Citation for the published publication:
Forkman, J & Piepho, H.P. (2015) Robustness of the simple parametric bootstrap method for the additive main effects and multiplicative interaction (AMMI) model. Presented at: The tenth working seminar on statistical methods in variety testing, Bedlewo, Poland, 30 June - 4 July, 2014, COBORU. This paper is also published in Biuletyn Oceny odmian.
http://www.coboru.pl/English/index_eng.aspx

Robustness of the simple parametric bootstrap method for the additive main effects and multiplicative interaction (AMMI) model

Published in Biuletyn Oceny Odmian, Vol. 34 (2015), p. 11–18.
http://www.coboru.pl

Johannes Forkman
Swedish University of Agricultural Sciences
Department of Crop Production Ecology
Box 7043, 750 07 Uppsala, Sweden
johannes.forkman@slu.se

Hans-Peter Piepho
University of Hohenheim
Institute of Crop Science
70 593 Stuttgart, Germany
hans-peter.piepho@uni-hohenheim.de

Summary: The additive main effects and multiplicative interaction (AMMI) model is a statistical model that is used for analysis of series of crop variety trials. This model can be fitted to a matrix of observations from a set of genotypes or crop varieties that have been investigated in a set of varying environments or locations. The model includes additive effects of genotypes and environments, and multiplicative effects of genotype-by-environment interaction. The multiplicative interaction terms are obtained through singular value decomposition. This paper describes the simple parametric bootstrap method, which can be used for testing significance of multiplicative terms. The simple parametric bootstrap method assumes that observations are normally distributed. Through simulation it is confirmed that the simple parametric bootstrap method performs well provided that the assumptions of normality and homogeneity of variance are fulfilled. However, when the distribution is non-normal, the frequency of Type I error is not maintained at the nominal significance level. The results of the simulation study suggest that a non-parametric bootstrap method would be needed.

Key words: Crop variety trials, genotype-environment interaction, heterogeneity, lognormal distribution, multi-environment trials, singular value decomposition

1. Introduction

Crop varieties are usually tested at several locations in order to investigate genotype-by-environment interaction. The additive main effects and multiplicative interaction (AMMI) model was proposed by Mandel (1971) and recommended for series of crop variety trials by Kempton (1984) and Gauch (1988, 1992). Using this model, genotype-by-environment interaction is studied through singular value decomposition. The interaction is written as a sum of multiplicative terms that are sorted according to decreasing importance. Thus, the first interaction term is the most important; the second is the second most important, and so on. Usually, only a few of the first terms are interpreted, since it is most likely that the last terms express random noise rather than any systematic pattern. Commonly in applications, two terms are retained, since this choice facilitates a two-dimensional graphical presentation in the form of a biplot (Gabriel, 1971). The adequacy of this rank-two approximation is often not assessed (Yang et al., 2009). Thus it might be that only the first multiplicative term is significant or that more than two terms would be needed to describe the actual interaction.

Forkman and Piepho (2014) reviewed tests of significance and proposed a new method to statistically test the multiplicative terms of the AMMI model. According to this *simple parametric bootstrap method*, the test statistic, which is computed from the data, is compared to a reference distribution that is derived through repeated sampling of matrices of standard normally distributed values. Section 2 describes the simple parametric bootstrap method for testing multiplicative terms in AMMI analyses of complete series of crop variety trials.

The simple parametric bootstrap method performs well with regard to power and frequency of Type I error (Forkman and Piepho, 2014). However, this result assumes that observations are normally distributed and homoscedastic. The present paper reports the results from a simulation study that investigates performance of the simple parametric bootstrap method when observations are non-normally distributed or heteroscedastic. Five distributions, which differ with regard to skewness, kurtosis and heterogeneity of variance, are investigated. Section 3 gives the details of the simulation study, and Section 4 presents the results.

