sciendo DOI: 10.2478/bile-2022-0003

Biometrical Letters
Vol. 59 (2022), No. 1, 23-46

Maximum Likelihood and L_2 Environmental Indices in Joint Regression Analysis

Dulce Gamito Pereira

Departamento de Matemática, Escola de Ciências e Tecnologia, Centro de Investigação em Matemática e Aplicações, Instituto de Investigação e Formação Avançada, Universidade de Évora, Rua Romão Ramalho 59, 7000-671 Évora, Portugal
e-mail: dgsp@uevora.pt

SUMMARY

This paper describes an iterative analysis of incomplete genotype × environment data. *L_2 environmental indices* were introduced to enable the use of Joint Regression Analysis (JRA) in analyzing experiments with incomplete blocks. We now show how, once normality of yields is assumed, the introduction of L_2 environmental indices provides a theoretical framework for Joint Regression Analysis. Using this framework, maximum likelihood estimators are obtained and likelihood ratio tests are derived. It is noted that the technique allows unequal weighting of data, and the special case of complete blocks is discussed.

Keywords: Joint Regression Analysis, maximum likelihood estimators, likelihood ratio tests, L_2 environmental indices.

1. Introduction

In their review of the literature on genotype × environment interaction, Aastveit & Mejza (1992) placed Joint Regression Analysis (JRA) in the group of models used for response stability when the environment changes. This group of models was further divided into:

- models based on regression techniques;
- models based on analysis of variance (ANOVA).

JRA clearly belongs to the first subgroup. More about genotype × environment interaction can be found in Kang & Gauch (1996).

JRA has been widely used in the analysis of *series of experiments* designed for cultivar comparison. These experiments must cover as wide an area as possible in order to increase the usability of the selected cultivars. Gusmão et al. (1989) introduced the concept of equipotential zones for cultivar yield pattern evolution in this context. Equipotential zones will be regions, as large as possible, in which JRA can be carried out.

The crucial problem in methods based on regression is the choice of the controlled variables, since the yield is usually taken as a dependent variable. In Linear Joint Regression Analysis (LJRA) a synthetic variable, the *environmental index*, is used to measure the productive capacity of each pair (location, year).

Initially conceived by Mooers (1921) and further developed by Yates & Cochran (1938), JRA was greatly improved by Finlay & Wilkinson (1963), who made the method practicable. The environmental index as introduced by Eberhart & Russel (1966) was measured, for each experiment, by its average yield. Next, for example, Gusmão (1985, 1986a, 1986b) showed that the precision in analyzing series of randomized block experiments was greatly increased by considering environmental indices for individual blocks instead of only one environmental index per experiment. With this proposal, if we have K experiments each with b blocks, we will have Kb supporting points per regression instead of only K such points.

Later on, LJRA was criticized by Westcott (1986) and Lin et al. (1986) for not considering specific environmental variables, but other authors, such as Becker & Leon (1988), argued that this criticism was not strong enough to justify discarding LJRA. In fact, the holistic approach to productive capacity integrated into this technique has performed quite well in the Portuguese case. Based on results from regional yield experiments and several simulated studies, Baeta et al. (1990) illustrated how genotype × environment interaction (as expressed by the cultivar × block interaction) may be responsible for the increase in estimates of the experimental error in the ANOVA of randomized complete block design (RCBD) experiments. In order to overcome such difficulties, and to increase the accuracy of the statistical analysis of RCBD trials, practical solutions were suggested, for the case when the F test for blocks is not significant (Gusmão et al. 1990) and when it is significant (Mexia et al. 1990, 1991).

Formerly, LJRA was applied to series of randomized block experiments. The block environmental indices were estimated, following Gusmão (1985, 1986a, 1986b, 1988), by the block average yields. This procedure had several shortcomings:

- it can only be used when every block contains all cultivars;
- it is difficult to accept an average yield as a controlled variable.

Despite these limitations, LJRA has, as pointed out above, led to quite good results.

Both of the shortcomings of LJRA mentioned above are overcome when *L_2 environmental indices* are used. We will present an algorithm for their estimation when incomplete block designs are used. In each block, every cultivar is absent or has a single replicate. This means that we will consider binary block designs only.

