Computational Statistics and Data Analysis 54 (2010) 1840–1857

Contents lists available at ScienceDirect

Computational Statistics and Data Analysis

journal homepage: www.elsevier.com/locate/csda

Series of randomized complete block experiments with non-normal data

Arne C. Bathke [a,*], Solomon W. Harrar [b], Haiyan Wang [c], Ke Zhang [d], Hans-Peter Piepho [e]

[a] Department of Statistics, University of Kentucky, Lexington, KY 40506-0027, USA
[b] Department of Mathematical Sciences, University of Montana, Missoula, MT 59812-0864, USA
[c] Department of Statistics, Kansas State University, Manhattan, KS 66506-0802, USA
[d] Department of Pathology, School of Medicine and Health Sciences, University of North Dakota, Grand Forks, ND 58202, USA
[e] Institut für Pflanzenernährung, Universität Hohenheim, D-70599 Stuttgart, Germany

ARTICLE INFO

Article history:
Received 23 June 2009
Received in revised form 9 February 2010
Accepted 10 February 2010
Available online 16 February 2010

Keywords:
Experimental design
Multi-environment trial
Multivariate analysis
Repeated measures
Split-plot design

ABSTRACT

Randomized complete block designs are common in agricultural and other experiments. In this manuscript, we derive asymptotic procedures as well as finite approximations, for the analysis of data arising from series of such experiments. We do not assume normality of the data, and the within-block covariance structures can be arbitrary (no restriction to compound symmetry). The methods are specifically designed for trials with many environments and few blocks per environment, such as multi-environment trials in variety testing and plant breeding. We consider fixed and random effects models for the environment factor. The methodology takes advantage of multivariate notation, and the questions of interest are formulated as profile analysis problems. Finite performance of the proposed procedures is examined in a simulation study, and application is demonstrated using data from a series of crop variety trials.

1. Introduction

Randomized complete block designs (RCBD) constitute one of the most important classes of experimental designs in agricultural research. Most of the published literature on RCBDs and closely related designs follows under the normal theory paradigm. However, often the type of data encountered in field trials does not justify assuming normality, in particular when ordinal ratings are involved. In that case, it is sometimes possible to resort to procedures that are asymptotically valid. The main goal of this manuscript is to extend the realm of asymptotic procedures that do not require normality of the observations.

Classical asymptotic results pertain to the situation where the number of blocks is large (see, e.g. Arnold, 1981, pp. 141–151). However, in field trial applications often the number of blocks is small, while there could be many treatment levels (Caliński et al., 2005; Smith et al., 2005). Thus, non-standard asymptotic results tailored to this situation are of particular interest in agricultural and plant science applications.

In this manuscript, we provide a multivariate profile analysis approach to the analysis of data generated from a design commonly considered in crop variety multi-environment field trials (MET). Here, varietal yield performance is compared across many different environments, with usually few blocks per environment.

For example, in a crop variety trial that will be considered in more detail in Section 6, at $a = 19$ environments, $t = 15$ treatments (different varieties of spring barley) are compared using a randomized block design with $n = 2$ blocks at each environment. Each variety is rated at each environment and each block.

* Corresponding author.
E-mail address: arne@uky.edu (A.C. Bathke).

0167-9473/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.csda.2010.02.007

With regard to this kind of data set, inferential statistical procedures should be aimed at answering the questions whether there is an interaction effect between environment and treatment, and whether there are effects due to the individual factors. Associated hypotheses corresponding to these questions can also be formulated, for example, as equal, parallel, and flat rating profiles across the environments. In addition, it may be appropriate to perform pairwise comparisons between all different treatments. In the subsequent sections, we will derive easily implementable inferential procedures to answer these questions that are asymptotically valid for the many-environment case (large a), but do not require a large number of blocks (small n) and do not rely on assuming normality of the data, nor a compound symmetry covariance structure within blocks. To our knowledge, asymptotic methods for this situation have not been derived before.

McIntosh (1983) is a highly quoted reference, in which appropriate F-ratios are listed for the analysis of combined experiments under the assumptions of an additive model with compound symmetry covariance structure within blocks, and normal error terms. Recently, different mixed model approaches have emerged, that allow for greater flexibility in modeling covariance structures (Denis et al., 1997; Smith et al., 2005; Caliński et al., 2005). However, inference in these models is also derived under the normality assumption. Caliński et al. (2005) also explicitly considered resolvable incomplete block designs. While the discussion in the present paper is limited to complete block designs, note the result by Speed et al. (1985) who pointed out that an ordinary complete block analysis can be used to analyze any resolvable incomplete block design, including designs with unequal block size. A Bayesian approach to estimation for MET data has been proposed by Theobald et al. (2002).

The example described above fits generally into the framework of the F1-LD-F1 design described by Brunner and Langer (1999), who have derived nonparametric inference procedures that also do not rely on normality of the data. However, those procedures were designed for the large n asymptotic situation. Thus, their results do not apply to crop variety trials with small number (n) of blocks per environment and typically many (a) environments and several (t) treatments.

In the following, we pursue the path of deriving large a asymptotics for the described designs, and then consider finite-sample approximations that are valid even when a is small.

To derive the asymptotic results, we follow a multivariate approach (Bathke and Harrar, 2008; Harrar and Bathke, 2008), and the different questions of interest are formulated as profile analysis problems (cf. Harrar, 2009). We noticed that it is not uncommon to find the same or very similar test statistics expressed in rather different notation. This often makes it difficult to appreciate the relations between different approaches (e.g., univariate vs. multivariate). In the following sections, we therefore attempt to provide 'translations' of the test statistics into different notational 'languages'. The summation notation seems to dominate the more applied literature, and it is most intuitive to statistics practitioners, but the other notations possess advantages: In particular, the multivariate notation is most useful in theoretical derivations and for programming purposes. It facilitates easy implementation of the proposed test statistics in statistical software packages that allow for matrix-based programming (e.g., SAS IML, R).

An issue that has been discussed controversially in the literature is whether the environment factor should be assumed as fixed or random, and an answer may depend upon the objective of the experiment. If the different environments can be justified as legitimate random samples from a target region, then a random effects model may be used (Yates and Cochran, 1938; Comstock and Moll, 1963; Kuehl, 2000, p. 295; Caliński et al., 2005; Littell et al., 2006, Ch. 6). We will consider both possibilities, fixed or random environment factor. A fixed effects model is considered in Section 2, while Section 3 discusses the changes that occur if the environment effect is assumed random instead of fixed.

Denote the number of environments by a, and let the number of blocks for each environment be n. The total number of blocks is $N = a \cdot n$. Each block is subjected to t treatments, e.g., different varieties of a crop species.

Note that the series of randomized block designs at different environments could also be regarded as having split-plot form. Here, the different environments correspond to the whole-plot factor, the different blocks at a environment correspond to whole-plot units, and the plots within a block correspond to sub-plot units. As opposed to usual split plots, however, the environment factor cannot be randomized.

An additive model equation with fixed environment factor is

$$Y_{ijk} = \mu + \lambda_i + C(\lambda)_{j(i)} + \tau_k + (\lambda\tau)_{ik} + \varepsilon_{ijk},$$

where μ is the overall mean, $\lambda_i,\ i = 1, \ldots, a$ denotes the (fixed) environment effect, $\tau_k,\ k = 1, \ldots, t$ is the (fixed) treatment effect, $(\lambda\tau)_{ik}$ is the (fixed) interaction effect between environment and treatment, $C(\lambda)_{j(i)},\ j = 1, \ldots, n$ describes the (random) block effect, nested within environment, and ε_{ijk} is the random error. The fixed effects are assumed to satisfy the usual identifiability constraints.

Block effect and error are commonly modeled to follow independent normal distributions (see, e.g., Littell et al., 2006, Chapter 6). In this manuscript, we will not assume normality of block effect or error (nor of the effects pertaining to the environment factor if it is modeled as random). Also, we will allow for a more general correlation structure between observations from the same block than the often assumed compound symmetry that would be implied by the above equations (cf. the simulation study in Section 5): The within-block covariance matrix is here assumed completely unstructured. As an alternative approach, one could fit different specific covariance structures to the data, and use some criterion to choose a particular structure (Guerin and Stroup, 2000; Littell, 2002). While such an approach certainly has the potential advantage of a power gain when the covariance structure with the best fit is close to correct, it can lead to poor results if the chosen covariance structure is wrong (note that it may merely be the best structure among those candidate models which yielded any results), and it is generally computationally intensive (Loughin, 2006). We will not pursue it here.

The general covariance structure used in this manuscript can be expressed by aggregating block effect and error into one random variable, $\tilde{\varepsilon}_{ijk}$ say, with the assumption that the mean zero random variables $\tilde{\varepsilon}_{ijk}$ have the same variance and are independent if they correspond to different blocks $j(i)$. In addition, we make the technical assumption that the fourth moments of $\tilde{\varepsilon}_{ijk}$ exist. The model above is then contained as a special case in this more general model formulation:

$$Y_{ijk} = \mu + \lambda_i + \tau_k + (\lambda\tau)_{ik} + \tilde{\varepsilon}_{ijk}. \quad (1)$$

In a similar way, a general model with random environment factor is

$$Y_{ijk} = \mu + L_i + \tau_k + (L\tau)_{ik} + \tilde{\varepsilon}_{ijk}, \quad (2)$$

where L_i and $(L\tau)_{ik},\ i = 1, \ldots, a,\ k = 1, \ldots, t$, are random. This model will be discussed in more detail in Section 3. Note that the additive model equation (2), together with the assumptions on $\tilde{\varepsilon}_{ijk}$ formulated in the paragraph before Eq. (1), implies equal covariances between observations from different blocks at the same environment. Caliński et al. (2005) allow for a more general within-environment covariance structure between blocks, but assume normality.

The model with fixed environment effect is formulated in multivariate notation as $\mathbf{Y} = \mathbf{ZB} + \mathbf{E}$, where $\mathbf{Y}$ is the $(N \times t)$ matrix of responses, $\mathbf{Z}$ is an $(N \times a)$ design matrix, $\mathbf{B}$ is the $(a \times t)$ parameter matrix, and $\mathbf{E}$ is the error matrix (with entries $\tilde{\varepsilon}_{ijk}$). The rows of $\mathbf{E}$ are assumed to be independent and follow a multivariate distribution with mean zero and positive definite covariance matrix $\boldsymbol{\Sigma}$. The canonical estimator for this covariance matrix is the residual MSE matrix for the responses.

In some situations, it is more convenient to make use of the univariate notation. That is, the multivariate model above can be rewritten as $\mathbf{y} = \tilde{\mathbf{Z}}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$, where $\mathbf{y} = \text{vec}(\mathbf{Y}')$ is a $N \cdot t$-dimensional column vector, $\tilde{\mathbf{Z}} = \mathbf{Z} \otimes \mathbf{I}_t$, $\boldsymbol{\beta} = \text{vec}(\mathbf{B}')$, and $\boldsymbol{\varepsilon} = \text{vec}(\mathbf{E}')$, using the identity $\tilde{\mathbf{Z}}\boldsymbol{\beta} = (\mathbf{Z} \otimes \mathbf{I}_t)\,\text{vec}(\mathbf{B}') = \text{vec}(\mathbf{B}'\mathbf{Z}')$. We will use both, the multivariate and the univariate notation in this manuscript.

Here and in the following, for any natural number d, $\mathbf{I}_d$ denotes the identity matrix of dimension d, $\mathbf{1}_d$ a d-dimensional column vector of ones, $\mathbf{J}_d = \mathbf{1}_d \cdot \mathbf{1}_d'$ a square matrix of ones, and $\mathbf{P}_d = \mathbf{I}_d - d^{-1}\mathbf{J}_d$ the projector onto the space orthogonal to the vector $\mathbf{1}_d$.

In the next section, asymptotic results will be derived for testing main effects due to environment and treatment, as well as their interaction, assuming that environment is a fixed factor with many levels. Section 3 discusses the changes when environment is considered random instead of fixed. A method for pairwise comparisons of all treatments is presented in Section 4. Results from simulation studies in which the performance of the proposed tests is evaluated, in comparison with normal theory F tests, are summarized in Section 5 and Section 6 illustrates application of the procedures with a real data set. The Appendix contains many of the more technical results.

2. Asymptotic results for fixed environment factor

Consider the example described in Section 1. Here, in a environments, t different varieties (treatments) of a crop species are compared using a randomized block design with n blocks nested within each environment. The multivariate model equation to describe this and similar situations is $\mathbf{Y} = \mathbf{ZB} + \mathbf{E}$, where the rows of $\mathbf{E}$ are independent with mean zero and covariance matrix $\boldsymbol{\Sigma}$, assumed to be positive definite. In this section, the effects due to environment and treatment are considered fixed, they are described by the parameter matrix $\mathbf{B}$. The block effect is random, and it is reflected by the general unstructured covariance matrix $\boldsymbol{\Sigma}$.

Model (1) can be translated into this additive multivariate model: The design matrix is $\mathbf{Z} = \mathbf{I}_a \otimes \mathbf{1}_n$ and the ith row and kth column entry of the parameter matrix $\mathbf{B}$ is $\lambda_i + \tau_k + (\lambda\tau)_{ik}$.

The typical hypotheses to be tested within this model are those of no main effects due to environment or treatment, and no interaction effect between these two factors. In addition, it may be of interest to test the corresponding simple factor effects. These different hypotheses are explained in more detail in the respective subsections below.

In the following, the $(n \times t)$ sub-matrix corresponding to the ith environment is denoted by $\mathbf{Y}_i$, and $\mathbf{Y}_{ij} = (Y_{ij1}, \ldots, Y_{ijt})$ is the $(1 \times t)$ response vector associated with the jth block at the ith environment.