2. The simple parametric bootstrap method

Let y_{ij} denote the observation from ith environment and jth genotype. Consider the additive main effects model

$$y_{ij} = \mu + \alpha_i + \beta_j + e_{ij}, \tag{1}$$

where μ is an intercept, and α_i and β_j are main effects of the ith environment and jth genotype, respectively, and e_{ij} is the residual associated with y_{ij}. When J genotypes have been observed in I environments, the observations y_{ij} can be compiled into an $I \times J$ matrix **Y**. Possible genotype-by-environment interaction effects are included in the errors e_{ij} in (1). AMMI analysis explores this interaction using singular value decomposition (SVD). The SVD is carried out on the $I \times J$ matrix $\hat{\mathbf{E}}$ of residuals from a fit of (1). Explicitly, the matrix $\hat{\mathbf{E}}$**,** with elements $\hat{e}_{ij}$, is obtained from **Y** through

$$\hat{e}_{ij} = y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..}.$$

where $\bar{y}_{i.} = \sum_{j=1}^{J} y_{ij}/J$, $\bar{y}_{.j} = \sum_{i=1}^{I} y_{ij}/I$ and $\bar{y}_{..} = \sum_{i=1}^{I}\sum_{j=1}^{J} y_{ij}/(IJ)$. The SVD of $\hat{\mathbf{E}}$ provides $M = \min(I-1,\ J-1)$ positive singular values $\hat{\lambda}_1, \hat{\lambda}_2, \ldots, \hat{\lambda}_M$. We shall use the convention that the first singular value is the largest; the second is the second largest, and so on. Thus $\hat{\lambda}_1 > \hat{\lambda}_2 > \ldots > \hat{\lambda}_M > 0$. The SVD also provides left- and right singular vectors corresponding to the singular values. We shall let $\hat{\gamma}_{ik}$ denote the kth element of the ith left-singular vector of $\hat{\mathbf{E}}$, and $\hat{\delta}_{jk}$ the kth element of the jth right-singular vector of $\hat{\mathbf{E}}$.

Let **E** denote the $I \times J$ matrix of elements e_{ij}. The simple parametric bootstrap method was derived using the model

$$\mathbf{E} = \mathbf{\Theta} + \mathbf{R}$$

where $\mathbf{\Theta}$ is a fixed part, and **R** is a random part. Specifically it was assumed that $\mathbf{\Theta}$ has κ singular values that are not zero and **R** is a matrix of independent identically normally distributed values. Under the null hypothesis, the AMMI model can be written as

$$y_{ij} = \mu + \alpha_i + \beta_j + \gamma_{i1}\lambda_1\delta_{j1} + \gamma_{i2}\lambda_2\delta_{j2} + \cdots + \gamma_{i\kappa}\lambda_\kappa\delta_{j\kappa} + r_{ij},$$

where $(\gamma_{i1}, \gamma_{i2}, \ldots, \gamma_{i\kappa})$ is the ith left-singular vector, $(\delta_{j1}, \delta_{j2}, \ldots, \delta_{j\kappa})$ is the $j$the right-singular vector, $\lambda_1, \lambda_2, \ldots \lambda_\kappa$ are the singular values, and r_{ij} is a normally distributed residual error. An AMMI model with κ multiplicative terms is called an AMMIκ model.

We are interested in the value κ, which can be regarded as the true dimensionality of the genotype-by-environment interaction. This is the dimensionality when the random errors (i.e. **R**) are ignored. Thus we would like to test the null hypothesis that κ equals some specified value K, that is,

$$H_0: \kappa = K,$$

against the alternative hypothesis

$$H_1\colon \kappa > K.$$

The simple parametric bootstrap method for AMMI analysis is performed as follows (Forkman and Piepho, 2014):

1. Compute the singular values $\hat{\lambda}_1, \hat{\lambda}_2, ..., \hat{\lambda}_M$ of $\hat{\mathbf{E}}$ and let $T = \hat{\lambda}_{K+1}^2 / \sum_{k=K+1}^{M} \hat{\lambda}_k^2$.
2. For b = 1, 2, ..., B, where B is large, do the following:
    1. Sample an $(I - 1 - K) \times (J - 1 - K)$ matrix $\hat{\mathbf{E}}_b^B$ of independent $\mathrm{N}(0, 1)$ distributed errors.
    2. Compute the singular values $\hat{\lambda}_{b1}, \hat{\lambda}_{b2}, ..., \hat{\lambda}_{bL}$ of $\hat{\mathbf{E}}_b^B$, where $L = \min(I - 1 - K,\ J - 1 - K)$, and from these compute $T_b = \hat{\lambda}_{b1}^2 / \sum_{k=1}^{L} \hat{\lambda}_{bk}^2$.
3. Estimate the p-value as the observed frequency of T_b larger than T.