For completeness' sake, we point out that Digby (1979) gave an iterative numerical method for fitting a joint regression model to an incomplete two-way table without replicates, but did not analyze the statistical properties of the obtained estimators. Ng & Grunwald (1997) gave another iterative method for the adjustment of joint regressions. Both methods were compared in Ng & Williams (2001). Applying our technique to the data in this paper, we obtained comparable results.

Using a model similar to the one introduced by Digby (1979) and assuming the yields to be normal and homoscedastic, we show that the zigzag algorithm leads to maximum likelihood estimators. Moreover, we derive likelihood ratio tests for the degree of the regressions to be adjusted and for equality of the coefficients. As to the degree, we point out that Mexia et al. (1999, 2001) showed that the precision might only slightly increase with the degree of the adjusted polynomials. Thus, the corresponding test is included only for completeness' sake.

2. Model

We assume, following Gusmão (1985), that for each block in a series of experiments there is a distinct environment. Moreover, we restrict ourselves to linear regressions, since (see Mexia et al. 2001) no great improvement was obtained by considering polynomial regression. Thus if the j^{th} cultivar is present in the i^{th} block, its mean yield will be expressed as

$$E\left(Y_{ij}\right)=\alpha_j+\beta_j x_i \ , j=1,...,J,\ i=1,...,b,$$

where α_j and β_j are the regression coefficients for the j^{th} cultivar, and x_i , $i=1,...,b$ is the environmental index for the i^{th} block. We are thus led to the model

$$Y_{ij}=\alpha_j+\beta_j x_i+\varepsilon_{ij} \ , j=1,...,J,\ i=1,...,b,$$

very similar to that proposed by Digby (1979). The main difference is that we will assume the error ε_{ij} to be normal and homoscedastic. This assumption will enable us to obtain maximum likelihood estimators and to perform likelihood ratio tests.

To adjust this model by least squares, we seek to minimize

$$S\left(\boldsymbol{\alpha}^J,\boldsymbol{\beta}^J,\mathbf{x}^b\right)=\sum_{i=1}^{b}\sum_{j=1}^{J}p_{ij}\left(Y_{ij}-\alpha_j-\beta_j x_i\right)^2 .$$

The weights p_{ij} are 1 when the j^{th} cultivar is present in the i^{th} block and 0 otherwise.

To carry out the minimization, the zigzag algorithm (see Mexia et al. 1999) may be used. This is an iterative algorithm that starts with a set of initial values $x_{0,i}$, $i=1,...,b$, for the environmental indices. It is clear that the choice of these values is important. In the complete case we can often start (see Mexia et al. 1999) with the block mean yields. Moreover, if we have a design with resolvable super-blocks, we can use, for each block, the average yield of the corresponding super-block.

In each iteration, minimization is first carried out for the vectors $\boldsymbol{\alpha}^J$ and $\boldsymbol{\beta}^J$ of coefficients, and subsequently for the vector $\mathbf{x}^b$ of environmental indices. Lastly, the vector $\mathbf{x}^b$ is standardized, so that the range of environmental indices is kept unchanged.

The zigzag algorithm is summarized in the following scheme:

ZIGZAG ALGORITHM

Calculation of the initial values for the environmental indices $\mathbf{x}_0^b$

Range $[a_0\ ;\ b_0]$ with $a_0 = Min\{x_{01},...,x_{0b}\}$ and $b_0 = Max\{x_{01},...,x_{0b}\}$

↓

Minimize the function $S\left(\boldsymbol{\alpha}^J,\boldsymbol{\beta}^J \middle| \mathbf{x}_0^b\right)$: $\begin{cases} \tilde{\alpha}\left(\mathbf{x}_0^b\right) \\ \tilde{\beta}\left(\mathbf{x}_0^b\right) \end{cases}$

↓

To minimize $S\left(\mathbf{x}_J^b \middle| \tilde{\boldsymbol{\alpha}}^J\left(\mathbf{x}_0^b\right),\tilde{\boldsymbol{\beta}}^J\left(\mathbf{x}_0^b\right)\right)$ minimize the functions $h_i\left(x \middle| \tilde{\boldsymbol{\alpha}}^J,\tilde{\boldsymbol{\beta}}^J\right) = \sum_{j=1}^{J} p_{ij}\left(Y_{ij} - \tilde{\alpha}_j - \tilde{\beta}_j x_i\right)^2$, $i = 1,...,b$, to obtain the new vector $\mathbf{x}'^b_0$.