2.1. Parallel treatment profiles (no treatment–environment interaction)

The first natural question in analyzing data following the described model is whether the treatment profiles can be considered parallel. This corresponds to testing whether there is an interaction effect between environment (whole-plot factor) and treatment (time, sub-plot-factor). Formally, this null hypotheses can be written in multivariate form as $H_0 : \mathbf{P}_a\mathbf{B}\mathbf{P}_t = \mathbf{0}$, or in univariate form as $(\mathbf{P}_a \otimes \mathbf{P}_t)\boldsymbol{\beta} = \mathbf{0}$. Note that the former $\mathbf{0}$ denotes a matrix, while the latter denotes a vector. In terms of the additive components used in model equation (1), the null hypothesis of no interaction can also be expressed as $H_0 : (\lambda\tau)_{ik} \equiv 0,\ i = 1, \ldots, a,\ k = 1, \ldots, t$. That is, the (i, k) element of the parameter matrix $\mathbf{B}$ is, under null hypothesis, $\lambda_i + \tau_k$.

To test this hypothesis, consider the ANOVA-type test statistic, which is the ratio of the interaction mean square and the pooled error mean square. The test statistic can formally be defined in several different ways. The multivariate notation has advantages in methodological derivations and for implementation in statistical software, but univariate and summation notation appear to be used by a wider audience.

(a) Using a univariate quadratic form to define the treatment sum of squares:

$$F_{AB} = \frac{N\bar{\mathbf{y}}_{\cdot}'(\mathbf{P}_a \otimes \mathbf{P}_t)\bar{\mathbf{y}}_{\cdot}}{\operatorname{tr}\left[(\mathbf{P}_a \otimes \mathbf{P}_t)\hat{\mathbf{V}}\right]}, \quad \text{where } \bar{\mathbf{y}}_{\cdot} = \left(\mathbf{I}_a \otimes \frac{1}{n}\mathbf{1}_n' \otimes \mathbf{I}_t\right)\mathbf{y} = \operatorname{vec}\left[\mathbf{Y}'\left(\mathbf{I}_a \otimes \frac{1}{n}\mathbf{1}_n\right)\right], \tag{3}$$

$$\text{and} \quad \hat{\mathbf{V}} = \mathbf{I}_a \otimes \mathbf{Y}'\left(\mathbf{I}_a \otimes \frac{1}{n-1}\mathbf{P}_n\right)\mathbf{Y} = \mathbf{I}_a \otimes \left(\sum_{i=1}^{a} \mathbf{Y}_i' \frac{1}{n-1}\mathbf{P}_n\mathbf{Y}_i\right). \tag{4}$$

(b) Expressing numerator and denominator as traces of matrix-valued quadratic forms:

$$F_{AB} = \frac{\frac{1}{a-1}\operatorname{tr}\left[\mathbf{P}_t\mathbf{Y}'\left(\mathbf{P}_a \otimes \frac{1}{n}\mathbf{J}_n\right)\mathbf{Y}\right]}{\frac{1}{a(n-1)}\operatorname{tr}\left[\mathbf{P}_t\mathbf{Y}'(\mathbf{I}_a \otimes \mathbf{P}_n)\mathbf{Y}\right]}. \tag{5}$$

(c) Employing the univariate notation in terms of $\mathbf{y} = \operatorname{vec}(\mathbf{Y}')$:

$$F_{AB} = \frac{\frac{1}{a-1}\mathbf{y}'\left(\mathbf{P}_a \otimes \frac{1}{n}\mathbf{J}_n \otimes \mathbf{P}_t\right)\mathbf{y}}{\frac{1}{a(n-1)}\mathbf{y}'(\mathbf{I}_a \otimes \mathbf{P}_n \otimes \mathbf{P}_t)\mathbf{y}}. \tag{6}$$

(d) And finally, writing the test statistic in summation notation:

$$F_{AB} = \frac{B_{AB}}{W_{AB}}, \quad \text{where } B_{AB} = \frac{n}{(a-1)(t-1)}\sum_{i=1}^{a}\sum_{k=1}^{t}\left(\bar{Y}_{i\cdot k} - \bar{Y}_{i\cdot\cdot} - \bar{Y}_{\cdot\cdot k} + \bar{Y}_{\cdot\cdot\cdot}\right)^2 \tag{7}$$

$$\text{and} \quad W_{AB} = \frac{1}{at(n-1)}\sum_{i=1}^{a}\left[\sum_{j=1}^{n}\sum_{k=1}^{t}(Y_{ijk} - \bar{Y}_{i\cdot k})^2 - \frac{1}{t-1}\sum_{j=1}^{n}\sum_{k\neq k'}(Y_{ijk} - \bar{Y}_{i\cdot k})(Y_{ijk'} - \bar{Y}_{i\cdot k'})\right].$$

The following proposition states that these four forms are all equivalent, which might not be obvious at first sight. The proof can be found in the Appendix.

Proposition 1. *The four different forms of F_{AB} defined in displays (3), (5)–(7), are equivalent.*

Note that Brunner and Langer (1999) proposed, in the large n asymptotic context, a different variance–covariance matrix estimator. They defined the block-diagonal matrix

$$\hat{\mathbf{V}}_B = \bigoplus_{i=1}^{a} \mathbf{Y}_i' \frac{1}{n-1}\mathbf{P}_n\mathbf{Y}_i. \tag{8}$$

However, this estimator actually results here in the same denominator of the ANOVA-type statistic F_{AB}, namely

$$\operatorname{tr}\left[(\mathbf{P}_a \otimes \mathbf{P}_t)\left(\bigoplus_{i=1}^{a} \mathbf{Y}_i' \frac{1}{n-1}\mathbf{P}_n\mathbf{Y}_i\right)\right] = (a-1)\operatorname{tr}\left[\mathbf{P}_t\mathbf{Y}'\left(\mathbf{I}_a \otimes \frac{1}{n-1}\mathbf{P}_n\right)\mathbf{Y}\right].$$

Next, we derive the asymptotic (large a) distribution of F_{AB}.

Theorem 1. *As $a \to \infty$, the ANOVA-type test for interaction, F_{AB}, follows asymptotically a normal distribution,*

$$\sqrt{a}(F_{AB} - 1)\frac{1}{\sqrt{\hat{\tau}}} \xrightarrow{d} N(0, 1), \quad \textit{where } \hat{\tau} = \frac{2n}{n-1}\frac{\operatorname{tr}(\mathbf{P}_t\hat{\boldsymbol{\Sigma}})^2}{(\operatorname{tr}\mathbf{P}_t\hat{\boldsymbol{\Sigma}})^2} \quad \textit{and} \quad \hat{\boldsymbol{\Sigma}} = \frac{1}{a(n-1)}\left[\mathbf{Y}'(\mathbf{I}_a \otimes \mathbf{P}_n)\mathbf{Y}\right].$$

The proof of Theorem 1 can be found in the Appendix.

Note that the asymptotic (large a) distribution of the standardized F ratio does not depend on the underlying population distribution of the response variable. This could motivate approximating the small a sampling distribution of F_{AB} by its normal theory counterpart, assuming a spherical covariance structure, resulting in an F-distribution with numerator degrees of freedom $df_1 = (a-1)(t-1)$ and denominator degrees of freedom $df_2 = a(n-1)(t-1)$ (see, e.g. McIntosh, 1983).

An alternative small-sample approximation for F_{AB} that does not have to rely on covariance sphericity can be constructed based on a moment matching approach. Assume that the sampling distribution of F_{AB} can be approximated by a scaled χ^2-distribution, $F_{AB} \sim \chi_f^2/f$. Then, the variance of $\sqrt{a}(F_{AB} - 1)$ equals $2a/f$. Setting this equal to the estimated variance $\hat{\tau}$, we obtain

$$\hat{f} = \frac{2a(n-1)}{2n}\frac{\left[\operatorname{tr}(\mathbf{P}_t\hat{\boldsymbol{\Sigma}})\right]^2}{\operatorname{tr}(\mathbf{P}_t\hat{\boldsymbol{\Sigma}})^2} = \frac{n-1}{n}\frac{a}{a-1}\frac{\left(\operatorname{tr}\left[(\mathbf{P}_a \otimes \mathbf{P}_t)\hat{\mathbf{V}}\right]\right)^2}{\operatorname{tr}\left[(\mathbf{P}_a \otimes \mathbf{P}_t)\hat{\mathbf{V}}\right]^2}.$$

Brunner and Langer (1999, p. 141) recommended the following similarly constructed degrees of freedom estimator for a χ^2-distribution approximation of F_{AB} in the large n context.

$$\hat{f}_{Bru} = \frac{\left(\mathrm{tr}\left[(\mathbf{P}_a \otimes \mathbf{P}_t)\hat{\mathbf{V}}_B\right]\right)^2}{\mathrm{tr}\left[(\mathbf{P}_a \otimes \mathbf{P}_t)\hat{\mathbf{V}}_B\right]^2}.$$

However, recall that they used a covariance matrix estimator $\hat{\mathbf{V}}_B$ that was defined in a slightly different way (cf. display (8)).

It should be mentioned that instead of the above ANOVA-type statistic, tests for interaction can also be derived based on the usual multivariate test statistics, such as Wilks' Lambda, and the Lawley–Hotelling and Bartlett–Nanda–Pillai criteria. We will only explain the idea briefly, and refer to Harrar (2009) for details on this approach and for small-sample approximations.

Denote $\mathbf{L}$ the matrix that satisfies $\mathbf{P}_t = \mathbf{L}\mathbf{L}'$ such that $\mathbf{L}'\mathbf{L} = \mathbf{I}_{t-1}$ and define $\mathbf{Z} = \mathbf{Y}\mathbf{L}$. Notice that F_{AB} in (5) can then be rewritten as

$$F_{AB} = \frac{\frac{1}{a-1}\mathrm{tr}\left[\mathbf{Z}'\left(\mathbf{P}_a \otimes \frac{1}{n}\mathbf{J}_n\right)\mathbf{Z}\right]}{\frac{1}{a(n-1)}\mathrm{tr}\left[\mathbf{Z}'\left(\mathbf{I}_a \otimes \mathbf{P}_n\right)\mathbf{Z}\right]}.$$

In this form, it can be seen directly that F_{AB} is identical to the multivariate ANOVA-type statistic that has been considered by Bathke and Harrar (2008) in the large a framework. In that paper, the role of $\mathbf{X}$ is assumed by a matrix of componentwise ranks, but Harrar and Bathke (2008) also explicitly derived large a asymptotic results for the multivariate ANOVA-type statistic based on a matrix of original observations. Following the notation in Harrar and Bathke (2008), we define $\mathbf{H} = (a-1)^{-1}\mathbf{Z}'\left(\mathbf{P}_a \otimes \frac{1}{n}\mathbf{J}_n\right)\mathbf{Z}$ and $\mathbf{G} = [a(n-1)]^{-1}\mathbf{Z}'\left(\mathbf{I}_a \otimes \frac{1}{n-1}\mathbf{P}_n\right)\mathbf{Z}$ as the hypothesis and error mean squares and cross products. Based on these matrix-valued quadratic forms, the ANOVA-type statistic is $\mathrm{tr}(\mathbf{H})/\mathrm{tr}(\mathbf{G})$, and other common multivariate test statistics can be expressed as, for example, $\mathrm{tr}(\mathbf{H}\mathbf{G}^{-1})$ (Lawley–Hotelling), $\mathrm{tr}\left[\mathbf{H}(\mathbf{H}+\mathbf{G})^{-1}\right]$ (Bartlett–Nanda–Pillai), and $-\log\det[\mathbf{G}(\mathbf{H}+\mathbf{G})^{-1}]$ (likelihood ratio or Wilks' Lambda). Harrar (2009) considered the three latter types of tests and derived asymptotic results and small-sample approximations for different profile analysis hypotheses.

2.2. Equal profile averages (no environment effect)

Given that the average treatment profiles are parallel (no treatment–environment interaction), the next question of interest is whether the profile averages are equal (no environment effect, $\lambda_i = 0$, $i = 1, \ldots, a$). The corresponding hypothesis can be written as $H_0 : \mathbf{P}_a\mathbf{B}\mathbf{J}_t = \mathbf{0}$ (multivariate formulation of model) or $H_0 : (\mathbf{P}_a \otimes \mathbf{J}_t)\boldsymbol{\beta} = \mathbf{0}$ (univariate formulation). Note that the hypothesis of equal profile averages may be meaningless in case of treatment profiles that are not parallel, that is, in presence of treatment–environment interaction. For example, one could imagine three different treatment profiles at different environments, such that one environment has an increasing profile, another is decreasing, and the third one is flat. All three treatment profiles could have the same average. However, only focusing on this average would obviously ignore the more interesting fact that the treatment profiles are different, and it could lead to misleading conclusions. This can be avoided either by testing for the environment simple effect, or by carefully evaluating the treatment–environment interaction first.

