be used to estimate variance components in any mixed and random effects model approach with an appropriate sample size.

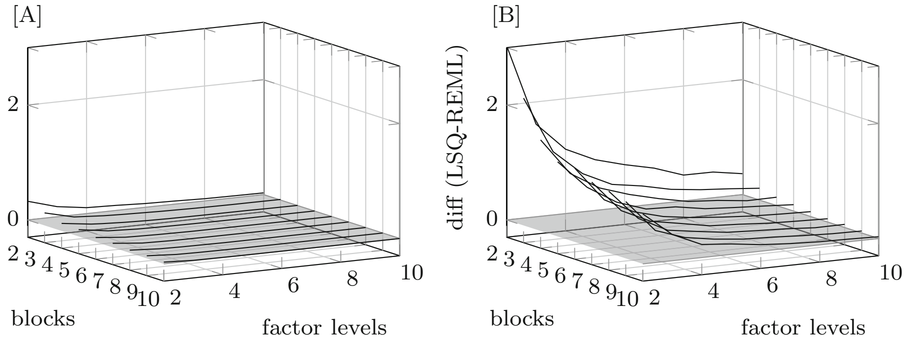

Fig. 7. Comparison of REML and LSQ estimation of variance components for interaction assuming $\sigma_{ab}^2 = \sigma^2$ ([A]) and $\sigma_{ab}^2 = 9\sigma^2$ ([B])

6 Conclusions

The method presented here allows for testing interaction in different kind of block designs. It can also be used for estimating variance components of interaction and main effects for mixed and random effect models in block designs. There are some additional advantages in compare to other methods:

- It is based on a linear model and as such comprises all previously developed non linear models.
- It is very easy to use and any statistical package like SAS, R, SPSS can be used with just a few simple additional calculations.
- There are no uncommon restrictions.
- It can be used not only for common block designs or Latin Squares, but for any statistical model that has a suitable block structure as for instance Split-Plot Designs or maybe even incomplete block designs (although one has to use least-square means here).
- The power of the method is good in respect to interaction and even more for the main effect, if interaction exists.

Block analysis is based on at least a two factorial design (Table 1). Depending on the model (fixed, random, mixed) main effects are tested against either the error sum of squares or that of the interaction. For a fixed and a special kind of mixed effects models (Table 1, $\kappa = 0$) an assessment of the main and block effect must be made by means of the error. In all other situations you have to test against means squares of the interaction. Interaction itself always has to be tested against the error.

So far, means squares for the error term were not or only available under the very restrictive assumptions of non linear models. With the method presented here this test now is possible even in a random or mixed effect surrounding (Table 1, $\kappa = 1$). Although the basic assumption ($\kappa = 1$) is different to that of the fixed model, it doesn't really matter for estimating variance components.

References

1. Alin, A., Kurt, S.: Testing non-additivity (interaction) in two-way ANOVA tables with no replication. Stat. Methods Med. Res. **15**, 63–85 (2006)
2. Barhdadi, A., Froda, S.: An exact test for additivity in two-way tables under biadditive modelling. Commun. Stat. Theor. Methods **39**, 1960–1978 (2010)
3. Cochran, W.G., Cox, D.R.: Eperimental Designs, 2nd edn. Wiley, New York (1957)
4. Gollob, H.F.: A statistical model which combines features of factor analysis and analysis of variance. Psychometrica **33**, 73–115 (1968). https://doi.org/10.1007/BF02289676
5. Johnson, D.E., Graybill, F.A.: An analysis of a two-way model with interaction and no replication. J. Am. Stat. Assoc. **67**, 862–869 (1972)
6. Karabatos, G.: Additivity test. In: Everitt, B.S., Howell, D.C. (eds.) Encyclopedia of Statistics in Behavioral Science, Wiley, Heidelberg, pp. 25–29 (2005)
7. Mandel, J.: Non-additivity in two-way analysis of variance. J. Am. Stat. Assoc. **56**, 878–888 (1961)
8. Mandel, J.: The partitioning of interaction in analysis of variance. J. Res. Nat. Bur. Stand. B Math. Sci. **3**, 309-328 (1969)
9. Mandel, J.: A new analysis of variance model for non-additive data. Technometrics **13**, 1–18 (1971)
10. Rasch, D., Rusch, T., Simeckova, M., Kubinger, K.D., Moder, K., Simecek, P.: Tests of additivity in mixed and fixed effect two-way ANOVA models with single sub-class numbers. Stat. Papers **50**(4), 905–916 (2009)
11. Scheffé, H.: The Analysis of Variance, 2nd edn., pp. 45–56. Springer, New York (1959)
12. Tukey, J.W.: One degree of freedom for non-additivity. Biometrics **5**, 232–242 (1949)