↓

Standardize the vector of environmental indices to ensure unchanged range, with $a'_0 = Min\{x'_{01},...,x'_{0b}\}$; $b'_0 = Max\{x_{01},...,x'_{0b}\}$

Take $x_{1i} = a_0 + \dfrac{b_0 - a_0}{b'_0 - a'_0}\left(x'_{0i} - a'_0\right)$

↓

Thus we obtain the vector $\mathbf{x}_1^b$.

↓

Repeat this procedure until successive sums of sums of squares of weighted residuals differ by less than a fixed amount.

In the appendix, we consider the behavior of this algorithm.

We point out that (see Pereira & Mexia 2003) we could use other non-negative weights in order to take into account the agronomical relevance of the different blocks.

3. Maximum likelihood estimators

Since we work with large samples, and the main goal of JRA is connected with inference in the regression coefficients, and not with variance components, we will use the usual maximum likelihood estimators.

Let us now assume that the yields are independent and normally distributed, with variances σ^2. We take the pairs (α_j, β_j), $j = 1,..., J$, of coefficients, θ_i, $i = 1,...,b$ the environmental indices, and N the number of observed yields. Now the likelihood function will be

$$L\left(\boldsymbol{\alpha}^J, \boldsymbol{\beta}^J, \boldsymbol{\theta}^b, \sigma^2\right) = \frac{e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{b}\sum_{j=1}^{J} p_{ij}\left(Y_{ij}-\alpha_j-\beta_j\theta_i\right)^2}}{(2\pi)^{\frac{N}{2}}\sigma^N}.$$

The log-likelihood will then be

$$\ell\left(\boldsymbol{\alpha}^J, \boldsymbol{\beta}^J, \boldsymbol{\theta}^b, \sigma^2\right) = -\frac{1}{2\sigma^2}\sum_{i=1}^{b}\sum_{j=1}^{J} p_{ij}\left(Y_{ij}-\alpha_j-\beta_j\theta_i\right)^2 - \frac{N}{2}\ln\sigma^2 - \frac{N}{2}\ln(2\pi).$$

From

$$\frac{\partial \ell}{\partial \sigma^2} = \frac{1}{2\left(\sigma^2\right)^2}\sum_{i=1}^{b}\sum_{j=1}^{J} p_{ij}\left(Y_{ij}-\alpha_j-\beta_j\theta_i\right)^2 - \frac{N}{2\sigma^2},$$

we obtain

$$\hat{\sigma}^2 = \frac{1}{N}\sum_{i=1}^{b}\sum_{j=1}^{J} p_{ij}\left(Y_{ij}-\hat{\alpha}_j-\hat{\beta}_j\hat{\theta}_i\right)^2,$$

where " ^ " indicates, as usual, maximum likelihood estimators. To obtain the $(\hat{\alpha}_j, \hat{\beta}_j)$, $j = 1,..., J$, and the $\hat{\theta}_i$, $i = 1,...,b$, we must minimize

$$S\left(\boldsymbol{\alpha}^J, \boldsymbol{\beta}^J, \boldsymbol{\theta}^b\right) = \sum_{i=1}^{b}\sum_{j=1}^{J} p_{ij}\left(Y_{ij}-\alpha_j-\beta_j\theta_i\right)^2.$$

Thus, our weighted nonlinear least square estimators are also maximum likelihood estimators, once normality is assumed. Moreover, the zigzag algorithm can continue to be used.

These results extend, naturally, to polynomial regressions; thus:

- least square estimators of the coefficients and environmental indices are also maximum likelihood estimators, under assumed normality;
- moreover, with *m* the degree of the polynomial regression, we obtain

$$\hat{\sigma}_m^2 = \frac{1}{N}\sum_{i=1}^{b}\sum_{j=1}^{J} p_{ij}\left(Y_{ij} - \sum_{\ell=0}^{m}\hat{\beta}_{j\ell}\hat{\theta}_i^{\ell}\right)^2 .$$

We considered the general case, so that these results also hold for the complete block case.