To test the hypothesis of equal profile averages, an ANOVA-type statistic can be defined in the following four ways described in the previous section.

$$\text{(a)}\quad F_A = \frac{N\bar{\mathbf{y}}_{\cdot}'(\mathbf{P}_a \otimes \mathbf{J}_t)\bar{\mathbf{y}}_{\cdot}}{\mathrm{tr}\left((\mathbf{P}_a \otimes \mathbf{J}_t)\hat{\mathbf{V}}\right)}, \quad \text{where } \bar{\mathbf{y}}_{\cdot} \text{ is defined in (3)} \tag{9}$$

$$\text{(b)}\quad F_A = \frac{\frac{1}{a-1}\mathrm{tr}\left[\mathbf{J}_t\mathbf{Y}'\left(\mathbf{P}_a \otimes \frac{1}{n}\mathbf{J}_n\right)\mathbf{Y}\right]}{\frac{1}{a(n-1)}\mathrm{tr}\left[\mathbf{J}_t\mathbf{Y}'(\mathbf{I}_a \otimes \mathbf{P}_n)\mathbf{Y}\right]} \qquad \text{(c)}\quad F_A = \frac{\frac{1}{a-1}\mathbf{y}'\left(\mathbf{P}_a \otimes \frac{1}{n}\mathbf{J}_n \otimes \frac{1}{t}\mathbf{J}_t\right)\mathbf{y}}{\frac{1}{a(n-1)}\mathbf{y}'\left(\mathbf{I}_a \otimes \mathbf{P}_n \otimes \frac{1}{t}\mathbf{J}_t\right)\mathbf{y}}$$

$$\text{(d)}\quad F_A = \frac{B_A}{W_A}, \quad \text{where } B_A = \frac{n}{a-1}\sum_{i=1}^{a}\sum_{k=1}^{t}(\bar{Y}_{i\cdot\cdot} - \bar{Y}_{\cdot\cdot\cdot})^2$$

$$\text{and } W_A = \frac{1}{at(n-1)}\sum_{i=1}^{a}\sum_{j=1}^{n}\sum_{k,k'}^{t}(Y_{ijk} - \bar{Y}_{i\cdot k})(Y_{ijk'} - \bar{Y}_{i\cdot k'}).$$

An alternative approach to deriving a statistic for testing parallel average treatment profiles consists of aggregating all t observations per block into a univariate random variable, $X_{ij} = t^{-1}\mathbf{Y}_{ij}\mathbf{1}_t$, and analyzing the resulting data using results from Bathke (2002). That is, let $\mathbf{X} = t^{-1}\mathbf{Y}\mathbf{1}_t$ be the vector of aggregated responses and define

$$\text{(e)}\quad F_A = \frac{\frac{1}{a-1}\mathbf{X}'\left(\mathbf{P}_a \otimes \frac{1}{n}\mathbf{J}_n\right)\mathbf{X}}{\frac{1}{a(n-1)}\mathbf{X}'(\mathbf{I}_a \otimes \mathbf{P}_n)\mathbf{X}}. \tag{10}$$

Each of these formulations represents the ratio between environment mean squares and block (environment) mean squares. As in the previous section, we first state the equivalence between these different formulations. For the proof, see the Appendix.

Proposition 2. *The five different forms of F_A defined in displays (9)–(10) are equivalent.*

As stated in Proposition 2, the test for main effect of environment reduces to the normal theory F-test statistic presented, for example, in Rencher (2002). The asymptotic (large a) distribution of F_A can then be obtained directly from Theorem 1 in Bathke (2002), (see also Boos and Brownie, 1995). That is, $\sqrt{a}(F_A-1)$ is asymptotically normal with mean zero and variance $2n/(n-1)$. Obviously, we should obtain the same outcome when deriving the asymptotic distribution of F_A by following the procedure used in the derivation of the test for parallel profiles (Section 2.1). Indeed, this route results in

$$\sqrt{a}(F_A-1)\frac{1}{\sqrt{\tau}} \xrightarrow{d} N(0,1) \quad \text{where } \tau = \frac{2n}{n-1}\frac{\mathrm{tr}(\mathbf{J}_t\boldsymbol{\Sigma})^2}{[\mathrm{tr}(\mathbf{J}_t\boldsymbol{\Sigma})]^2} = \frac{2n}{n-1},$$

noticing that $\mathrm{tr}(\mathbf{J}_t\boldsymbol{\Sigma})^2 = (\mathbf{1}_t'\boldsymbol{\Sigma}\mathbf{1}_t)^2 = [\mathrm{tr}(\mathbf{J}_t\boldsymbol{\Sigma})]^2$.

Small-sample approximations for F_A can be constructed similar to Section 2.1, resulting in $F_A \sim F[a-1, a(n-1)]$ (normal theory result under sphericity), or $F_A \sim \chi_f^2/f$ where $f = a(n-1)/n$ (moment method).

Srivastava (1987) derived the normal theory likelihood ratio statistic for testing equality of profiles given parallelism of profiles. It is

$$\Lambda_A = \frac{\mathbf{1}_t'(\mathbf{S}_h+\mathbf{S}_e)^{-1}\mathbf{1}_t}{\mathbf{1}_t'\mathbf{S}_e^{-1}\mathbf{1}_t}, \quad \text{where } \mathbf{S}_h = \mathbf{Y}'\left(\mathbf{P}_a \otimes \frac{1}{n}\mathbf{J}_n\right)\mathbf{Y} \text{ and } \mathbf{S}_e = \mathbf{Y}'(\mathbf{I}_a \otimes \mathbf{P}_n)\mathbf{Y}.$$

It was also shown that a monotone function of Λ_A has under normality an F distribution:

$$T_A = \frac{(N-a-t+1)}{(a-1)}\frac{(1-\Lambda_A)}{\Lambda_A} \sim F_{a-1,N-a-t+1}.$$

Under non-normality, Harrar and Xu (2007) derived an asymptotic (large n) expansion for the null distribution of T_A including terms of order up to $O(n^{-1})$. The large a asymptotic distribution under non-normality is given by the following theorem.

Theorem 2. *As $a \to \infty$, while n and t are fixed, T_A follows asymptotically a normal distribution,*

$$\sqrt{a}(T_A-1) \xrightarrow{d} N\left(0, \frac{2n}{n-1}\right).$$

The proof for Theorem 2 can be found in the Appendix.

It must be underscored again that the above test statistics are applicable only if we are testing for equal profile averages given that the average profiles are parallel (i.e. given that there is no treatment–environment interaction). It may be of interest, however, to test the equality of average profiles against the general alternative of profiles with unequal averages, commonly called 'testing the simple environment effect' to distinguish this question from 'testing the environment main effect' (equal profile averages). This can be done using either one of the following four multivariate tests: ANOVA-type, Lawley–Hotelling, Bartlett–Nanda–Pillai (Bathke and Harrar, 2008; Bathke et al., 2008), Wilks' Lambda (Liu et al., 2008). For each of these tests, the large a asymptotic distribution under non-normality and for rank-based tests have been established and details, as well as a discussion of small-sample approximations can be found in the mentioned publications. The methods described there are related to those introduced in this manuscript, as can be seen from writing the 'simple environment effect' ANOVA-type test statistic as

$$F_{A|B} = \frac{\frac{1}{t(a-1)}\mathrm{tr}\left[\mathbf{I}_t\mathbf{Y}'\left(\mathbf{P}_a \otimes \frac{1}{n}\mathbf{J}_n\right)\mathbf{Y}\right]}{\frac{1}{at(n-1)}\mathrm{tr}\left[\mathbf{I}_t\mathbf{Y}'(\mathbf{I}_a \otimes \mathbf{P}_n)\mathbf{Y}\right]} = \frac{\frac{1}{a-1}\mathrm{tr}\left[\mathbf{Y}'\left(\mathbf{P}_a \otimes \frac{1}{n}\mathbf{J}_n\right)\mathbf{Y}\right]}{\frac{1}{a(n-1)}\mathrm{tr}\left[\mathbf{Y}'(\mathbf{I}_a \otimes \mathbf{P}_n)\mathbf{Y}\right]}.$$

The Lawley–Hotelling, Bartlett–Nanda–Pillai, and Wilks' Lambda-type tests can now be defined as described at the end of Section 2.1.

2.3. Flat treatment profiles (no treatment effect)

Finally, we consider the hypothesis that the average treatment profiles are flat (no treatment effect; $\tau_k = 0$, $k = 1, \ldots, t$), which should be tested after investigating whether the profiles are parallel. As in the previous section, it should be pointed out that the interpretation of 'no treatment effect' may be meaningless in the presence of treatment–environment interaction (non-parallel profiles). The hypothesis of flatness can be written as $H_0 : \mathbf{J}_a\mathbf{B}\mathbf{P}_t = \mathbf{0}$ (multivariate formulation of model) or $H_0 : (\mathbf{J}_a \otimes \mathbf{P}_t)\boldsymbol{\beta} = \mathbf{0}$ (univariate formulation).

The ANOVA-type statistic for testing the hypothesis of no treatment effect (flat treatment profile) is the ratio between treatment mean squares and pooled error mean squares,

$$F_B = \frac{N\bar{\mathbf{y}}'_{\cdot\cdot}\mathbf{P}_t\bar{\mathbf{y}}_{\cdot\cdot}}{\mathrm{tr}(\mathbf{P}_t\hat{\boldsymbol{\Sigma}})} = \frac{\mathrm{tr}\left[\mathbf{P}_t\mathbf{Y}'\left(\frac{1}{a}\mathbf{J}_a \otimes \frac{1}{n}\mathbf{J}_n\right)\mathbf{Y}\right]}{\frac{1}{a(n-1)}\mathrm{tr}\left[\mathbf{P}_t\mathbf{Y}'(\mathbf{I}_a \otimes \mathbf{P}_n)\mathbf{Y}\right]} = \frac{\mathbf{y}'\left(\frac{1}{a}\mathbf{J}_a \otimes \frac{1}{n}\mathbf{J}_n \otimes \mathbf{P}_t\right)\mathbf{y}}{\frac{1}{a(n-1)}\mathbf{y}'(\mathbf{I}_a \otimes \mathbf{P}_n \otimes \mathbf{P}_t)\mathbf{y}} \tag{11}$$

where $\bar{\mathbf{y}}_{\cdot\cdot} = (a^{-1}\mathbf{1}'_a \otimes n^{-1}\mathbf{1}'_n \otimes \mathbf{I}_t)\mathbf{y} = \mathbf{Y}'(a^{-1}\mathbf{1}_a \otimes n^{-1}\mathbf{1}_n)$ is the t-dimensional vector of treatment means of the response variable, and $\hat{\boldsymbol{\Sigma}}$ is defined in Theorem 1. The definition suggests that the three ratios in Eq. (11) are identical, which can be formulated as

Proposition 3. *The three different forms of F_B defined in display (11) are equivalent.*

The proof follows using the same techniques as in the proofs of Propositions 1 and 2.

In deriving the large a asymptotic results, the test for flat treatment profiles provides a situation that is somewhat different from those considered before. This is due to the fact that the hypotheses of parallel and equal treatment profiles were formulated in terms of a factor with number of levels tending to infinity. Here, however, as $a \to \infty$, the sample sizes $N = a \cdot n$ for each treatment increase, while the number t of factor levels being compared remains fixed. Thus, the test for flat profiles can be constructed using the classical asymptotic techniques for large sample sizes. We obtain first the asymptotic χ^2-distribution of the Wald-type statistic

$$Q_B = N\bar{\mathbf{y}}'_{\cdot\cdot}\mathbf{P}_t(\mathbf{P}_t\hat{\boldsymbol{\Sigma}}\mathbf{P}_t)^{-}\mathbf{P}_t\bar{\mathbf{y}}_{\cdot\cdot} \xrightarrow{d} \chi^2_{t-1}, \quad \text{as } a \to \infty.$$

Note that, while it was possible to derive an asymptotic normal distribution of the ANOVA-type statistic in the previous sections, this is not the case in the asymptotic situation encountered here. In fact, the asymptotic sampling distribution of F_B is that of a weighted sum of χ^2 random variables, where the weights are unknown. However, a good and practical approximation to the sampling distribution is provided by a scaled χ^2_c-distribution, where c is substituted by a moment estimator $\hat{c}$. That is, $F_B \cdot \hat{c}$ has approximately a $\chi^2_{\hat{c}}$-distribution, where $\hat{c} = [\mathrm{tr}(\mathbf{P}_t\hat{\boldsymbol{\Sigma}})]^2/[\mathrm{tr}(\mathbf{P}_t\hat{\boldsymbol{\Sigma}})^2]$. The performance of both Wald-type and ANOVA-type statistics will be compared in a simulation study in Section 5, alongside $F_B \sim F[t-1, a(n-1)(t-1)]$, which is the exact sampling distribution under normality and sphericity of the covariance matrix.

We have already emphasized that the above tests for flatness of average profiles are appropriate only if it can be assumed that the average profiles are parallel. On the other hand, one might be interested in testing flatness against the more general alternative that the average profiles are not flat. To that end, consider the normal theory likelihood ratio statistic derived by Harrar and Xu (2007): $\Lambda_{B|A} = \det(\mathbf{L}'\mathbf{S}_e\mathbf{L})/\det(\mathbf{L}'(\mathbf{S}_h^* + \mathbf{S}_e)\mathbf{L})$, where $\mathbf{S}_h^* = \mathbf{Y}'(\mathbf{I}_a \otimes n^{-1}\mathbf{J}_n)\mathbf{Y}$, $\mathbf{S}_e = \mathbf{Y}'(\mathbf{I}_a \otimes \mathbf{P}_n)\mathbf{Y}$, and $\mathbf{L}$ is the matrix that satisfies $\mathbf{P}_t = \mathbf{L}\mathbf{L}'$ such that $\mathbf{L}'\mathbf{L} = \mathbf{I}_{t-1}$. Then, we obtain the following result.

Theorem 3. *As $a \to \infty$, while n and t are fixed, $\Lambda_{B|A}$ follows asymptotically a normal distribution,*

$$\sqrt{a}\left[-n\log\Lambda_{B|A} + n(t-1)\log\left(\frac{n-1}{n}\right)\right] \xrightarrow{d} N\left(0, \frac{2n(t-1)}{n-1}\right).$$

The proof for Theorem 3 can be found in the Appendix.

3. Random environment factor

When the different environments can be regarded as a random sample, and when generalizations are to be made beyond those particular environments used in the series of experiments, it is appropriate to model the environment factor as random instead of fixed (see, e.g., Yates and Cochran, 1938; Comstock and Moll, 1963; Kuehl, 2000, p. 295; Caliński et al., 2005; Littell et al., 2006, Ch. 6). As we will see below, this may only have practical implications in the calculations for the test on flat treatment profiles (no treatment effect).

