Evaluation of interaction effects with user-defined contrasts in the cell means model

Andreas Kitsche (kitsche@biostat.uni-hannover.de)

Leibniz Universität Hannover

Problem

Experiments in horticultural and agricultural research are often set up with two or more treatment factors. Using the analysis of variance and the corresponding F-tests for the main and interaction effects offers only global inference. If the interactions are the parameters of interest, methods are required detecting the source of significant global interaction. Gabriel et. al. provide a procedure to make inferences on a finite set of interaction contrasts. Unfortunately in most cases this selected family of intercation contrasts does not cover the research question and is in addition not implemented in standard statistical software.

The presented approach takes both the structure of each factor and the research question into account by building user defined contrasts. This leads to a small subset of all possible interaction contrasts and therefore improves the interpretability. Simultaneous inference for this user specified interaction contrasts is available by using quantiles of the multivariate t-distribution. In addition to adjusted p-values we recommend the use of simultaneous confidence intervals to present the direction, magnitude and the relevance of the comparison of interest.

The Model

We assume the following fixed effects model for a completely randomized two-factorial design that reflects the effects of treatments A and B and the AB-interaction in terms of the **ANOVA model**:

$$Y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \epsilon_{ijk}$$

with $i = 1, 2, \ldots, I$, $j = 1, 2, \ldots, J$ and $k = 1, 2, \ldots, n_{ij}$ The corresponding **cell means model** is given by:

$$Y_{ijk} = \mu_{ij} + \epsilon_{ijk}$$

with μ_{ij} the mean of the cell with treatment level i and j. The effects of the factorial model can be reformulated in sense of the cell means model as follows:

- main effect A: $\alpha_i = \mu_{i.} - \mu_{..}$
- main effect B: $\beta_j = \mu_{.j} - \mu_{..}$
- interaction effect: $(\alpha\beta)_{ij} = \mu_{ij} - \mu_{i.} - \mu_{.j} + \mu_{..}$
- simple A_i effect for level B_j: $\alpha_{i(B_j)} = \alpha_i + (\alpha\beta)_{ij} = \mu_{ij} - \mu_{.j}$

Simultaneous Inference

For simplicity the parameter vector of the cell means model is reformulated as:$(\mu_{11}, \ldots, \mu_{1J}, \mu_{21}, \ldots, \mu_{2J}, \mu_{I1}, \ldots \mu_{IJ}) = (\eta_1, \ldots, \eta_S)$. Interest is in simultaneous estimation of $m = 1, \ldots, M$ linear combinations of the parameters η_s: $\theta_m = \sum_{s=1}^{S} c_{ms}\eta_s$. The goal is to test the hypotheses:

$$H_0 : \bigcap_{m=1}^{M} \theta_m = l_m \qquad H_A : \bigcup_{m=1}^{M} \theta_m \neq l_m$$

This test can be performed using the test-statistic: $T_m = \frac{\theta_m - l_m}{\sqrt{s_m}}$ where s_m is the variance estimate of θ_m and l_m is a user specified constant chosen by the researcher (often set to 0). Simultaneous (1-α) confidence intervals for θ_m are given by:

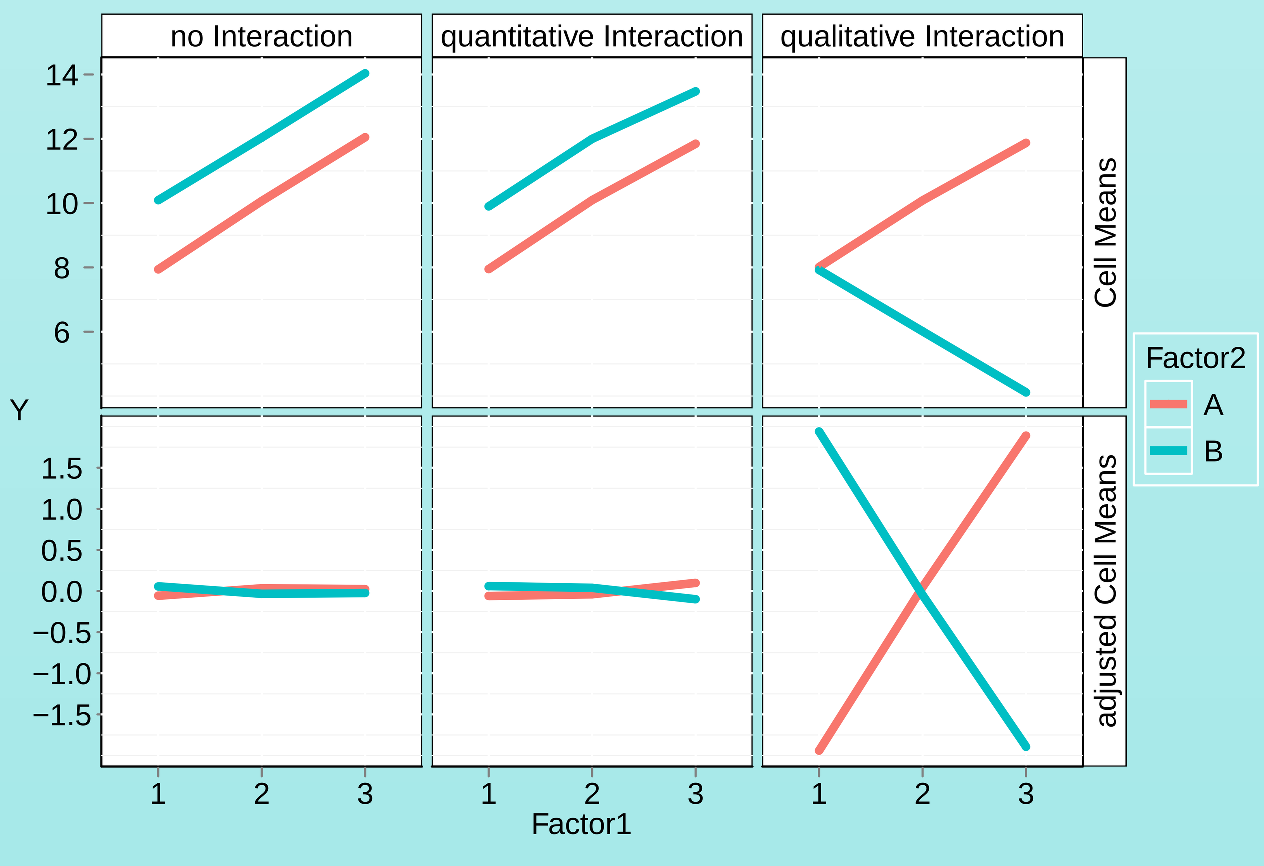
$$\theta_m \pm q_{(M,\mathbf{R},1-\alpha,2-sided)} \cdot \sqrt{s_m}.$$

The equicoordinate quantile q is taken from a M-variate t-distribution with correlation matrix R with elements depending on the chosen contrasts and the sample size:

$$\rho_{mm'} = \frac{\sum_{s=1}^{S} c_{ms}c_{m's}/n_s}{\sqrt{(\sum_{s=1}^{S} c_{ms}^2/n_s)(\sum_{s=1}^{S} c_{m's}^2/n_s)}}$$

Graphical Representation

A common graphical representation of interactions is to plot the observed cell means and connect them by straight lines. The use of adjusted observed cell means, corresponding to the interaction effects of the ANOVA model, is another tool to properly characterize the nature of the interaction effects.



Selected References

[1] K. R. Gabriel, J. Putter, and Y. Wax. Simultaneous Confidence Intervals for Product-Type Interaction Contrasts. *Journal of the Royal Statistical Society. Series B (Methodological)*, 35(2):234–244, 1973.

[2] T. Hothorn, F. Bretz, and P. Westfall. Simultaneous inference in general parametric models. *Biometrical Journal*, 50(3):346–363, 2008.

[3] Roger G. Petersen. *Design and analysis of experiments*, volume 66 of *Statistics*. Dekker, New York, NY, 1985.