4. Likelihood ratio tests

4.1. The general case

As a polynomial of degree *t* is a special case of a polynomial of degree $t+1$, we can (see Fisz 1963, p. 580) assume the degree $m \le t+1$ to test $H_0 : m \le t$ (formally by this hypothesis we mean $H_0 : \beta_{t+1} = 0$). Let ω be the parametric space associated with H_0, while the total parametric space, assuming that the degree of the regression does not exceed $t+1$, will be Ω.

With S_m the sum of squares of residuals, when degree *m* polynomial regressions are adjusted, the corresponding maximum of the log-likelihood is

$$\overline{\ell}_m = -\frac{1}{2\hat{\sigma}_m^2}S_m - \frac{N}{2}\ln(2\pi) - \frac{N}{2}\ln\hat{\sigma}_m^2 = -\frac{N}{2} - \frac{N}{2}\ln(2\pi) - \frac{N}{2}\ln\frac{S_m}{N}$$

Now, if H_0 holds, the limit distribution, when $N \to +\infty$, of

$$\lambda = -2\left(\overline{\ell}_t - \overline{\ell}_{t+1}\right) = N\ln\frac{S_t}{S_{t+1}}$$

is (see Mood 1974, p. 441) the central chi-square distribution with *J* degrees of freedom, since the last coefficients of the *J* regressions will be null.

The most important case is that for $t = 1$, in which we test for linearity.

In what follows, we restrict ourselves to linear regressions, with pairs (α_j, β_j), $j = 1,..., J$, of coefficients. We start by testing the hypothesis $H_0^{'}: \beta_1 = ... = \beta_J$, and let β be the common value of these coefficients when $H_0^{'}$ holds. We redefine ω and Ω in accordance with the hypothesis to be tested. Then $\overline{\ell}_1$ will be the unrestricted maximum of the log-likelihood; this is the maximum associated with Ω. Moreover, to obtain the maximum of the likelihood, when $H_0^{'}$ holds, we have to maximize

$$\ell\left(\boldsymbol{\alpha}^J, \beta\mathbf{1}^J, \boldsymbol{\theta}^b, \sigma^2\right) = -\frac{1}{2\sigma^2}\sum_{i=1}^{b}\sum_{j=1}^{J} p_{ij}\left(Y_{ij} - \alpha_j - \beta\theta_i\right)^2 - \frac{N}{2}\ln(2\pi) - \frac{N}{2}\ln\sigma^2 .$$

Now

$$\frac{\partial\ell}{\partial\sigma^2} = \frac{1}{2\left(\sigma^2\right)^2}\sum_{i=1}^{b}\sum_{j=}^{J} p_{ij}\left(Y_{ij} - \alpha_j - \beta\theta_i\right)^2 - \frac{N}{2\sigma^2},$$

so that

$$\hat{\sigma}_\omega^2 = \frac{1}{N}\sum_{i=1}^{b}\sum_{j=1}^{J} p_{ij}\left(Y_{ij} - \hat{\alpha}_{j\omega} - \hat{\beta}_\omega\hat{\theta}_{i\omega}\right)^2,$$

where the index ω indicates that these estimators are obtained maximizing the likelihood in ω. To obtain them we have only to minimize

$$S_\omega\left(\boldsymbol{\alpha}^J, \beta, \boldsymbol{\theta}^b\right) = \sum_{i=1}^{b}\sum_{j=1}^{J} p_{ij}\left(Y_{ij} - \alpha_j - \beta\theta_i\right)^2 .$$

To simplify our notation, we omit the index ω from the estimators. We will again use a zigzag type algorithm, in each cycle minimizing first for $\left(\boldsymbol{\alpha}^J, \beta\mathbf{1}^J\right)$ and then for $\boldsymbol{\theta}^b$. As a starting point, we take $\boldsymbol{\theta}_1^b = \hat{\boldsymbol{\theta}}^b$, that is, we use the unrestricted maximum likelihood estimators of the environmental indices. We thus obtain

$$\begin{cases} \dfrac{\partial S_\omega}{\partial\alpha_j} = -2\displaystyle\sum_{i=1}^{b} p_{ij}\left(Y_{ij} - \alpha_j - \beta\theta_i\right) \; ; \; j = 1,...,J \\ \\ \dfrac{\partial S_\omega}{\partial\beta} = -2\displaystyle\sum_{i=1}^{b}\sum_{j=1}^{J} p_{ij}\theta_i\left(Y_{ij} - \alpha_j - \beta\theta_i\right) \end{cases} .$$