In the simple additive models stated in displays (1) and (2), the change from fixed to random environment factor is carried out by replacing the fixed components λ_i and $(\lambda\tau)_{ik}$ with mean zero random variables L_i and $(L\tau)_{ik}$.

In multivariate notation, the model with random environment effect can be decomposed as $\mathbf{Y} = \mathbf{1}_N\boldsymbol{\tau} + \mathbf{Z}_r\mathbf{B}_r + \mathbf{E}$. Here, $\boldsymbol{\tau} = (\tau_1, \ldots, \tau_t)$ is the vector of fixed treatment effects, $\mathbf{Z}_r = \mathbf{I}_a \otimes \mathbf{1}_n$ is the $(N \times a)$ design matrix for the random part, and $\mathbf{B}_r$ is the $(a \times t)$ matrix of random effects, while $\mathbf{E}$ is the error matrix. As in the fixed environment model, the rows of $\mathbf{E}$ are assumed to be independent with mean zero and covariance matrix $\boldsymbol{\Sigma}$. The matrix of random effects has in its ith row and kth column entry the element $L_i + (L\tau)_{ik}$.

A classical controversy in this context is with regard to whether identifiability constraints should be imposed on the interaction term $(L\tau)_{ik}$. For details on this discussion, we refer to the review paper by Samuels et al. (1991), and, particularly appropriate in the context of multi-environment trials, Basford et al. (2004). Concurring with these and other authors, we have chosen the restricted formulation in which the interaction terms are random variables that add up to zero across the levels of the fixed factor. This formulation allows for a more straightforward interpretation of model parameters and the variance component corresponding to $(L\tau)_{ik}$. However, note that the discussion of restricted versus unrestricted interaction only matters with regard to testing the environmental main effect, while tests for the other two hypotheses considered in this manuscript are unaffected. In this regard, see for example Tables 3 and 5 in McLean et al. (1991) (note the typographical error in Table 5 where in the row for source R, MSR/MSE should read MSR/MSFR).

Furthermore, we limit the discussion here to the population-based formulation of the random environment model. Alternatives are provided, for example, by restriction error models (Anderson, 1970), or randomization-based models (Kempthorne, 1952, p. 135; see also Caliński et al., 2005). Also, the additive model formulation (2) and the assumptions on the error term implicitly introduce a covariance structure in which pairs of observations from different blocks at the same environment have the same covariance. More general covariance structures were considered under normality by, for example, Denis et al. (1997), Piepho (1997), Caliński et al. (2005), and Smith et al. (2005).

Let us now consider the different hypotheses for which tests were derived under the fixed effects model in Section 2.

Parallel treatment profiles (no treatment–environment interaction). Under the null hypothesis of no interaction effect, the variance component $(L\tau)_{ik}$ is zero. It is straightforward to see, for example using model equations (1) and (2), along with the summation notation (7), that changing the environment effect from fixed to random does not change the null distribution of the test statistic F_{AB}. The environment effect simply cancels in both numerator and denominator of F_{AB}.

Equal profile averages (no environment effect). Similarly, under the null hypothesis of no environment effect, and assuming the restricted mixed model, the distribution of the test statistic F_A does not change either, when environment is assumed random instead of fixed. Note that the multivariate methods from Section 3.1 of Gupta et al. (2006) can be used to test the environment effect hypothesis in a model without interaction.

Flat treatment profiles (no treatment effect). The test statistic for treatment effect changes when the environment factor and its interaction with treatment are assumed random. In particular, in this case the correct denominator term is different from that in the fixed model. Instead of the pooled error mean squares, the appropriate denominator term is now given by the treatment by environment interaction mean squares. The resulting test statistic is

$$F_B^* = \frac{\operatorname{tr}\left[\mathbf{P}_t\mathbf{Y}'\left(\frac{1}{a}\mathbf{J}_a \otimes \frac{1}{n}\mathbf{J}_n\right)\mathbf{Y}\right]}{\frac{1}{a-1}\operatorname{tr}\left[\mathbf{P}_t\mathbf{Y}'\left(\mathbf{P}_a \otimes \frac{1}{n}\mathbf{J}_n\right)\mathbf{Y}\right]} = \frac{\mathbf{y}'\left(\frac{1}{a}\mathbf{J}_a \otimes \frac{1}{n}\mathbf{J}_n \otimes \mathbf{P}_t\right)\mathbf{y}}{\frac{1}{a-1}\mathbf{y}'\left(\mathbf{P}_a \otimes \frac{1}{n}\mathbf{J}_n \otimes \mathbf{P}_t\right)\mathbf{y}}.$$

This can also be rewritten as

$$F_B^* = \frac{\bar{\mathbf{y}}_{\cdot}'(\frac{1}{a}\mathbf{J}_a \otimes \mathbf{P}_t)\bar{\mathbf{y}}_{\cdot}}{\frac{1}{a-1}\bar{\mathbf{y}}_{\cdot}'(\mathbf{P}_a \otimes \mathbf{P}_t)\bar{\mathbf{y}}_{\cdot}}, \quad \text{where } \bar{\mathbf{y}}_{\cdot} = \left(\mathbf{I}_a \otimes \frac{1}{n}\mathbf{1}_n' \otimes \mathbf{I}_t\right)\mathbf{y}.$$

Clearly, this is identical to the F test for treatment effect in a single randomized complete block design where the responses $\bar{Y}_{i\cdot k},\ i = 1, \ldots, a,\ k = 1, \ldots, t$, which form the vector $\bar{\mathbf{y}}_{\cdot}$, correspond to averages over blocks within a environment $\times$ treatment combination of the original design, and the different blocks in the new design correspond to different environments in the original layout. Thus, it can be immediately seen that, under null hypothesis, $(t-1) \cdot F_B^*$ asymptotically has a χ^2_{t-1}-distribution, as $a \to \infty$. For small a, the distribution may be approximated by $F[t-1, (a-1)(t-1)]$.

4. Pairwise comparisons of treatments

If the test for treatment main effect from Section 2.3 provides evidence for an overall difference between the t treatments, the researcher is interested in a more detailed analysis, for example in form of an all-pairwise comparison (or in some cases multiple comparisons with a standard treatment). Such an analysis can be carried out by applying the test statistic F_B from Section 2.3 to every pair of treatments. Here, an appropriate multiple comparison adjustment (e.g., the Holm procedure, Holm, 1979) should be performed to ensure that the pairwise comparison procedure keeps the experimentwise error rate. Finally, after calculating the multiplicity adjusted p-values, a concise representation of all significant results can be generated using the letter-based algorithm introduced by Piepho (2004). This heuristic algorithm starts with a display that has the same letter for each treatment. Significant differences are sequentially inserted into this initial display until all significant pairwise differences are truthfully represented. Finally, a sweeping step removes excessive letters in the display. See Table 5 for results from applying this algorithm to an example data set. The letter-based algorithm uses a set of p-values that can be provided by any pairwise tests. Thus, it is as conservative, or as liberal, as the procedure used to calculate these p-values. We suggest to use a set of p-values that meets the experimentwise alpha level.

Note that the line displays or letter displays commonly available in most statistical software packages do not guarantee a correct display of all significant differences, unless the standard errors for all comparisons are constant, which is clearly not the case when the test statistic F_B is applied to each treatment pair. The algorithm by Piepho (2004) does not require this assumption and guarantees truthful representation of all significant differences.

5. Simulation study

We have conducted several simulation experiments to confirm the asymptotic results and study the finite performance of the tests described above. The simulations have been carried out with SAS 9.2, and each of the results presented here is

Table 1

Simulated α-levels [in percent] of tests for treatment–environment interaction, F_{AB}^{n}, F_{AB}^{c2}, F_{AB}^{c2Bru}, and F_{AB}^{sph}, $t = 4, 14$, $a = 5, 10, 15, 20, 30, 50$, and $n = 4, 20$, nominal α is 5%. Underlying distributions are multivariate normal. Covariance matrix for $t = 4$ is given in text $(\mathbf{\Sigma})$ (12); covariance structure for $t = 14$ has equal correlation 0.5 between variables, and variance j for the jth variable.

$t = 4$	F_{AB}^{n}		F_{AB}^{c2}		F_{AB}^{c2Bru}		F_{AB}^{sph}	
a	$n = 4$	$n = 20$	$n = 4$	$n = 20$	$n = 4$	$n = 20$	$n = 4$	$n = 20$
5	9.8	8.4	7.7	6.7	5.2	4.8	7.1	7.2
10	8.8	8.1	7.1	6.7	4.3	5.3	7.7	8.2
15	7.9	7.0	6.6	5.8	3.7	4.8	7.7	7.6
20	7.3	6.7	6.3	5.7	3.5	4.6	7.5	7.8
30	7.1	6.5	6.0	5.5	3.0	4.7	7.6	7.8
50	6.6	5.7	5.9	5.0	2.8	4.2	7.7	7.5
$t = 14$	F_{AB}^{n}		F_{AB}^{c2}		F_{AB}^{c2Bru}		F_{AB}^{sph}	
a	$n = 4$	$n = 20$	$n = 4$	$n = 20$	$n = 4$	$n = 20$	$n = 4$	$n = 20$
5	4.9	7.3	3.8	6.1	0.7	3.3	7.7	7.2
10	5.0	6.5	4.2	5.7	0.4	2.7	7.4	7.4
15	5.3	5.8	4.5	5.2	0.4	2.6	7.5	6.9
20	5.5	6.0	4.9	5.4	0.2	2.2	7.6	7.5
30	5.4	5.8	4.7	5.3	0.2	2.4	7.5	7.3
50	5.6	4.6	5.2	4.6	0.3	2.3	7.5	6.6

Table 2

Simulated α-levels [in percent] of tests for environment effect, F_{A}^{n}, F_{A}^{c2}, T_{A}^{n}, T_{A}^{F}, and F_{A}^{sph}, $t = 4, 14$, $a = 5, 10, 15, 20, 30, 50$, and $n = 4, 20$, nominal α is 5%. Underlying distributions are multivariate normal with covariance structures as in Table 1.

$t = 4$	F_{A}^{n}		F_{A}^{c2}		T_{A}^{n}		T_{A}^{F}		F_{A}^{sph}	
a	$n = 4$	$n = 20$	$n = 4$	$n = 20$	$n = 4$	$n = 20$	$n = 4$	$n = 20$	$n = 4$	$n = 20$
5	11.1	9.6	8.8	7.2	13.0	9.7	4.8	5.2	4.6	5.1
10	10.2	7.7	8.0	6.1	10.3	8.1	5.1	5.1	4.9	5.1
15	8.9	6.7	7.2	5.2	10.1	6.9	5.1	4.8	5.1	4.5
20	8.6	7.4	6.8	6.0	8.7	7.0	5.1	5.1	4.9	5.4
30	7.9	6.8	6.6	5.6	7.7	6.5	4.8	4.8	5.0	5.2
50	7.6	6.3	6.4	5.2	7.4	6.5	5.1	5.2	5.4	4.9
$t = 14$	F_{A}^{n}		F_{A}^{c2}		T_{A}^{n}		T_{A}^{F}		F_{A}^{sph}	
a	$n = 4$	$n = 20$	$n = 4$	$n = 20$	$n = 4$	$n = 20$	$n = 4$	$n = 20$	$n = 4$	$n = 20$
5	11.9	8.9	9.5	6.8	34.2	8.8	5.3	4.8	5.4	4.9
10	9.9	7.5	7.8	5.8	13.4	8.4	5.2	5.3	5.1	4.6
15	9.6	6.6	7.7	5.1	10.8	6.9	4.7	4.6	5.1	4.4
20	8.7	7.3	6.9	5.6	9.6	6.8	5.0	4.7	5.0	5.1
30	7.8	6.7	6.4	5.3	8.6	6.5	5.0	4.7	4.9	4.8
50	7.6	5.8	6.3	4.7	7.8	5.9	5.3	4.9	5.2	4.5

generally based on 10,000 simulation runs per setting (except for the configuration with $t = 14$, $a = 50$, and $n = 20$, where the number of simulations was 1000). The following tests have been investigated:

- The test statistic for testing parallel treatment profiles (no treatment–environment interaction) F_{AB}, compared to quantiles from a normal distribution as described in Theorem 1, and using χ^2 approximations with estimated degrees of freedom $\hat{f}$ and $\hat{f}_{Bru}$ (see Section 2.1). These tests are denoted as F_{AB}^{n}, F_{AB}^{c2}, and F_{AB}^{c2Bru}, respectively.
- The test statistics for testing equal profile averages (no environment effect) F_A and T_A. While F_A can be compared to normal or χ^2 quantiles, rejection regions for T_A can be based on the normal (Theorem 2) or F distribution (see Section 2.2). In the following these are denoted as F_{A}^{n}, F_{A}^{c2}, T_{A}^{n}, and T_{A}^{F}.
- The ANOVA-type and Wald-type tests for flat treatment profiles (no treatment effect), F_{B}^{c2} and Q_B, respectively, each using χ^2-quantiles (see Section 2.3).
- The test $\Lambda_{B|A}$ for simple effect of treatment against a general alternative, using normal quantiles as stated in Theorem 3.
- The F-tests using degrees of freedom that are valid under a normal model with compound symmetric (or spherical) covariance matrix structure. These are $F_{AB} \sim F[(a-1)(t-1), a(n-1)(t-1)]$, $F_A \sim F[a-1, a(n-1)]$, and $F_B \sim F[t-1, a(n-1)(t-1)]$, and they are denoted by the superscript *sph* in the following tables and graphs.