Interaction Contrasts

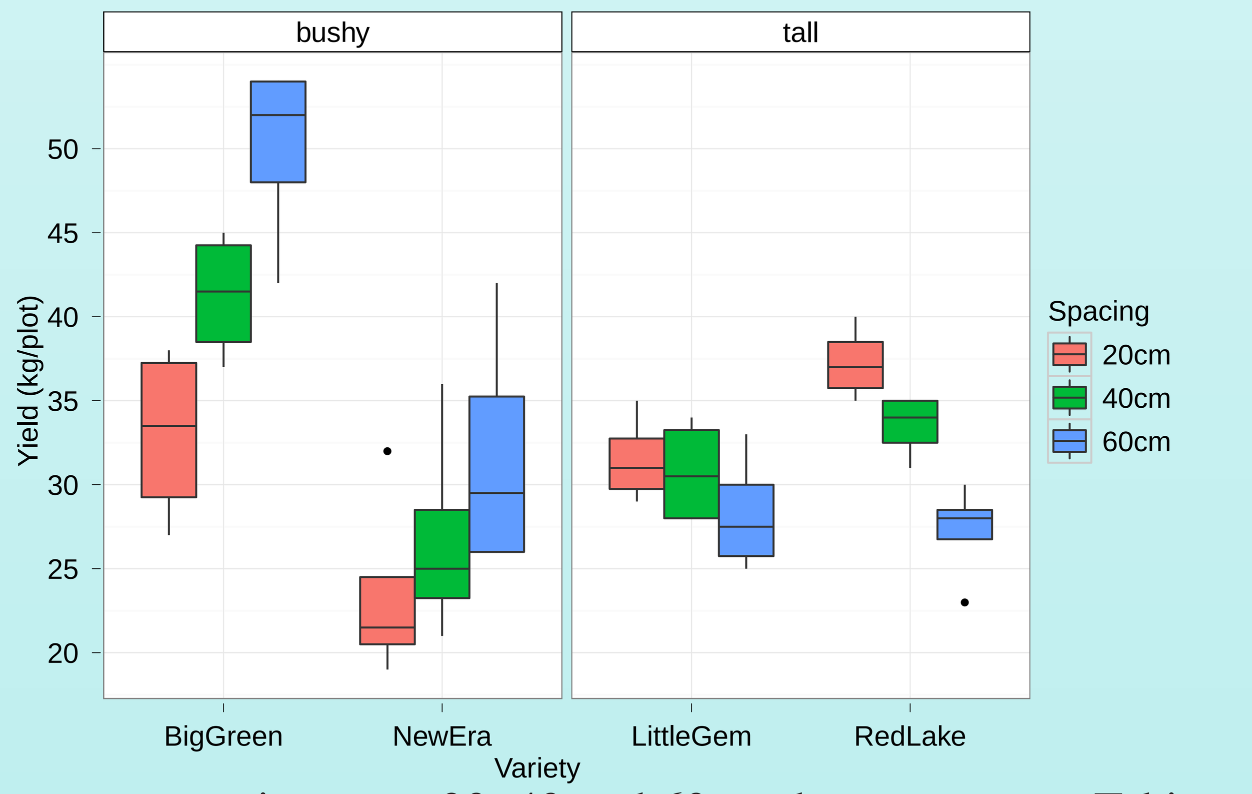
Gabriel et. al. distinguished several types of interaction contrasts. The most commonly known are the interaction residuals, that are the interaction effects of the ANOVA model, and the tetrad contrasts, $\mu_{ij} - \mu_{i'j} - \mu_{ij'} + \mu_{i'j'} (i \neq j', i \neq j')$,which are direct products of pairwise differences.

A more appropriate approach is to define a set of contrasts for each factor, leading to a contrast matrix for each factor. Building the direct (Kronecker) product of this matrices with respect to the factorial ordering leads to an interaction contrast matrix that reflects the treatment structure of each factor and takes the research question into account. This approach has the following characteristics

- enables more detailed inferences than the global F-test
- takes subsets of all possible contrasts
- reflects the structure of each experimental factor
- specifies the research question
- simultaneous inference is possible using the quantiles of the multivariate t-distribution (Hothorn et. al.)

Example

In Petersen (1985) the effect of row spacing on the yield of different varieties of bush beans was investigated. The selected four varieties differ such that "New Era" and "Big Green" form low, bushy plants and the two varieties "Little Gem" and "Red Lake" form erect plants with few branches. The chosen row spacing were 20, 40 and 60 cm between rows. Taking the experimental layout of the trial into account leads to the following user defined contrast matrices:



$$\mathbf{C}^{\mathbf{Variety}} = \begin{pmatrix} -0.5 & -0.5 & 0.5 & 0.5 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & -1 & 1 \end{pmatrix} \qquad \mathbf{C}^{\mathbf{Spacing}} = \begin{pmatrix} 1 & -1 & 0 \\ 1 & 0 & -1 \\ 0 & 1 & -1 \end{pmatrix}$$