Equating to zero the $\frac{\partial S_\omega}{\partial \alpha_j}$, $j=1,\dots,J$, we obtain $\beta = a_j - b_j\alpha_j\ ;\ j=1,\dots,J$ with

$$\begin{cases} a_j = \dfrac{\sum_{i=1}^{b} p_{ij}Y_{ij}}{\sum_{i=1}^{b} p_{ij}\theta_i}\ ;\ j=1,\dots,J \\ \\ b_j = \dfrac{\sum_{i=1}^{b} p_{ij}}{\sum_{i=1}^{b} p_{ij}\theta_i}\ ;\ j=1,\dots,J \end{cases},$$

thus obtaining $a_j - b_j\alpha_j = a_1 - b_1\alpha_1\ ;\ j=2,\dots,J$. We can thus write $\alpha_j = c_j\alpha_1 + d_j\ ;\ j=2,\dots,J$, where

$$c_j = \frac{b_1}{b_j}$$

and

$$d_j = \frac{(a_j - a_1)}{b_j};\ j=2,\dots,J$$

Equating $\sum_{j=1}^{J}\frac{\partial S_\omega}{\partial \alpha_j}$ to zero, we further obtain

$$-\sum_{i=1}^{b}\sum_{j=1}^{J} p_{ij}Y_{ij} + \sum_{j=1}^{J}\left(\sum_{i=1}^{b} p_{ij}\right)\alpha_j + \left(\sum_{i=1}^{b}\sum_{j=1}^{J} p_{ij}\theta_i\right)\beta = 0$$

so that, since $c_1 = 1$ and $d_1 = 0$, we arrive at

$$\beta = \frac{\sum_{i=1}^{b}\sum_{j=1}^{J} p_{ij}Y_{ij} - \sum_{j=1}^{J}\left(\sum_{i=1}^{b} p_{ij}\right)(c_j\alpha_1 + d_j)}{\sum_{i=1}^{b}\sum_{j=1}^{J} p_{ij}\theta_i}.$$

Besides this, a second expression for β may be obtained by equating $\frac{\partial S_\omega}{\partial \beta}$ to zero. This expression will be

$$\beta = \frac{\sum_{i=1}^{b}\sum_{j=1}^{J} p_{ij}\theta_i Y_{ij} - \sum_{j=1}^{J}\left(\sum_{i=1}^{b} p_{ij}\theta_i\right)\left(c_j\alpha_1 + d_j\right)}{\sum_{i=1}^{b}\sum_{j=1}^{J} p_{ij}\theta_i^2}.$$

From both expressions for β, we obtain

$$\alpha_1 = \frac{\dfrac{\sum_{i=1}^{b}\sum_{j=1}^{j} p_{ij}Y_{ij} - \sum_{j=1}^{J}\left(\sum_{i=1}^{b} p_{ij}\right)d_j}{\sum_{i=1}^{b}\sum_{j=1}^{J} p_{ij}\theta_i} - \dfrac{\sum_{i=1}^{b}\sum_{j=1}^{J} p_{ij}\theta_i Y_{ij} - \sum_{j=1}^{J}\left(\sum_{i=1}^{b} p_{ij}\theta_i\right)d_j}{\sum_{i=1}^{b}\sum_{j=1}^{J} p_{ij}\theta_i^2}}{\dfrac{\sum_{j=1}^{J}\left(\sum_{i=1}^{b} p_{ij}\right)c_j}{\sum_{i=1}^{b}\sum_{j=1}^{J} p_{ij}\theta_i} - \dfrac{\sum_{j=1}^{J}\left(\sum_{i=1}^{b} p_{ij}\theta_i\right)c_j}{\sum_{i=1}^{b}\sum_{j=1}^{J} p_{ij}\theta_i^2}}.$$

Starting from here, we obtain the $\alpha_2,\dots,\alpha_J$ and β for the first iteration. Next, we maximize with respect to the environmental indices, which is simple, since we have to minimize $\sum_{i=1}^{b} h_i\left(\theta_i\right)$, with

$$h_i\left(\theta_i\right) = \sum_{j=1}^{J} p_{ij}\left(Y_{ij} - \alpha_j - \beta\theta_i\right)^2 \; ; \; i = 1,\dots,b,$$

and it is straightforward to obtain

$$\theta_i = \frac{\sum_{j=1}^{J} p_{ij}\left(Y_{ij} - \alpha_j\right)}{\beta\sum_{j=1}^{J} p_{ij}} \; ; \; i = 1,\dots,b.$$

We can, as before, standardize these values in order to keep the range of environmental indices unchanged, and then carry out a second iteration and so on.