For the α-level simulations under null hypothesis, we have considered several combinations of values of a, n, and t. In the following, results are presented for a selection of values, namely all combinations of $a = 5, 10, 15, 20, 30, 50$, $n = 4, 20$, and $t = 4, 14$. The data has been generated as either multivariate normal or as coming from other distributions, with identical expectations per cell, and different correlation structures within block (but independence

Table 3
Simulated α-levels [in percent] of tests for main and simple treatment effect, F_B^{c2}, Q_B, $\Lambda_{B|A}$, and F_B^{sph}, $t = 4, 14$, $a = 5, 10, 15, 20, 30, 50$, and $n = 4, 20$, nominal α is 5%. Underlying distributions are multivariate normal with covariance structures as in Table 1.

$t = 4$	F_B^{c2}		Q_B		$\Lambda_{B\|A}$		F_B^{sph}	
a	$n = 4$	$n = 20$	$n = 4$	$n = 20$	$n = 4$	$n = 20$	$n = 4$	$n = 20$
5	6.0	5.2	13.4	5.8	13.2	6.9	6.8	6.8
10	4.8	5.1	8.2	5.6	9.6	6.9	5.9	6.6
15	5.2	5.1	7.5	5.4	8.9	6.6	6.5	6.8
20	5.1	5.0	6.5	5.1	8.3	6.2	6.5	6.7
30	5.3	4.6	6.4	5.1	7.2	5.8	6.8	6.2
50	4.8	4.9	5.7	5.0	6.9	6.1	6.2	6.5
$t = 14$	F_B^{c2}		Q_B		$\Lambda_{B\|A}$		F_B^{sph}	
a	$n = 4$	$n = 20$	$n = 4$	$n = 20$	$n = 4$	$n = 20$	$n = 4$	$n = 20$
5	3.0	4.4	91.9	13.4	98.2	13.4	7.1	6.8
10	3.9	4.8	46.4	8.7	66.7	10.2	7.2	6.9
15	4.0	4.8	28.7	6.7	47.9	9.2	7.1	6.7
20	4.7	4.9	21.3	6.2	38.7	7.9	7.6	6.8
30	4.7	5.0	14.2	5.8	28.3	8.0	7.1	7.0
50	4.8	5.8	9.8	5.5	20.6	6.6	6.9	7.2

Table 4
Analysis of spring barley data. Test statistics, degrees of freedom, and p-values using different methods discussed in this paper. A is environment, B is variety.

		Test statistic	Numerator df	Denominator df	p-value
Yield	F_{AB}^{c2}	2.632	62.8	∞	<0.0001
	F_{AB}^{sph}	2.632	252	266	<0.0001
	T_A^F	35.217	18	5	0.0005
	F_A^{sph}	52.022	18	19	<0.0001
	F_B^{c2}	26.304	6.6	∞	<0.0001
	F_B^{sph}	26.304	14	266	<0.0001
Ears	F_{AB}^{c2}	1.285	61.6	∞	0.0656
	F_{AB}^{sph}	1.285	252	266	0.0220
	T_A^F	12.745	18	5	0.0052
	F_A^{sph}	17.153	18	19	<0.0001
	F_B^{c2}	8.713	6.5	∞	<0.0001
	F_B^{sph}	8.713	14	266	<0.0001
MNAFG	F_{AB}^{c2}	1.305	39.3	∞	0.0956
	F_{AB}^{sph}	1.305	224	238	0.0218
	T_A^F	6.459	16	3	0.0750
	F_A^{sph}	52.397	16	17	<0.0001
	F_B^{c2}	3.251	4.6	∞	0.0077
	F_B^{sph}	3.251	14	238	0.0001

Table 5
Spring barley data. Average responses for the 15 varieties, and grouping results from letter-based pairwise comparison procedure, based on different test statistics discussed in this paper. For each test (column), varieties that share a letter did not perform significantly different at an experimentwise 5% level.

Variety	Yield			Ears			MNAFG		
	Mean	F_B^{c2}	F_B^{sph}	Mean	F_B^{c2}	F_B^{sph}	Mean	F_B^{c2}	F_B^{sph}
1	6.691	a	a	108.211	ab	abe	1.471	abc	ab
2	6.956	ab	ab	112.132	ade	abd	1.529	abc	ab
3	7.241	cd	bcd	107.132	ad	ad	1.500	abc	ab
4	7.068	bc	bc	111.553	ade	abd	1.588	bcd	ab
5	7.694	gi	fgi	128.974	f	g	1.618	ce	bd
6	7.517	efgh	defh	119.447	ef	bfg	1.500	abc	ab
7	7.368	de	de	104.868	a	ab	1.412	ab	ab
8	7.346	df	ce	113.184	ade	abd	1.471	ac	ad
9	7.614	fi	ei	107.605	ad	abd	1.471	abc	ab
10	7.409	dfg	ceg	116.816	de	def	1.412	a	ac
11	7.331	de	ce	105.789	ac	ac	1.471	abc	ab
12	7.767	i	gi	112.711	bcde	abd	1.471	ac	ad
13	7.386	cef	cef	110.868	ade	abd	1.441	ab	ab
14	7.466	df	def	112.211	ade	abd	1.618	be	bc
15	7.761	hi	hi	117.895	cdf	cdg	1.441	ad	ab

Table 6

Simulated α-levels [in percent] of tests for treatment–environment interaction, F^n_{AB}, F^{c2}_{AB}, F^{c2Bru}_{AB}, and F^{sph}_{AB}, $t = 4$, $a = 5, 10, 15, 20, 30, 50$, and $n = 4, 20$, nominal α is 5%. Underlying Beta (0.5, 0.5) and exponential (1) distributions, mixed correlation structure as in the upper part of Table 1.

Beta	F^n_{AB}		F^{c2}_{AB}		F^{c2Bru}_{AB}		F^{sph}_{AB}	
a	$n=4$	$n=20$	$n=4$	$n=20$	$n=4$	$n=20$	$n=4$	$n=20$
5	10.6	7.8	8.7	6.2	6.1	4.7	8.1	6.4
10	8.9	7.6	7.4	6.2	4.8	5.0	7.9	7.7
15	7.7	7.0	6.6	5.8	4.1	4.9	7.5	7.7
20	7.7	6.5	6.5	5.4	3.9	4.7	7.8	7.6
30	7.5	6.5	6.6	5.8	3.8	5.1	8.0	7.8
50	6.6	5.8	6.0	5.1	3.5	4.5	7.6	7.6
Exp	F^n_{AB}		F^{c2}_{AB}		F^{c2Bru}_{AB}		F^{sph}_{AB}	
a	$n=4$	$n=20$	$n=4$	$n=20$	$n=4$	$n=20$	$n=4$	$n=20$
5	8.4	8.0	6.7	6.3	3.9	4.0	6.4	6.9
10	8.0	7.5	6.5	6.1	2.7	3.7	7.2	7.6
15	7.1	7.1	5.8	5.9	2.0	3.7	7.0	7.8
20	7.9	6.5	6.7	5.5	2.1	3.8	8.3	7.5
30	7.2	6.3	6.1	5.5	2.0	3.6	8.0	7.8
50	6.9	6.2	6.1	5.4	1.6	3.2	8.1	7.8

Table 7

Simulated α-levels [in percent] of tests for environment effect, F^n_A, F^{c2}_A, T^n_A, T^F_A, and F^{sph}_A, $t = 4$, $a = 5, 10, 15, 20, 30, 50$, and $n = 4, 20$, nominal α is 5%. Underlying Beta (0.5, 0.5) and exponential (1) distributions, mixed correlation structure as in the upper part of Table 2.

Beta	F^n_A		F^{c2}_A		T^n_A		T^F_A		F^{sph}_A	
a	$n=4$	$n=20$	$n=4$	$n=20$	$n=4$	$n=20$	$n=4$	$n=20$	$n=4$	$n=20$
5	12.1	8.7	9.6	6.5	13.2	9.5	5.0	5.1	5.0	4.9
10	10.0	7.6	8.1	6.0	10.7	7.9	5.3	5.1	5.2	4.9
15	9.7	7.5	7.7	6.0	9.5	7.7	5.4	5.3	5.6	5.2
20	8.8	7.3	7.4	5.9	8.9	6.9	5.1	4.9	5.3	5.1
30	8.1	6.9	6.5	5.5	8.2	6.7	4.9	4.9	5.1	5.0
50	7.6	6.5	6.1	5.5	7.9	6.5	5.3	5.0	5.1	5.2
Exp	F^n_A		F^{c2}_A		T^n_A		T^F_A		F^{sph}_A	
a	$n=4$	$n=20$	$n=4$	$n=20$	$n=4$	$n=20$	$n=4$	$n=20$	$n=4$	$n=20$
5	11.8	8.9	9.0	6.7	12.9	8.8	5.0	4.9	4.7	4.6
10	10.1	7.6	8.1	5.8	10.9	7.7	5.3	4.6	5.1	4.7
15	9.2	7.5	7.5	5.9	9.7	7.6	5.4	5.1	5.1	5.1
20	8.8	7.0	7.1	5.6	9.3	6.7	5.7	4.6	5.3	4.9
30	7.4	6.8	6.2	5.6	8.1	6.6	5.1	4.9	4.9	5.2
50	7.5	6.4	6.6	5.4	7.6	6.4	5.3	4.9	5.5	5.1

across blocks). In particular, we have focused on covariance matrices that do not fall into the class of spherical structures (as, e.g., compound symmetry), in order to confirm that the newly derived test procedures remain valid under general covariance structures. Tables 1–3 show α-level simulation results for $t = 4$ and $t = 14$, and all combinations of the above values of a and n. The exemplary setting chosen for $t = 4$ is that of a multivariate normal distribution with covariance matrix

$$\mathbf{\Sigma} = \begin{bmatrix} 1.0 & 0.2 & 0.0 & 0.5 \\ 0.2 & 1.0 & 0.3 & -0.3 \\ 0.0 & 0.3 & 1.0 & -0.4 \\ 0.5 & -0.3 & -0.4 & 1.0 \end{bmatrix}, \tag{12}$$

while the results for $t = 14$ were obtained from a multivariate normal distribution with equal correlations of 0.5, but unequal variances: The jth diagonal element of the covariance matrix equals j. In general, the simulation results under null hypothesis varied only slightly across the different simulated distributions. Results for other distributions are presented in Tables 6–8.

For the power simulation results presented in this paper, we have chosen the setting $a = 15$, $n = 4$, $t = 4$, and created additive effects as follows. For the main effect simulations, observations in the ith factor level were shifted by $i \cdot x$, where x ranged between 0 and either 0.1 (environment) or 0.25 (treatment). For the interaction effect, the shift matrix $\begin{bmatrix} 1 & 0 & -1 \\ -1 & 0 & 1 \end{bmatrix} \otimes x \cdot \mathbf{J}(t/2, n \cdot a/3)$ was added to the $(N \times t)$ matrix of responses. Here, $\mathbf{J}(r, c)$ denotes a matrix of ones with dimension $r \times c$, and x ranged between 0 and 0.5.

The α-level simulation results suggest caution in using Q_B or $\Lambda_{B|A}$, which performed extremely liberal in situations where t was larger and a was moderate. Also, the tests based on asymptotic normality, F^n_A, T^n_A, as well as F^n_{AB} were generally not

Table 8
Simulated α-levels [in percent] of tests for main and simple treatment effect, F_B^{c2}, Q_B, $\Lambda_{B|A}$, and F_B^{sph}, $t = 4$, $a = 5, 10, 15, 20, 30, 50$, and $n = 4, 20$, nominal α is 5%. Underlying Beta (0.5, 0.5) and exponential (1) distributions, mixed correlation structure as in the upper part of Table 3.

Beta	F_B^{c2}		Q_B		$\Lambda_{B\|A}$		F_B^{sph}	
a	$n = 4$	$n = 20$	$n = 4$	$n = 20$	$n = 4$	$n = 20$	$n = 4$	$n = 20$
5	6.4	5.2	14.0	6.1	13.9	6.9	7.2	6.5
10	5.6	4.8	8.4	5.4	10.6	6.8	6.7	6.6
15	5.2	4.7	7.5	5.3	9.3	6.4	6.5	6.6
20	5.6	5.1	7.1	5.2	8.0	6.2	6.7	6.7
30	5.3	4.9	6.2	5.1	7.7	6.1	6.9	6.6
50	4.7	5.0	5.3	4.8	6.7	5.5	6.0	6.6
Exp	F_B^{c2}		Q_B		$\Lambda_{B\|A}$		F_B^{sph}	
a	$n = 4$	$n = 20$	$n = 4$	$n = 20$	$n = 4$	$n = 20$	$n = 4$	$n = 20$
5	5.6	5.0	16.0	6.4	13.0	7.3	6.7	6.5
10	5.2	4.6	10.3	5.8	9.3	6.3	6.7	6.0
15	5.4	4.7	8.3	5.5	8.4	6.5	6.8	6.6
20	5.0	4.6	7.4	5.7	8.2	6.0	6.6	6.3
30	4.9	4.5	6.5	5.0	7.5	6.0	6.3	6.2
50	4.9	4.6	6.3	4.7	7.1	5.9	6.7	6.2

as well approximated by their limiting sampling distribution, as the corresponding small-sample approximations. The F-tests assuming normality and spherical covariance matrix should not be used for testing interaction or treatment effect when these assumptions (in particular the sphericity) are violated: Both F_{AB}^{sph} and F_B^{sph} were generally too liberal in those simulations where the covariance structure was not spherical. On the other hand, these covariance structures did not have an effect on the simulated α-levels of the F-test for environment, F_A^{sph}. Thus, when testing the treatment–environment interaction hypothesis (parallel treatment profiles), only F_{AB}^{c2} should be considered. A similar test using a different estimated covariance matrix, F_{AB}^{c2Bru}, became too conservative in simulations with larger a or t. Regarding the test for an environment effect (equal profile averages), both T_A^F and F_A^{sph} met the α-level well in the simulation study, with F_A^{c2} requiring larger samples to meet the nominal level. When testing the treatment effect (flat treatment profiles), only F_B^{c2} remains as a valid choice, though it performed somewhat conservative in simulation settings where a was smaller than t.