3. Design of the simulation study

Three real datasets that were also investigated by Forkman and Piepho (2014) were used as a basis for the simulations:

1. The New York soybean dataset with seven genotypes investigated in ten environments (Gauch, 1992; Zobel et al., 1988).
2. The international maize and wheat improvement center (CIMMYT) maize dataset with nine genotypes in 20 environments (Cornelius et al., 1996; Crossa et al., 2002).
3. The Ontario winter wheat dataset with 18 genotypes in nine environments (Yan et al., 2007; Yan and Tinker, 2006; Yang et al., 2009).

For each real dataset, statistics were computed as specified in Table 1.

< Table 1 close to here >

Type I error rates were estimated through simulation. For each investigated distribution (see below) and model, 100,000 sample datasets were generated. The simplest model included only an intercept and a random error. With this model, the significances of the additive main effects were tested at significance level 0.05. For $K = 0, 1, 2, 3$ and 4, the AMMIK model estimated from the data was used for generation of sample datasets, and the significance of the $(K + 1)$th term was tested, also at level 0.05. The simple parametric bootstrap method used $B = 1000$ bootstrap samples.

Simulations were made using six distributions as specified by Table 2. These distributions differ with regard to skewness, kurtosis and heterogeneity of variance. For a random variable X, skewness and kurtosis are defined as $E((X-E(X))^3)/\sigma^3$ and $E((X-E(X))^4)/\sigma^4-3$, respectively, where σ is the square root of $E((X-E(X))^2)$. For the normal distribution, skewness and kurtosis are 0.

< Table 2 close to here >

The first distribution is a normal distribution with homogeneous variance. Forkman and Piepho (2014) used this distribution in their simulation, so this part is a repetition of their study. In biological data, the standard deviation is often increasing with the mean, whereas the coefficient of variation, that is, the standard deviation divided by the mean, is approximately constant. The second distribution is a normal distribution with homogeneous coefficient of variation.

When the coefficient of variation is approximately constant, it is often more plausible to assume a lognormal distribution than a normal. The third distribution is a lognormal distribution with homogeneous coefficient of variation. The lognormal distribution that was used for simulation when testing additive effects has expected value and variance equal to the estimates $\hat{\mu}$ and $\hat{\sigma}^2$, respectively, that are specified in Table 1. The coefficient of variation is $\hat{\sigma}/\hat{\mu}$. The lognormal distribution used for simulation when testing multiplicative effects has expected value $\hat{\mu}_{ij(K)}$ and variance $\hat{\sigma}^2_{(K)}\hat{\mu}^2_{ij(K)}/\hat{\mu}^2$. The coefficient of variation is $\hat{\sigma}_{(K)}/\hat{\mu}$. Thus, the coefficients of variation were the same for the lognormal distribution as for the heterogeneous normal distribution. The lognormal distribution is positively skewed. Table 3 lists coefficient of variation, skewness and kurtosis of the simulated lognormal distributions. For the lognormal distribution, skewness and kurtosis have limit 0 as the error variance on the logarithmic scale decreases. Thus, it is no coincidence that skewness and kurtosis decreases (Table 3) as the complexity of the model increases, that is, when more multiplicative terms are added.

< Table 3 close to here >

The fourth and fifth distributions are gamma(α,β) distributions, where α is the shape parameter, and β the scale parameter. The expected value of the gamma(α,β) distribution is $\alpha\beta$, and the variance is $\alpha\beta^2$. Thus, the used distributions are centered about $\hat{\mu}$ and $\hat{\mu}_{ij(K)}$ with variances $\hat{\sigma}^2$ and $\hat{\sigma}^2_{(K)}$ when evaluating tests of additive and multiplicative effects, respectively. Skewness and kurtosis of a gamma distribution with shape parameter $\alpha=4$ are 1 and $3/2$, respectively. The exponential distribution is a gamma distribution with shape parameter $\alpha=1$. This distribution has skewness 2 and kurtosis 6.

The sixth distribution is a uniform distribution centred about the mean. The variance of the $U(a,b)$ distribution is $(b-a)^2/12$, so the used distributions have variances $\hat{\sigma}^2$ and $\hat{\sigma}^2_{(K)}$, for tests of additive and multiplicative effects, respectively. The uniform distribution has skewness 0 and kurtosis $-6/5$.