With S_ω° the final value of S_ω, we obtain $\hat{\sigma}_\omega^2 = \frac{S_\omega^\circ}{N}$, and the maximum of the log-likelihood in ω will be

$$\overline{\ell}_{\omega} = -\frac{N}{2} - \frac{N}{2}\ln\left(2\pi\right) - \frac{N}{2}\ln\frac{S_{\omega}^{\circ}}{N}.$$

When N is large (see again Mood et al. 1974, p. 441) the statistic $\lambda = N\ln\frac{S_{\omega}^{\circ}}{S_1}$ is distributed as a central chi-square with $J-1$ degrees of freedom, when $H_0^{'}:\beta_1 = ... = \beta_J = \beta$ holds.

Similarly, to test the hypothesis of equality of the intercepts $H_0^{''}:\alpha_1 = ... = \alpha_J = \alpha$, we have to minimize

$$S_{\omega}\left(\alpha\mathbf{1}^J,\ \boldsymbol{\beta}^J, \boldsymbol{\theta}^b\right) = \sum_{i=1}^{b}\sum_{j=1}^{J} p_{ij}\left(Y_{ij} - \alpha - \beta_j\theta_i\right)^2,$$

to obtain $\hat{\alpha}_{\omega}$, $\hat{\boldsymbol{\beta}}_{\omega}^J$ and $\hat{\boldsymbol{\theta}}_{\omega}^b$, where ω continues to represent the parameter space associated with the tested hypothesis. Again we will use the zigzag algorithm, minimizing alternately for $\left(\alpha\mathbf{1}^J,\ \boldsymbol{\beta}^J\right)$ and for $\boldsymbol{\theta}^b$, starting by taking $\boldsymbol{\theta}_1^b = \hat{\boldsymbol{\theta}}^b$.

Equating $\sum_{j=1}^{J}\frac{\partial S_{\omega}}{\partial\beta_j}$ to zero, we obtain

$$\alpha = \frac{\sum_{i=1}^{b}\sum_{j=1}^{J} p_{ij}\theta_i Y_{ij} - \sum_{i=1}^{b}\sum_{j=1}^{J} p_{ij}\theta_i^2\left(c_j\beta_1 + d_j\right)}{\sum_{i=1}^{b}\sum_{j=1}^{J} p_{ij}\theta_i},$$

where

$$c_j = \frac{\sum_{i=1}^{b} p_{i1}\theta_i^2 \sum_{i=1}^{b} p_{ij}\theta_i}{\sum_{i=1}^{b} p_{i1}\theta_i \sum_{i=1}^{b} p_{ij}\theta_i^2} \text{ and } d_j = \frac{\left(\frac{\sum_{i=1}^{b} p_{ij}Y_{ij}\theta_i}{\sum_{i=1}^{b} p_{ij}\theta_i} - \frac{\sum_{i=1}^{b} p_{i1}Y_{i1}\theta_i}{\sum_{i=1}^{b} p_{i1}\theta_i}\right)\sum_{i=1}^{b} p_{ij}\theta_i}{\sum_{i=1}^{b} p_{ij}\theta_i^2};\ j = 2,...,J\,.$$

Likewise, from $\frac{\partial S_{\omega}}{\partial\alpha} = 0$, we obtain

$$\alpha = \frac{\sum_{i=1}^{b}\sum_{j=1}^{J} p_{ij}Y_{ij} - \sum_{i=1}^{b}\sum_{j=1}^{J} p_{ij}\theta_i\left(c_j\beta_1 + d_j\right)}{\sum_{i=1}^{b}\sum_{j=1}^{J} p_{ij}},$$