Several additional simulations were run with the intention to closely imitate the data presented in Section 6. These data present a particular challenge since the number of blocks is rather small, $n = 2$. For $a = 19, n = 2, t = 15$ with positively correlated continuous response variable, as well as $a = 17, n = 2, t = 15$ with discrete response, we obtained the following range of simulated α-levels at a nominal $\alpha = 0.05$, across different continuous and discrete population distributions: F_{AB}^{c2} (from 0.03 to 0.08), F_{AB}^{sph} (0.08 to 0.19) T_A^F (0.047 to 0.055), F_A^{sph} (0.042 to 0.060), F_B^{c2} (0.03 to 0.06), F_B^{sph} (0.06 to 0.14).

Based on these results, we have selected those tests for power simulations that performed well in the α-level simulations. For each of interaction and treatment effect hypotheses, only one of the tests met the nominal level under the respective null hypothesis (F_{AB}^{c2}, F_B^{c2}). Therefore, the most interesting situation for a power comparison of different test statistics is the one in which the environment hypothesis is tested. Here, under null hypothesis both T_A^F and F_A^{sph} met the nominal level. On the other hand, F_A^{c2} was liberal, unless n and a were both moderate to large. The power simulations showed an advantage for T_A^F over F_A^{sph} across a large range of location shift alternatives (Fig. 1). Since F_{AB}^{sph} and F_B^{sph} were too liberal under null hypothesis, they should not be recommended under non-sphericity, and it is not surprising that they also displayed slightly higher power under alternative than their competitors (Figs. 2 and 3).

In conclusion, for each of the three hypotheses under consideration, it is possible to improve upon the common normal theory F-tests by choosing one of the newly proposed test statistics. While F_{AB}^{c2} and F_B^{c2} are the only ones that reliably met the α-level, T_A^F showed higher power than its normal theory counterpart.

6. Data example

To exemplify our methods, we analyzed a series of spring barley trials from official German variety tests conducted by the Bundessortenamt (BSA), Hannover. The data analyzed comprise trials from 19 locations in Germany which were performed in 2006. The trials were laid out with two different intensities, defined by the level of fertilization and plant protection. For each of the two intensities, 15 varieties were randomized according to a randomized complete block design (RCBD) with two replicates. The two replicates of the two intensities were completely randomized, resulting in a split-plot design with intensity as main plot factor. The treatment factor was barley variety. The central interest in the analyses of these trials is in the performance per intensity level, and thus analysis of series of trials is routinely done separately for the intensities according to the underlying RCBDs per environment. We followed the same practice, presenting results only for the low intensity level. The trait analyzed was kernel yield per plot (Yield). Two other response variables that we considered were

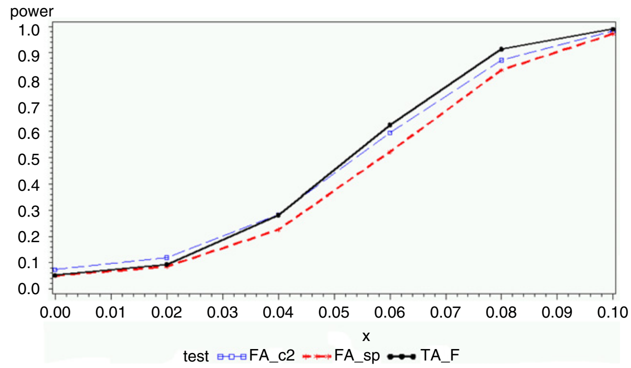

Fig. 1. Simulated power of tests for environment effect, F_A^{c2}, F_A^{sph}, and T_A^F, $t = 4$, $a = 15$, and $n = 4$, nominal α is 5%.

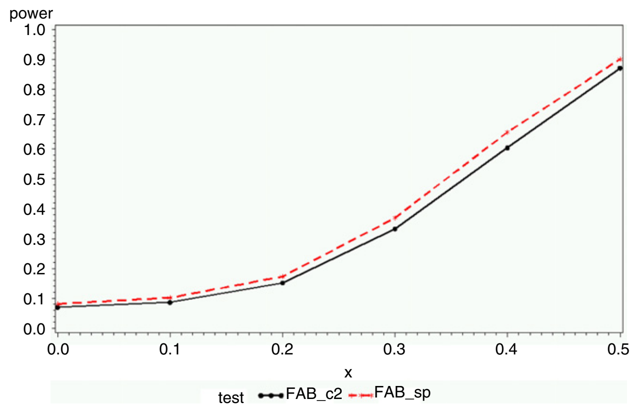

Fig. 2. Simulated power of tests for treatment–environment interaction, F_{AB}^{c2} and F_{AB}^{sph}, $t = 4$, $a = 15$, and $n = 4$, nominal α is 5%.

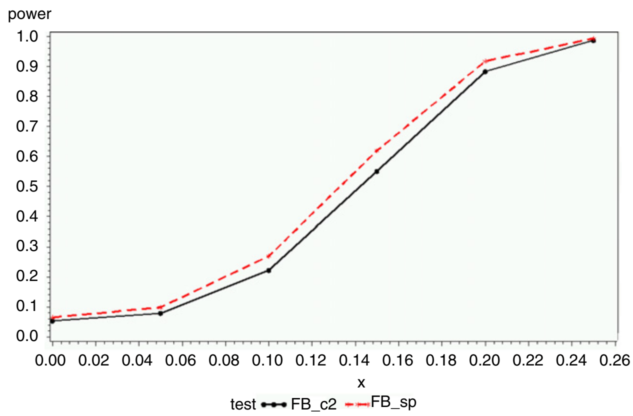

Fig. 3. Simulated power of tests for treatment effect, F_B^{c2} and F_B^{sph}, $t = 4$, $a = 15$, and $n = 4$, nominal α is 5%.

the number of ears per running meter (Ears), and an ordinal score for status of plants (MNAFG), which was obtained at 17 locations.

The results for an overall analysis of these variables are displayed in Table 4. Shapiro–Wilk tests for normality of the residuals after fitting a linear model resulted in p-values of $p < 0.0001$, $p = 0.0006$ (Ears), and $p < 0.0001$ (MNAFG), suggesting that the assumption of normality was questionable for all three variables.

The analysis provided strong evidence for presence of main effects due to both environment and variety in all three response variables, as well as in the case of Yield a strong interaction effect between the two factors. The interaction effect was borderline for Ears and MNAFG. Note that F_{AB}^{sph} rejected the interaction null hypotheses for both of these variables, but this test had been shown not to meet the α-level. On the other hand, F_{AB}^{c2} did not reject either of them.

In order to obtain more detailed information on the difference between the varieties, we have done a follow-up analysis using the letter-based algorithm mentioned in Section 4. To this end, the test statistic F_B^{c2} was calculated for each pair of varieties, the p-values were adjusted using the Holm procedure (Holm, 1979), and then fed into the algorithm that creates a display of all significant pairwise differences (Piepho, 2004). The resulting groupings for the different variables are shown in Table 5, together with the grouping based on p-values from F_B^{sph} instead of F_B^{c2}. Even though the latter test statistic met the nominal α-level more reliably in the simulation study, it resulted in more significant pairwise differences. In general, however, both test statistics led to similar groupings.

7. Discussion

In this paper, we have presented inferential methods for the analysis of data from multi-environment trials with many environments, or, more generally, from series of randomized complete block experiments. Here, we did not assume normality of the data, nor a particular within-block covariance structure (such as, for example, compound symmetry).

We have derived the asymptotic distribution of appropriate test statistics, and investigated the performance of different finite approximations in a simulation study. When testing treatment effect and treatment–environment interaction, the tests proposed in this manuscript performed in general more reliably than classical F-tests, which tended to become too liberal when the within-block covariance structure deviated from sphericity.

Future extensions of the work presented here involve generalization to covariance matrix heterogeneity across environments, and to tests based on rank-scores, which are expected to show superior performance in the presence of highly skewed or heavy-tailed response variable distributions. Also, recent work on dimensionally stable degrees of freedom estimators for high-dimensional data (Brunner, 2009; Chi and Muller, 2009) may lead to improvements particularly in the situation when the number of tested treatments is larger than the number of environments.

Acknowledgement

We wish to thank Dr. F. Laidig (Bundessortenamt Hannover, Germany) for providing the barley data.

Appendix

Proof of Proposition 1. The equivalence between (5) and (6) follows immediately from the identity $\mathrm{tr}(A'BCD') = (\mathrm{vec}\, A)'(D \otimes B)(\mathrm{vec}\, C)$ (see, e.g., Harville, 2007, Theorem 16.2.2).

Utilizing properties of the vec operator, the numerator of (3) can be rewritten as

$$\begin{aligned} N\bar{\mathbf{y}}_\cdot'(\mathbf{P}_a \otimes \mathbf{P}_t)\bar{\mathbf{y}}_\cdot &= N\,\mathrm{vec}\left[\mathbf{Y}'\left(\mathbf{I}_a \otimes \frac{1}{n}\mathbf{1}_n\right)\right]'(\mathbf{P}_a \otimes \mathbf{P}_t)\,\mathrm{vec}\left(\mathbf{Y}'\left(\mathbf{I}_a \otimes \frac{1}{n}\mathbf{1}_n\right)\right) \\ &= N\,\mathrm{vec}\left[\mathbf{Y}'\left(\mathbf{I}_a \otimes \frac{1}{n}\mathbf{1}_n\right)\right]'\mathrm{vec}\left[\mathbf{P}_t\mathbf{Y}'\left(\mathbf{I}_a \otimes \frac{1}{n}\mathbf{1}_n\right)\mathbf{P}_a\right] \\ &= N\,\mathrm{tr}\left[\left(\mathbf{I}_a \otimes \frac{1}{n}\mathbf{1}_n'\right)\mathbf{Y}\mathbf{P}_t\mathbf{Y}'\left(\mathbf{I}_a \otimes \frac{1}{n}\mathbf{1}_n\right)\mathbf{P}_a\right] = a\,\mathrm{tr}\left[(\mathbf{Y}\mathbf{P}_t\mathbf{Y}')\left(\mathbf{P}_a \otimes \frac{1}{n}\mathbf{J}_n\right)\right]. \end{aligned}$$

Substituting the empirical variance–covariance matrix $\hat{\mathbf{V}}$ that is used for the estimation of $\boldsymbol{\Sigma}$ by the expression for $\hat{\mathbf{V}}$ given above in display (4), we obtain the following denominator of (3).

$$\mathrm{tr}\left((\mathbf{P}_a \otimes \mathbf{P}_t)\left[\mathbf{I}_a \otimes \mathbf{Y}'\left(\mathbf{I}_a \otimes \frac{1}{n-1}\mathbf{P}_n\right)\mathbf{Y}\right]\right) = (a-1)\mathrm{tr}\left[\mathbf{P}_t\mathbf{Y}'\left(\mathbf{I}_a \otimes \frac{1}{n-1}\mathbf{P}_n\right)\mathbf{Y}\right].$$

Thus, the equivalence of (3) and (5) can be seen directly by assembling these terms as

$$\frac{N\bar{\mathbf{y}}_\cdot'(\mathbf{P}_a \otimes \mathbf{P}_t)\bar{\mathbf{y}}_\cdot}{\mathrm{tr}\left[(\mathbf{P}_a \otimes \mathbf{P}_t)\hat{\mathbf{V}}\right]} = \frac{a\,\mathrm{tr}\left[\mathbf{P}_t\mathbf{Y}'\left(\mathbf{P}_a \otimes \frac{1}{n}\mathbf{J}_n\right)\mathbf{Y}\right]}{(a-1)\mathrm{tr}\left[\mathbf{P}_t\mathbf{Y}'\left(\mathbf{I}_a \otimes \frac{1}{n-1}\mathbf{P}_n\right)\mathbf{Y}\right]} = \frac{\frac{1}{a-1}\mathrm{tr}\left[\mathbf{P}_t\mathbf{Y}'\left(\mathbf{P}_a \otimes \frac{1}{n}\mathbf{J}_n\right)\mathbf{Y}\right]}{\frac{1}{a(n-1)}\mathrm{tr}\left[\mathbf{P}_t\mathbf{Y}'(\mathbf{I}_a \otimes \mathbf{P}_n)\mathbf{Y}\right]}.$$