4. Results of the simulation study

Table 4 presents the results of the simulation study in terms of observed frequencies of Type I errors.

< Table 4 close to here >

For each dataset, the first and the second rows refers to tests of main effects of environments and genotypes, respectively. These observed frequencies are close to the nominal significance level 0.05, which confirms the well-known (e.g. Moore and McCabe, 2003) robustness of analysis of variance F-tests.

The simple parametric bootstrap method performed well when the assumptions of normality and homoscedasticity were fulfilled. For all datasets and all null models, the Type I error rate was close to 0.05. This outcome was expected, since the same good performance of the simple parametric bootstrap method was observed by Forkman and Piepho (2014).

When the distribution was normal with a heterogeneous variance, the Type I error rate was consistently too high, but decreasing with the number of multiplicative terms. The lognormal distribution, with the same coefficients of variation as the heterogeneous normal distribution, performed similarly. The simulated lognormal distributions were only slightly skewed, so the high Type I error rates observed for the lognormal distribution was probably mainly due to heterogeneity.

The gamma distributions also produced too high frequencies of Type I errors. The gamma distribution with shape parameter $\alpha = 4$ performed better than the heavily skewed exponential distribution, that is, the gamma distribution with shape parameter $\alpha = 1$. Notice that in these cases, variance was homogeneous.

The uniform distribution produced too low Type I error rates. This distribution and the exponential are extreme cases that are not plausible in crop variety trials.

5. Conclusion

The results of the simulation study indicate that the simple parametric bootstrap method has excellent performance with regard to the Type I error rate, provided that the data is normally distributed and homoscedastic. When the distribution of the observations is non-normal or the data is heteroscedastic, the simple parametric bootstrap method does not maintain the nominal Type I error rate. In these situations, the reference distribution, which is obtained from sampling of normally distributed values, does not match the distribution of the test statistic. For this reason, it is in practice essential that the data is approximately normally distributed with approximately homogeneous variance. Observations might need to be transformed before analysis. It is an open question how a similar method performs if resampling is made from the correct distribution and not from a standard normal distribution. It is also concluded that a non-parametric bootstrap method for AMMI analysis would be useful.

Acknowledgement

H.P. Piepho acknowledges support by DFG grant PI 377/17-1.

References

Cornelius, P. L., Crossa, J., and Seyedsadr, M. S. (1996). Statistical tests and estimators of multiplicative models for genotype-by-environment interaction. In *Genotype-by-environment interaction* (M. S. Kang and H. G. Gauch, eds.), pp. 199-231. CRC Press, Boca Raton.

Crossa, J., Cornelius, P. L., and Yan, W. K. (2002). Biplots of linear-bilinear models for studying crossover genotype x environment interaction. *Crop Sci.* **42**, 619-633.

Forkman, J., and Piepho, H. P. (2014). Parametric bootstrap methods for testing multiplicative terms in GGE and AMMI models. *Biometrics* 70, 639–647.

Gabriel, K. R. (1971). The Biplot Graphic Display of Matrices with Application to Principal Component Analysis. *Biometrika* **58**, 453-467.

Gauch, H. G. (1988). Model selection and validation for yield trials with interaction. *Biometrics* **44**, 705-715.

Gauch, H. G. (1992). *Statistical analysis of regional yield trials: AMMI analysis of factorial designs*. Elsevier, Amsterdam.

Kempton, R. (1984). The use of biplots in interpreting variety by environment interactions. *J. Agric. Sci.* **103**, 123-135.

Mandel, J. (1971). A New Analysis of Variance Model for Non-Additive Data. *Technometrics* **13**, 1-18.

Moore, D. S., and McCabe, G. P. (2003). *Introduction to the Practice of Statistics,* 4th ed. W. H. Freeman, New York.

Yan, W., Kang, M. S., Ma, B. L., Woods, S., and Cornelius, P. L. (2007). GGE biplot vs. AMMI analysis of genotype-by-environment data. *Crop Sci.* **47**, 643-655.

Yan, W., and Tinker, N. A. (2006). Biplot analysis of multi-environment trial data: Principles and applications. *Canad. J. Plant Sci.* **86**, 623-645.