therefore

$$\beta_1 = \frac{\dfrac{\sum_{i=1}^{b}\sum_{j=1}^{J} p_{ij}\theta_i Y_{ij} - \sum_{i=1}^{b}\sum_{j=1}^{J} p_{ij}\theta_i^2 d_j}{\sum_{i=1}^{b}\sum_{j=1}^{J} p_{ij}\theta_i} - \dfrac{\sum_{i=1}^{b}\sum_{j=1}^{J} p_{ij} Y_{ij} - \sum_{i=1}^{b}\sum_{j=1}^{J} p_{ij}\theta_i d_j}{\sum_{i=1}^{b}\sum_{j=1}^{J} p_{ij}}}{\dfrac{\sum_{i=1}^{b}\sum_{j=1}^{J} p_{ij}\theta_i^2 c_j}{\sum_{i=1}^{b}\sum_{j=1}^{J} p_{ij}\theta_i} - \dfrac{\sum_{i=1}^{b}\sum_{j=1}^{J} p_{ij}\theta_i c_j}{\sum_{i=1}^{b}\sum_{j=1}^{J} p_{ij}}}.$$

It is now straightforward to obtain α and the $\beta_2,...,\beta_J$ for the first iteration. Now to maximize for the environmental indices we have to minimize

$$h_i(\theta_i) = \sum_{j=1}^{J} p_{ij}\left(Y_{ij} - \alpha - \beta_j\theta_i\right)^2 \;;\; i = 1,...,b,$$

giving

$$\theta_i = \frac{\sum_{j=1}^{J} p_{ij}\left(Y_{ij} - \alpha\right)\beta_j}{\sum_{j=1}^{J} p_{ij}\beta_j^2} \;;\; i = 1,...,b.$$

Next, we can standardize, as before, and repeat the iteration. We repeat the iteration until the decrease in the value of S_ω is sufficiently small. We can proceed as for the hypothesis $H_0^{'}: \beta_1 = ... = \beta_J$.

Although of lesser interest, we can also consider the hypothesis

$$H_0^{'''}: \begin{cases} \alpha_1 = ... = \alpha_J = \alpha \\ \beta_1 = ... = \beta_J = \beta \end{cases}$$

of "equal" regressions. The unrestricted parameter space Ω and the corresponding log-likelihood maximum will be the same as for the previous tests.

Now, to obtain $\hat{\alpha}_\omega$, $\hat{\beta}_\omega$ and $\hat{\boldsymbol{\theta}}_\omega^b$, we again use the zigzag algorithm to minimize

$$S_\omega\left(\alpha\mathbf{1}^J, \beta\mathbf{1}^J, \boldsymbol{\theta}^b\right) = \sum_{i=1}^{b}\sum_{j=1}^{J} p_{ij}\left(Y_{ij} - \alpha - \beta\theta_i\right)^2$$

alternately with respect to $\left(\alpha\mathbf{1}^J, \beta\mathbf{1}^J\right)$ and to $\boldsymbol{\theta}^b$. Just as before, we take $\boldsymbol{\theta}_1^b = \hat{\boldsymbol{\theta}}^b$ and, to simplify the notation, we omit the index ω.

We now have

$$\begin{cases} \dfrac{\partial S_\omega}{\partial \alpha} = -2\sum_{i=1}^{b} p_{ij}\left(Y_{ij} - \alpha - \beta\theta_i\right) \\ \\ \dfrac{\partial S_\omega}{\partial \beta} = -2\sum_{i=1}^{b}\sum_{j=1}^{J} p_{ij}\theta_i\left(Y_{ij} - \alpha - \beta\theta_i\right) \end{cases}$$

and thus, equating these derivatives to zero, we obtain the solutions of this system:

$$\begin{cases} \alpha = \dfrac{\sum_{i=1}^{b}\sum_{j=1}^{J} p_{ij}\theta_i^2 \sum_{i=1}^{b}\sum_{j=1}^{J} p_{ij}Y_{ij} - \sum_{i=1}^{b}\sum_{j=1}^{J} p_{ij}\theta_i \sum_{i=1}^{b}\sum_{j=1}^{J} p_{ij}\theta_i Y_{ij}}{\sum_{i=1}^{b}\sum_{j=1}^{J} p_{ij} \sum_{i=1}^{b}\sum_{j=1}^{J} p_{ij}\theta_i^2 - \left(\sum_{i=1}^{b}\sum_{j=1}^{J} p_{ij}\theta_i\right)^2} \\ \\ \beta = \dfrac{\sum_{i=1}^{b}\sum_{j=1}^{J} p_{ij} \sum_{i=1}^{b}\sum_{j=1}^{J} p_{ij}\theta_i Y_{ij} - \sum_{i=1}^{b}\sum_{j=1}^{J} p_{ij}\theta_i \sum_{i=1}^{b}\sum_{j=1}^{J} p_{ij}Y_{ij}}{\sum_{i=1}^{b}\sum_{j=1}^{J} p_{ij} \sum_{i=1}^{b}\sum_{j=1}^{J} p_{ij}\theta_i^2 - \left(\sum_{i=1}^{b}\sum_{j=1}^{J} p_{ij}\theta_i\right)^2} \end{cases}.$$