Now, consider numerator and denominator of (7). The between block mean sum of squares is

$$\begin{aligned} B_{AB} &= \frac{n}{(a-1)(t-1)}\sum_{i=1}^{a}\sum_{k=1}^{t}\left(\bar{Y}_{i\cdot k} - \bar{Y}_{i\cdot\cdot} - \bar{Y}_{\cdot\cdot k} + \bar{Y}_{\cdot\cdot\cdot}\right)^2 \\ &= \frac{1}{(a-1)(t-1)}\mathbf{y}'\left(\mathbf{P}_a \otimes \frac{1}{n}\mathbf{J}_n \otimes \mathbf{P}_t\right)\mathbf{y} \end{aligned}$$

and the within block mean sum of squares is

$$\begin{aligned}W_{AB} &= \frac{1}{at(n-1)}\sum_{i=1}^{a}\left[\sum_{j=1}^{n}\sum_{k=1}^{t}(Y_{ijk}-\bar{Y}_{i\cdot k})^2 - \frac{1}{t-1}\sum_{j=1}^{n}\sum_{k\neq k'}(Y_{ijk}-\bar{Y}_{i\cdot k})(Y_{ijk'}-\bar{Y}_{i\cdot k'})\right]\\
&= \frac{1}{a(n-1)(t-1)}\sum_{i=1}^{a}\left[\sum_{j=1}^{n}\sum_{k=1}^{t}(Y_{ijk}-\bar{Y}_{i\cdot k})^2 - t\sum_{j=1}^{n}(\bar{Y}_{ij\cdot}-\bar{Y}_{i\cdot\cdot})^2\right]\\
&= \frac{1}{a(n-1)(t-1)}\left[\mathbf{y}'(\mathbf{I}_a\otimes\mathbf{P}_n\otimes\mathbf{I}_t)\mathbf{y} - \mathbf{y}'\left(\mathbf{I}_a\otimes\mathbf{P}_n\otimes\frac{1}{t}\mathbf{J}_t\right)\mathbf{y}\right]\\
&= \frac{1}{a(n-1)(t-1)}\mathbf{y}'(\mathbf{I}_a\otimes\mathbf{P}_n\otimes\mathbf{P}_t)\mathbf{y}\end{aligned}$$

proving the equivalence of (7) and (6). □

Proof of Proposition 2. First note that the equivalences between (a), (b), and (c) can be seen as in the proof of Proposition 1. The numerator of (a) can be rewritten as (see, e.g., Harville, 2007, Section 16.2)

$$\begin{aligned}N\bar{\mathbf{y}}_{\cdot}'(\mathbf{P}_a\otimes\mathbf{J}_t)\bar{\mathbf{y}}_{\cdot} &= N\,\mathrm{vec}\left[\mathbf{Y}'\left(\mathbf{I}_a\otimes\frac{1}{n}\mathbf{1}_n\right)\right]'(\mathbf{P}_a\otimes\mathbf{J}_t)\,\mathrm{vec}\left[\mathbf{Y}'\left(\mathbf{I}_a\otimes\frac{1}{n}\mathbf{1}_n\right)\right]\\
&= N\,\mathrm{tr}\left[\left(\mathbf{I}_a\otimes\frac{1}{n}\mathbf{1}_n'\right)\mathbf{Y}\mathbf{J}_t\mathbf{Y}'\left(\mathbf{I}_a\otimes\frac{1}{n}\mathbf{1}_n\right)\mathbf{P}_a\right]\\
&= N\,\mathrm{tr}\left(\mathbf{1}_t\mathbf{Y}'\mathbf{P}_a\otimes\frac{1}{n^2}\mathbf{J}_n\mathbf{Y}\mathbf{1}_t\right)\\
&= a\mathbf{X}'\left(\mathbf{P}_a\otimes\frac{1}{n}\mathbf{J}_n\right)\mathbf{X}t^2,\end{aligned}$$

and the denominator can be expressed as

$$\begin{aligned}\mathrm{tr}\left[(\mathbf{P}_a\otimes\mathbf{J}_t)\hat{\mathbf{V}}\right] &= \mathrm{tr}\left((\mathbf{P}_a\otimes\mathbf{J}_t)\left[\mathbf{I}_a\otimes\mathbf{Y}'\left(\mathbf{I}_a\otimes\frac{1}{n-1}\mathbf{P}_n\right)\mathbf{Y}\right]\right)\\
&= \mathrm{tr}\left[\mathbf{P}_a\otimes\mathbf{J}_t\mathbf{Y}'\left(\mathbf{I}_a\otimes\frac{1}{n-1}\mathbf{P}_n\right)\mathbf{Y}\right]\\
&= (a-1)\mathbf{1}_t'\mathbf{Y}'\left(\mathbf{I}_a\otimes\frac{1}{n-1}\mathbf{P}_n\right)\mathbf{Y}\mathbf{1}_t\\
&= \frac{a-1}{n-1}\mathbf{X}'(\mathbf{I}_a\otimes\mathbf{P}_n)\mathbf{X}t^2.\end{aligned}$$

Clearly, (a) and (e) are identical.

Furthermore, the between environment mean sum of squares in (d) is

$$B_A = \frac{n}{a-1}\sum_{i=1}^{a}\sum_{k=1}^{t}(\bar{Y}_{i\cdot\cdot}-\bar{Y}_{\cdot\cdot\cdot})^2 = \frac{nt}{a-1}\sum_{i=1}^{a}(\bar{Y}_{i\cdot\cdot}-\bar{Y}_{\cdot\cdot\cdot})^2 = \frac{1}{a-1}\mathbf{y}'\left(\mathbf{P}_a\otimes\frac{1}{n}\mathbf{J}_n\otimes\frac{1}{t}\mathbf{J}_t\right)\mathbf{y},$$

and the within-environment mean sum of squares can be written as

$$\begin{aligned}W_A &= \frac{1}{at(n-1)}\sum_{i=1}^{a}\sum_{j=1}^{n}\sum_{k,k'}^{t}(Y_{ijk}-\bar{Y}_{i\cdot k})(Y_{ijk'}-\bar{Y}_{i\cdot k'}) = \frac{t}{a(n-1)}\sum_{i=1}^{a}\sum_{j=1}^{n}(\bar{Y}_{ij\cdot}-\bar{Y}_{i\cdot\cdot})^2\\
&= \frac{1}{a(n-1)}\mathbf{y}'\left(\mathbf{I}_a\otimes\mathbf{P}_n\otimes\frac{1}{t}\mathbf{J}_t\right)\mathbf{y},\end{aligned}$$

which finishes the proof of equivalence. □

Proof of Theorem 1. Using (5), the ANOVA-type test statistic for interaction can be written as

$$F_{AB} = \frac{\frac{1}{a-1}\mathrm{tr}\left[\mathbf{P}_t\mathbf{Y}'\left(\mathbf{P}_a\otimes\frac{1}{n}\mathbf{J}_n\right)\mathbf{Y}\right]}{\frac{1}{a(n-1)}\mathrm{tr}\left[\mathbf{P}_t\mathbf{Y}'(\mathbf{I}_a\otimes\mathbf{P}_n)\mathbf{Y}\right]}.$$

Then,

$$\begin{aligned}\sqrt{a}(F_{AB}-1) &= \frac{\sqrt{a}}{\hat{s}^2}\left(\frac{1}{a-1}\mathrm{tr}\left[\mathbf{P}_t\mathbf{Y}'\left(\mathbf{P}_a\otimes\frac{1}{n}\mathbf{J}_n\right)\mathbf{Y}\right]-\frac{1}{a(n-1)}\mathrm{tr}\left[\mathbf{P}_t\mathbf{Y}'(\mathbf{I}_a\otimes\mathbf{P}_n)\mathbf{Y}\right]\right)\\ &= \frac{\sqrt{a}}{\hat{s}^2}\mathrm{tr}\left(\mathbf{P}_t\mathbf{Y}'\left[\frac{1}{a(n-1)}\mathbf{I}_a\otimes(\mathbf{J}_n-\mathbf{I}_n)-\frac{1}{a(a-1)n}(\mathbf{J}_a-\mathbf{I}_a)\otimes\mathbf{J}_n\right]\mathbf{Y}\right),\end{aligned} \tag{13}$$

where $\hat{s}^2=\frac{1}{a(n-1)}\mathrm{tr}\left[\mathbf{P}_t\mathbf{Y}'(\mathbf{I}_a\otimes\mathbf{P}_n)\mathbf{Y}\right]=\frac{1}{a(n-1)}\mathrm{tr}\left(\mathbf{P}_t\sum_{i=1}^{a}\mathbf{Y}_i'\mathbf{P}_n\mathbf{Y}_i\right)$

is an almost surely consistent ($a\to\infty$) estimator of $\mathrm{tr}(\mathbf{P}_t\boldsymbol{\Sigma})$.

Considering the random term in the trace of (13), the second summand in the square brackets is of asymptotically ($a\to\infty$) vanishing order, so the limiting distribution is determined by

$$\sqrt{a}\mathbf{Y}'\left[\frac{1}{a(n-1)}\mathbf{I}_a\otimes(\mathbf{J}_n-\mathbf{I}_n)\right]\mathbf{Y}=\frac{1}{\sqrt{a}}\sum_{i=1}^{a}\frac{1}{n-1}\mathbf{Y}_i'(\mathbf{J}_n-\mathbf{I}_n)\mathbf{Y}_i.$$

Let $Q_i=\mathbf{Y}_i'(\mathbf{J}_n-\mathbf{I}_n)\mathbf{Y}_i$. Without loss of generality, the matrix $\mathbf{Y}$ can, under H_0, be taken to be centered, that is, we can assume that $\mathrm{E}(\mathbf{Y})=\mathbf{0}$. Then, it is clear that $\mathrm{E}(Q_i)=0$. Using Lemma 1 from Bathke and Harrar (2008, JSPI), the variance of Q_i can be calculated as follows.

$$\begin{aligned}\mathrm{Var}\left(\mathrm{vec}\left[\mathbf{Y}_i'(\mathbf{J}_n-\mathbf{I}_n)\mathbf{Y}_i\right]\right) &= \mathrm{Cov}\left(\mathrm{vec}\left[\mathbf{Y}_i'(\mathbf{J}_n-\mathbf{I}_n)\mathbf{Y}_i\right],\mathrm{vec}\left[\mathbf{Y}_i'(\mathbf{J}_n-\mathbf{I}_n)\mathbf{Y}_i\right]\right)\\ &= \sum_{j\neq j'}(1)(\mathbf{I}_{t^2}+K_{t,t})(\boldsymbol{\Sigma}\otimes\boldsymbol{\Sigma})=n(n-1)(\mathbf{I}_{t^2}+K_{t,t})(\boldsymbol{\Sigma}\otimes\boldsymbol{\Sigma}).\end{aligned}$$

Finally, using properties of the commutation matrix and the vec operator (see, e.g., Schott, 2005, Chapter 8)

$$\begin{aligned}\mathrm{Var}\left[\mathrm{tr}(\mathbf{P}_tQ_i)\right] &= \mathrm{vec}(\mathbf{P}_t)'\,\mathrm{Var}\left[\mathrm{vec}(Q_i)\right]\mathrm{vec}(\mathbf{P}_t)\\ &= n(n-1)\,\mathrm{vec}(\mathbf{P}_t)'(\mathbf{I}_{t^2}+K_{t,t})(\boldsymbol{\Sigma}\otimes\boldsymbol{\Sigma})\,\mathrm{vec}(\mathbf{P}_t)\\ &= 2n(n-1)\left[\mathrm{tr}(\mathbf{P}_t\boldsymbol{\Sigma})^2\right].\end{aligned}$$

Appealing to the Central Limit Theorem, we obtain an asymptotic normal distribution with expectation zero and variance

$$\frac{2n}{n-1}\frac{\mathrm{tr}(\mathbf{P}_t\boldsymbol{\Sigma})^2}{[\mathrm{tr}(\mathbf{P}_t\boldsymbol{\Sigma})]^2},$$

which can be consistently estimated by replacing $\boldsymbol{\Sigma}$ with $\hat{\boldsymbol{\Sigma}}=\frac{1}{a(n-1)}[\mathbf{Y}'(\mathbf{I}_a\otimes\mathbf{P}_n)\mathbf{Y}]$.

Altogether, we have derived the following asymptotic (large a) results for the interaction test.

$$\sqrt{a}(F_{AB}-1)\frac{1}{\sqrt{\hat{\tau}}}\xrightarrow{d}N(0,1)\quad\text{where }\hat{\tau}=\frac{2n}{n-1}\frac{\mathrm{tr}(\mathbf{P}_t\hat{\boldsymbol{\Sigma}})^2}{[\mathrm{tr}(\mathbf{P}_t\hat{\boldsymbol{\Sigma}})]^2}.\quad\square$$

Proof of Theorem 2. Define

$$\begin{aligned}\mathbf{U}_a &= \frac{1}{\sqrt{a}}\left[\boldsymbol{\Sigma}^{-1/2}\mathbf{Y}'\left(\mathbf{I}_a\otimes\frac{1}{n}\mathbf{J}_n\right)\mathbf{Y}\boldsymbol{\Sigma}^{-1/2}-\mathbf{I}_t\right]\quad\text{and}\\ \mathbf{V}_a &= \frac{1}{\sqrt{a}}\left[\boldsymbol{\Sigma}^{-1/2}\mathbf{Y}'(\mathbf{I}_a\otimes\mathbf{P}_n)\mathbf{Y}\boldsymbol{\Sigma}^{-1/2}-(n-1)\mathbf{I}_t\right].\end{aligned}$$