Yang, R. C., Crossa, J., Cornelius, P. L., and Burgueno, J. (2009). Biplot Analysis of Genotype x Environment Interaction: Proceed with Caution. *Crop Sci.* **49**, 1564-1576.

Zobel, R. W., Wright, M. J., and Gauch, H. G. (1988). Statistical analysis of a yield trial. *Agron. J.* **80**, 388-393.

Table 1

Statistic	Equation
Overall mean	$\hat{\mu} = \bar{y}_{..}$
Environment effect	$\hat{\alpha}_i = \bar{y}_{i.} - \bar{y}_{..}$
Genotype effect	$\hat{\beta}_j = \bar{y}_{.j} - \bar{y}_{..}$
Sum of **K** multiplicative terms	$\hat{\theta}_{ij(K)} = \begin{cases} 0, & K = 0 \\ \sum_{k=1}^{K} \hat{\gamma}_{ik}\hat{\lambda}_k\hat{\delta}_{jk}\,, & 1 \leq K \leq M-2 \end{cases}$
Fixed part of the AMMI**K** model	$\hat{\mu}_{ij(K)} = \hat{\mu} + \hat{\alpha}_i + \hat{\beta}_j + \hat{\theta}_{ij(K)}, \quad 0 \leq K \leq M-2$
Variance of the observations	$\hat{\sigma}^2 = \frac{\sum_{i=1}^{I}\sum_{j=1}^{J}\left(y_{ij} - \bar{y}_{..}\right)^2}{IJ-1}$
Variance of the AMMI**K** model	$\hat{\sigma}_{(K)}^2 = \frac{1}{(I-1)(J-1)}\sum_{k=K+1}^{M}\hat{\lambda}_k^2\,, \quad 0 \leq K \leq M-2,$

Table 2

Distribution	Test of additive effects	Test of multiplicative effects
1. Normal distribution with homogeneous variance	$y_{ij}\sim\mathrm{N}(\hat{\mu},\hat{\sigma}^2)$	$y_{ij}\sim\mathrm{N}\left(\hat{\mu}_{ij(K)},\hat{\sigma}^2_{(K)}\right)$
2. Normal distribution with homogeneous coefficient of variation	$y_{ij}\sim\mathrm{N}(\hat{\mu},\hat{\sigma}^2)$	$y_{ij}\sim\mathrm{N}\left(\hat{\mu}_{ij(K)},\gamma^2_{(K)}\hat{\mu}^2_{ij(K)}\right)$, $\gamma_{(K)}=\hat{\sigma}_{(K)}/\hat{\mu}$
3. Lognormal distribution with homogeneous coefficient of variation	$y_{ij}\sim\ln\mathrm{N}(m,s^2)$, $s^2=\log(\hat{\sigma}^2/\hat{\mu}^2+1)$, $m=\log(\hat{\mu})-s^2/2$	$y_{ij}\sim\ln\mathrm{N}\left(m_{(K)},s^2_{(K)}\right)$, $s^2_{(K)}=\log(\hat{\sigma}^2_{(K)}/\hat{\mu}^2+1)$, $m_{(K)}=\log\hat{\mu}_{ij(K)}-s^2_{(K)}/2$
4. Gamma distribution with shape parameter $\boldsymbol{\alpha}=\mathbf{4}$	$y_{ij}\sim\hat{\mu}+X-2\hat{\sigma}$, $X\sim\mathrm{gamma}(4,\hat{\sigma}/2)$	$y_{ij}\sim\hat{\mu}_{ij(K)}+X-2\hat{\sigma}_{(K)}$, $X\sim\mathrm{gamma}\left(4,\hat{\sigma}_{(K)}/2\right)$
5. Exponential distribution	$y_{ij}\sim\hat{\mu}+X-\hat{\sigma}$, $X\sim\mathrm{gamma}(1,\hat{\sigma})$	$y_{ij}\sim\hat{\mu}_{ij(K)}+X-\hat{\sigma}_{(K)}$, $X\sim\mathrm{gamma}\left(1,\hat{\sigma}_{(K)}\right)$
6. Uniform distribution	$y_{ij}\sim\hat{\mu}+R$, $R\sim\mathrm{U}\left(-\sqrt{3\hat{\sigma}^2},\sqrt{3\hat{\sigma}^2}\right)$	$y_{ij}\sim\hat{\mu}_{ij(K)}+\mathrm{R}$, $\mathrm{R}\sim\mathrm{U}\left(-\sqrt{3\hat{\sigma}^2_{(K)}},\sqrt{3\hat{\sigma}^2_{(K)}}\right)$