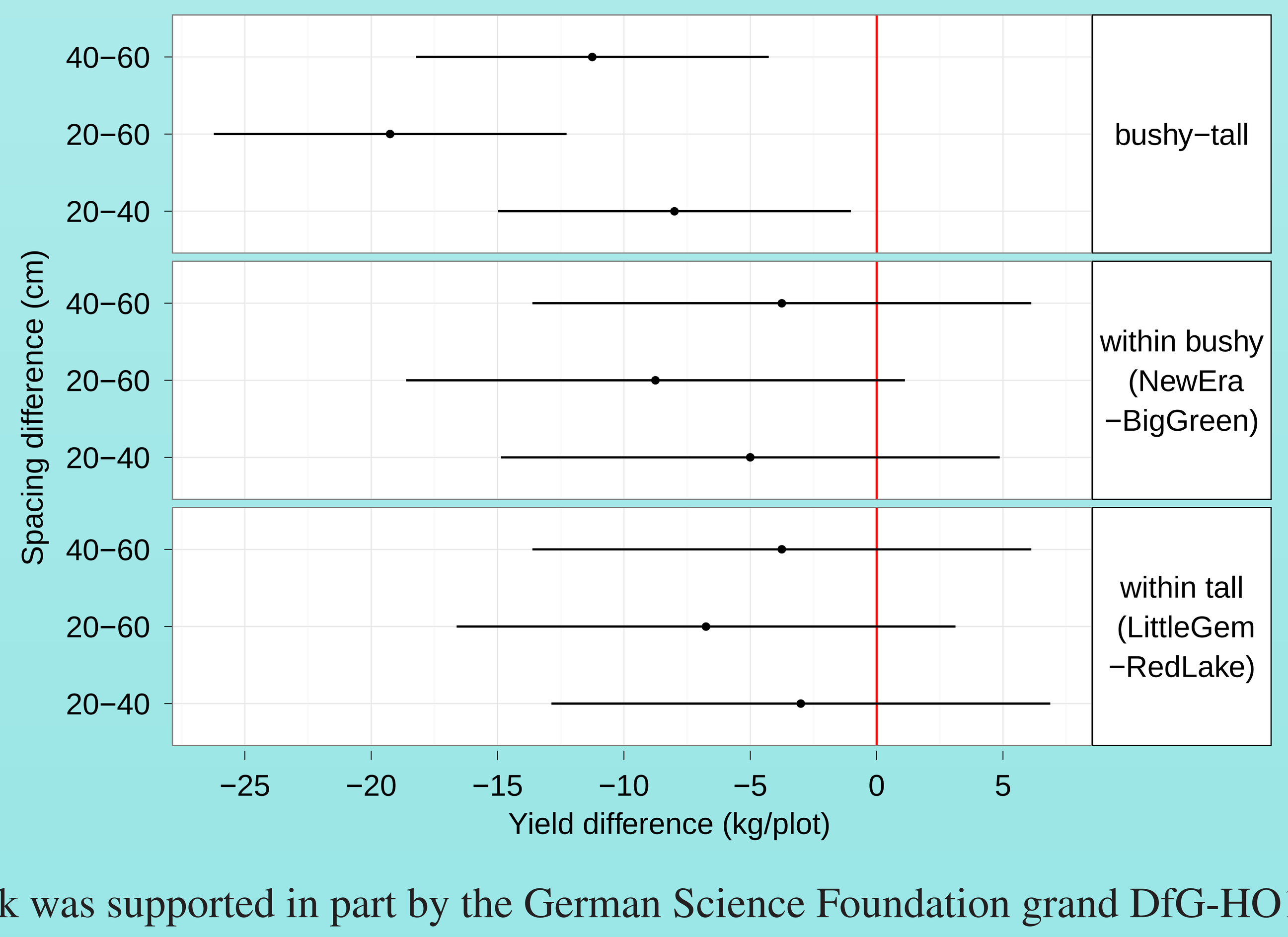
The corresponding interaction contrasts are given by the Kronecker product $C^{Variety} \otimes C^{Spacing}$:

$$\begin{array}{lcccccccccccc} Variety: & & BigGreen & & & NewEra & & & LittleGem & & & RedLake & \\ Spacing: & 20 & 30 & 40 & 20 & 30 & 40 & 20 & 30 & 40 & 20 & 30 & 40 \end{array}$$

$$\mathbf{C}^{\mathbf{Int}} = \begin{pmatrix} 0.5 & -0.5 & 0 & 0.5 & -0.5 & 0 & -0.5 & 0.5 & 0 & -0.5 & 0.5 & 0 \\ 0.5 & 0 & -0.5 & 0.5 & 0 & -0.5 & -0.5 & 0 & 0.5 & -0.5 & 0 & 0.5 \\ 0 & 0.5 & -0.5 & 0 & 0.5 & -0.5 & 0 & -0.5 & 0.5 & 0 & -0.5 & 0.5 \\ 1 & -1 & 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & -1 & -1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 & -1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & -1 & 1 \end{pmatrix}$$

Test-statistics, adjusted p-values and simultaneous confidence intervals:

t value	Pr(>t)
-3.140	0.02829
-7.556	< 1e-04
-4.416	0.00088
-1.388	0.72129
-2.429	0.14608
-1.041	0.90465
-0.833	0.96515
-1.874	0.39952
-1.041	0.90465



Fundings: This work was supported in part by the German Science Foundation grand DfG-HO1687.