The expansions

$$\begin{aligned}\left(\frac{1}{a}\boldsymbol{\Sigma}^{-1/2}\mathbf{S}_e\boldsymbol{\Sigma}^{-1/2}\right)^{-1} &= \left((n-1)\mathbf{I}_t+\frac{1}{\sqrt{a}}\mathbf{V}_a\right)^{-1}\\ &= \frac{1}{n-1}\left(\mathbf{I}_t-\frac{1}{\sqrt{a}(n-1)}\mathbf{V}_a+O_p(a^{-1})\right)\end{aligned} \tag{14}$$

and

$$\begin{aligned}\left(\frac{1}{a}\boldsymbol{\Sigma}^{-1/2}\mathbf{S}_h\boldsymbol{\Sigma}^{-1/2}+\frac{1}{a}\boldsymbol{\Sigma}^{-1/2}\mathbf{S}_e\boldsymbol{\Sigma}^{-1/2}\right)^{-1} &= \left(n\mathbf{I}_t+\frac{1}{\sqrt{a}}(\mathbf{U}_a+\mathbf{V}_a)+O_p(a^{-1})\right)^{-1}\\ &= \frac{1}{n}\left(\mathbf{I}_t-\frac{1}{n\sqrt{a}}(\mathbf{U}_a+\mathbf{V}_a)+O_p(a^{-1})\right)\end{aligned} \tag{15}$$

can be easily verified (see Schott, 2005, Section 9.6). Consider

$$\begin{aligned} T_A^* &= \frac{1-\Lambda_A}{\Lambda_A} \\ &= \frac{\mathbf{1}_t'\boldsymbol{\Sigma}^{-1/2}\left(\frac{1}{a}\boldsymbol{\Sigma}^{-1/2}\mathbf{S}_e\boldsymbol{\Sigma}^{-1/2}\right)^{-1}\boldsymbol{\Sigma}^{-1/2}\mathbf{1}_t - \mathbf{1}_t'\boldsymbol{\Sigma}^{-1/2}\left(\frac{1}{a}\boldsymbol{\Sigma}^{-1/2}\mathbf{S}_h\boldsymbol{\Sigma}^{-1/2}+\frac{1}{a}\boldsymbol{\Sigma}^{-1/2}\mathbf{S}_e\boldsymbol{\Sigma}^{-1/2}\right)^{-1}\boldsymbol{\Sigma}^{-1/2}\mathbf{1}_t}{\mathbf{1}_t'\boldsymbol{\Sigma}^{-1/2}\left(\frac{1}{a}\boldsymbol{\Sigma}^{-1/2}\mathbf{S}_h\boldsymbol{\Sigma}^{-1/2}+\frac{1}{a}\boldsymbol{\Sigma}^{-1/2}\mathbf{S}_e\boldsymbol{\Sigma}^{-1/2}\right)^{-1}\boldsymbol{\Sigma}^{-1/2}\mathbf{1}_t}. \end{aligned}$$

Substituting (14) and (15), and after quite a bit of algebra,

$$\sqrt{a}\left\{(n-1)T_A^*-1\right\} = \sigma^2\left(u_a - \frac{1}{n-1}v_a\right) + O_p(a^{-1/2}),$$

where

$$\sigma^{-2} = \mathbf{1}_t'\boldsymbol{\Sigma}^{-1}\mathbf{1}_t, \qquad u_a = \mathbf{1}_t'\boldsymbol{\Sigma}^{-1/2}\mathbf{U}_a\boldsymbol{\Sigma}^{-1/2}\mathbf{1}_t \quad \text{and} \quad v_a = \mathbf{1}_t'\boldsymbol{\Sigma}^{-1/2}\mathbf{V}_a\boldsymbol{\Sigma}^{-1/2}\mathbf{1}_t.$$

Now,

$$\sigma^2\left(u_a - \frac{1}{n-1}v_a\right) = \frac{1}{\sqrt{a}\sigma^{-2}}\sum_{i=1}^{a}\left(\mathcal{E}_i\boldsymbol{\Sigma}^{-1/2}\mathbf{1}_t\right)'\left[\frac{1}{n-1}(\mathbf{J}_n-\mathbf{I}_n)\right]\left(\mathcal{E}_i\boldsymbol{\Sigma}^{-1/2}\mathbf{1}_t\right),$$

where $\mathcal{E}_i = \mathbf{Y}_i\boldsymbol{\Sigma}^{-1/2}$. Finally, noting that $\lim_{a\to\infty}(a-1)^{-1}(N-a-t+1) = n-1$ we get, by virtue of Lemma 2 of Bathke (2002),

$$\sqrt{a}\,(T_A-1) \xrightarrow{d} N\left(0, \frac{2n}{n-1}\right)$$

as $a \to \infty$. □

Proof of Theorem 3. Let $\mathbf{U}_a$ and $\mathbf{V}_a$ be as defined in the proof of Theorem 2. Then,

$$-\log\Lambda_{B|A} = -\log\frac{|\mathbf{L}'\mathbf{S}_e\mathbf{L}|}{|\mathbf{L}'(\mathbf{S}_h^*+\mathbf{S}_e)\mathbf{L}|} = \log|\mathbf{I}_{t-1}+\mathbf{L}'\mathbf{S}_h^*\mathbf{L}(\mathbf{L}'\mathbf{S}_e\mathbf{L})^{-1}|. \tag{16}$$

Notice that

$$\begin{aligned} &\frac{1}{a}(\mathbf{L}'\boldsymbol{\Sigma}\mathbf{L})^{-1/2}\mathbf{L}'\mathbf{S}_h^*\mathbf{L}(\mathbf{L}'\boldsymbol{\Sigma}\mathbf{L})^{-1/2} = \mathbf{I}_{t-1} + \frac{1}{\sqrt{a}}(\mathbf{L}'\boldsymbol{\Sigma}\mathbf{L})^{-1/2}\mathbf{L}'\boldsymbol{\Sigma}^{1/2}\mathbf{U}_a\boldsymbol{\Sigma}^{1/2}\mathbf{L}(\mathbf{L}'\boldsymbol{\Sigma}\mathbf{L})^{-1/2} \quad \text{and} \\ &\frac{1}{a}(\mathbf{L}'\boldsymbol{\Sigma}\mathbf{L})^{-1/2}\mathbf{L}'\mathbf{S}_e\mathbf{L}(\mathbf{L}'\boldsymbol{\Sigma}\mathbf{L})^{-1/2} = (n-1)\mathbf{I}_{t-1} + \frac{1}{\sqrt{a}}(\mathbf{L}'\boldsymbol{\Sigma}\mathbf{L})^{-1/2}\mathbf{L}'\boldsymbol{\Sigma}^{1/2}\mathbf{V}_a\boldsymbol{\Sigma}^{1/2}\mathbf{L}(\mathbf{L}'\boldsymbol{\Sigma}\mathbf{L})^{-1/2}. \end{aligned}$$

Substituting these in (16) and expanding the logarithm of the determinant in Taylor series yields,

$$\begin{aligned} -\log\Lambda_{B|A} &= -(t-1)\log\left(\frac{n-1}{n}\right) \\ &\quad + \frac{1}{n\sqrt{a}}\operatorname{tr}\left((\mathbf{L}'\boldsymbol{\Sigma}\mathbf{L})^{-1/2}\mathbf{L}'\boldsymbol{\Sigma}^{1/2}\left(\mathbf{U}_a - \frac{1}{n-1}\mathbf{V}_a\right)\boldsymbol{\Sigma}^{1/2}\mathbf{L}(\mathbf{L}'\boldsymbol{\Sigma}\mathbf{L})^{-1/2}\right) + O_p\left(\frac{1}{a}\right). \end{aligned}$$

Therefore,

$$\begin{aligned} &\sqrt{a}\left[-n\log\Lambda_{B|A} + n(t-1)\log\left(\frac{n-1}{n}\right)\right] \\ &= \frac{1}{\sqrt{a}}\sum_{i=1}^{a}\operatorname{tr}\left((\mathbf{L}'\boldsymbol{\Sigma}\mathbf{L})^{-1/2}\mathbf{L}'\Sigma^{1/2}\mathcal{E}_i'\left[\frac{1}{n-1}(\mathbf{J}_n-\mathbf{I}_n)\right]\mathcal{E}_i\Sigma^{1/2}\mathbf{L}(\mathbf{L}'\boldsymbol{\Sigma}\mathbf{L})^{-1/2}\right). \end{aligned}$$

The asymptotic ($a \to \infty$) distribution of the right hand side can be seen to be $N(0, 2n(t-1)/(n-1))$. □

References

Anderson, V.L., 1970. Restriction errors for linear models (An aid to develop models for designed experiments). Biometrics 26, 255–268.
Arnold, S.F., 1981. The Theory of Linear Models and Multivariate Analysis. Wiley, New York, USA.
Basford, K.E., Federer, W.T., DeLacy, I.H., 2004. Mixed model formulations for multi-environment trials. Agronomy Journal 96 (1), 143–147.
Bathke, A.C., 2002. ANOVA for a large number of treatments. Mathematical Methods of Statistics 11 (1), 118–132.
Bathke, A.C., Harrar, S.W., 2008. Nonparametric methods in multivariate factorial designs for large number of factor levels. Journal of Statistical Planning and Inference 138 (3), 588–610.

Bathke, A.C., Harrar, S.W., Madden, L.V., 2008. How to compare small multivariate samples using nonparametric tests. Computational Statistics and Data Analysis 52 (11), 4951–4965.
Boos, D.D., Brownie, C., 1995. ANOVA and rank tests when the number of treatments is large. Statistics and Probability Letters 23, 183–191.
Brunner, E., 2009. Repeated measures under non-sphericity, in: Proceedings of the Sixth St. Petersburg Workshop on Simulation, June 28–July 4, 2009, pp. 605–609.
Brunner, E., Langer, F., 1999. Nichtparametrische Analyse Longitudinaler Daten. Oldenbourg, München, Germany.
Caliński, T., Czajka, S., Kaczmarek, Z., Krajewski, P., Pilarczyk, W., 2005. Analyzing multi-environment variety trials using randomization-derived mixed models. Biometrics 61, 448–455.
Chi, Y.-Y., Muller, K.E., 2009. The univariate approach to repeated measures for high dimension, Low Sample Size, Preprint.
Comstock, R.E., Moll, R.H., 1963. Genotype-environment interaction. In: Statistical Genetics and Plant Breeding, Publication 982. National Academy of Sciences—National Research Council, Washington, DC, USA, pp. 164–196.
Denis, J.-B., Piepho, H.P., van Eeuwijk, F.A., 1997. Modelling expectation and variance for genotype by environment data. Heredity 79, 162–171.
Guerin, L., Stroup, W.W., 2000. A simulation study to evaluate proc mixed analysis of repeated measures data, in: Proceedings of the 12th Annual Conference on Applied Statistics in Agriculture, Kansas State University, Manhattan, KS, USA.
Gupta, A.K., Harrar, S.W., Fujikoshi, Y., 2006. Asymptotics for testing hypothesis in some multivariate variance components model under non-normality. Journal of Multivariate Analysis 97 (1), 148–178.
Harrar, S.W., 2009. Asymptotics for tests on mean profiles, additional information and dimensionality under non-normality. Journal of Statistical Planning and Inference 139 (8), 2511–2862.
Harrar, S.W., Bathke, A.C., 2008. A nonparametric version of the Bartlett–Nanda–Pillai multivariate test: Asymptotics, approximations, and applications. American Journal of Mathematical and Management Sciences 28 (3–4), 309–335.
Harrar, S.W., Bathke, A.C., 2008. Nonparametric methods for unbalanced multivariate data and many factor levels. Journal of Multivariate Analysis 99 (8), 1635–1664.
Harrar, S.W., Xu, J., 2007. On the null distribution of some test statistics for profile analysis under general conditions, University of Montana Technical Report, 11.
Harville, D.A., 2007. Matrix Algebra from a Statistician's Perspective, second ed. Springer, Berlin, Germany.
Holm, S., 1979. A simple sequentially rejective multiple test procedure. Scandinavian Journal of Statistics 6, 65–70.
Kempthorne, O., 1952. The Design and Analysis of Experiments. Wiley, New York, USA.
Kuehl, R.O., 2000. Design of Experiments: Statistical Principles of Research Design and Analysis, second ed. Brooks/Cole, Pacific Grove.
Littell, R.C., 2002. Analysis of unbalanced mixed model data: A case study comparison of ANOVA versus REML/GS. Journal of Agricultural, Biological, and Environmental Statistics 7 (4), 472–490.
Littell, R.C., Milliken, G.A., Stroup, W.W., Wolfinger, R.D., Schabenberger, O., 2006. SAS for Mixed Models, second ed. SAS Institute, Cary.
Liu, C., Bathke, A.C., Harrar, S.W., 2008. A nonparametric version of Wilks' lambda—Asymptotic results and small sample approximations. In: Proceedings of the 2008 Joint Statistical Meetings. American Statistical Association, Alexandria, VA, USA.
Loughin, T.M., 2006. Improved experimental design and analysis for long-term experiments. Crop Science 46, 2492–2502.
McIntosh, M.S., 1983. Analysis of combined experiments. Agronomy Journal 75 (1), 153–155.
McLean, R.A., Sanders, W.L., Stroup, W.W., 1991. A unified approach to mixed linear models. The American Statistician 45 (1), 54–64.
Piepho, H.P., 1997. Analyzing genotype-environment data by mixed models with multiplicative effects. Biometrics 53, 761–766.
Piepho, H.P., 2004. An algorithm for a letter-based representation of all-pairwise comparisons. Journal of Computational and Graphical Statistics 13, 456–466.
Rencher, A.C., 2002. Methods of Multivariate Analysis, second ed. Wiley, New York.
Samuels, M.L., Casella, G., McCabe, G.P., 1991. Interpreting blocks and random factors, with discussion. Journal of the American Statistical Association 86 (415), 798–808.
Schott, J.R., 2005. Matrix Analysis for Statistics. Wiley, New York.
Smith, A.B., Cullis, B.R., Thompson, R., 2005. The analysis of crop cultivar breeding and evaluation trials: An overview of current mixed model approaches. Journal of Agricultural Science 143, 449–462.
Speed, T.P., Williams, E.R., Patterson, H.D., 1985. A note on the analysis of resolvable block designs. Journal of the Royal Statistical Society. Series B 47 (2), 357–361.
Srivastava, M.S., 1987. Profile analysis of several groups. Communications in Statistics, Theory and Methods 16 (3), 909–926.
Theobald, C., Talbot, M., Nabugoomu, F., 2002. A Bayesian approach to regional and local-area prediction from crop variety trials. Journal of Agricultural, Biological, and Environmental Statistics 7, 403–419.
Yates, F., Cochran, W.G., 1938. The analysis of groups of experiments. Journal of Agricultural Science 28, 556–580.