Table 3

Model under $\boldsymbol{H_0}$	CV	Skewness	Kurtosis
New York soybean data			
Intercept	0.336	1.05	2.00
AMMI$\mathbf{0}$	0.160	0.48	0.42
AMMI$\mathbf{1}$	0.067	0.20	0.07
AMMI$\mathbf{2}$	0.038	0.12	0.02
AMMI$\mathbf{3}$	0.029	0.09	0.01
AMMI$\mathbf{4}$	0.017	0.05	0.00
CIMMYT maize data			
Intercept	0.279	0.86	1.34
AMMI$\mathbf{0}$	0.132	0.40	0.28
AMMI$\mathbf{1}$	0.087	0.26	0.12
AMMI$\mathbf{2}$	0.071	0.21	0.08
AMMI$\mathbf{3}$	0.056	0.17	0.05
AMMI$\mathbf{4}$	0.041	0.12	0.03
Ontario winter wheat data			
Intercept	0.236	0.72	0.94
AMMI$\mathbf{0}$	0.091	0.28	0.13
AMMI$\mathbf{1}$	0.066	0.20	0.07
AMMI$\mathbf{2}$	0.049	0.15	0.04
AMMI$\mathbf{3}$	0.040	0.12	0.03
AMMI$\mathbf{4}$	0.032	0.10	0.02

Table 4

Model under $\boldsymbol{H_0}$	Normal homogen.	Normal heterogen.	Lognormal	Gamma $\boldsymbol{\alpha = 4}$	Gamma $\boldsymbol{\alpha = 1}$	Uniform
			New York soybean data			
Intercept[1]	0.050	0.051	0.049	0.048	0.046	0.051
Intercept[2]	0.050	0.050	0.049	0.048	0.045	0.051
AMMI**0**	0.049	0.187	0.207	0.100	0.252	0.023
AMMI**1**	0.051	0.144	0.146	0.084	0.191	0.028
AMMI**2**	0.050	0.086	0.086	0.074	0.147	0.033
AMMI**3**	0.049	0.073	0.072	0.063	0.102	0.040
AMMI**4**	0.049	0.060	0.060	0.056	0.073	0.047
			CIMMYT maize data			
Intercept[1]	0.049	0.050	0.050	0.050	0.049	0.051
Intercept[2]	0.050	0.051	0.050	0.051	0.047	0.050
AMMI**0**	0.050	0.190	0.210	0.113	0.329	0.020
AMMI**1**	0.050	0.179	0.184	0.099	0.272	0.024
AMMI**2**	0.048	0.142	0.146	0.089	0.224	0.026
AMMI**3**	0.048	0.120	0.120	0.082	0.189	0.029
AMMI**4**	0.049	0.110	0.111	0.076	0.161	0.033
			Ontario winter wheat data			
Intercept[1]	0.050	0.049	0.049	0.050	0.047	0.050
Intercept[2]	0.049	0.050	0.050	0.049	0.048	0.051
AMMI**0**	0.050	0.228	0.237	0.114	0.320	0.021
AMMI**1**	0.049	0.178	0.179	0.098	0.268	0.024
AMMI**2**	0.049	0.170	0.169	0.091	0.227	0.027
AMMI**3**	0.049	0.127	0.128	0.081	0.190	0.028
AMMI**4**	0.048	0.120	0.121	0.074	0.160	0.030

[1] On this row, Type I error rates refers to tests of main effects of environments

[2] On this row, Type I error rates refers to tests of main effects of genotypes

Table titles

Table 1. Definition of statistics computed for the three datasets that were used as a basis for the simulation study

Table 2. Distributions used in the simulation study

Table 3. Coefficient of variation (CV) assumed for the heterogeneous normal and lognormal distributions, and skewness and kurtosis of the lognormal distributions used in the simulation study

Table 4. Observed Type I error rates of the simple parametric bootstrap method for the six distributions that were specified in Table 2