In the second part of the iteration we have to minimize $\sum_{i=1}^{b} h_i\left(\theta_i\right)$, with

$$h_i\left(\theta_i\right) = \sum_{j=1}^{J} p_{ij}\left(Y_{ij} - \alpha - \beta\theta_i\right)^2 \ ; \ i = 1,\ldots,b,$$

and obtain

$$\theta_i = \frac{\sum_{j=1}^{J} p_{ij}\left(Y_{ij} - \alpha\right)}{\beta\sum_{j=1}^{J} p_{ij}} \ ; \ i = 1,\ldots,b.$$

As before, we standardize the estimated environmental indices and iterate until S_ω stabilizes. With S_ω° the final value of S_ω, if *N* is large and $H_0^{'''}$ holds, then the statistic $\lambda = N \ln \frac{S_\omega^\circ}{S_1}$ follows a central chi-square distribution, with $2(J-1)$ degrees of freedom.

We point out that the iteration procedure always converges, because we obtain a decreasing sequence of minima, bounded below by zero.

4.2. The complete case

In this case $N = bJ$, since all cultivars are present in all blocks, and $p_{ij} = 1$, $j = 1,..., J$, $i = 1,...,b$, which simplifies considerably the previous formulae. The adaptation of the tests for the hypothesis concerning the degree of the regressions is straightforward. Again, the most important test is the one for $t = 1$, that is, the test of linearity.

Let us consider the tests for the same three hypotheses as before, concerning the coefficients of linear regressions. We begin with

$$H_0 : \beta_1 = ... = \beta_J = \beta$$

where we must minimize

$$S_\omega\left(\boldsymbol{\alpha}^J,\ \beta\mathbf{1}^J, \boldsymbol{\theta}^b\right) = \sum_{i=1}^{b}\sum_{j=1}^{J}\left(Y_{ij} - \alpha_j - \beta\theta_i\right)^2 .$$

To do this we can proceed as before, except that now we obtain

$$\alpha_j = \alpha_1 + \frac{\sum_{i=1}^{b}\left(Y_{ij} - Y_{i1}\right)}{b} \ ;\ j = 2,..., J$$

In this case, with $T_i = \sum_{j=1}^{J} Y_{ij}$, $i = 1,...,b$, we have

$$\alpha_1 = \frac{\dfrac{\sum_{i=1}^{b} T_i - b\sum_{j=1}^{J} d_j}{J\sum_{i=1}^{b}\theta_i} - \dfrac{\sum_{i=1}^{b}\theta_i T_i - \sum_{i=1}^{b}\theta_i \sum_{j=1}^{J} d_j}{J\sum_{i=1}^{b} p_i\theta_i^2}}{\dfrac{b}{\sum_{i=1}^{b}\theta_i} - \dfrac{\sum_{i=1}^{b}\theta_i}{\sum_{i=1}^{b}\theta_i^2}}$$

so that the calculations are much simpler than in the general case. The second part of the iteration is also considerably simplified, since we can take

$$\theta_i = \frac{\sum_{j=1}^{J}\left(Y_{ij} - \alpha_j\right)}{J\beta} \; ; \; i = 1,\dots,b .$$

Next, to test

$$H_0^{''} : \alpha_1 = \dots = \alpha_J = \alpha$$

we must minimize

$$S_{\omega}\left(\alpha \mathbf{1}^J, \boldsymbol{\beta}^J, \boldsymbol{\theta}^b\right) = \sum_{i=1}^{b}\sum_{j=1}^{J}\left(Y_{ij} - \alpha - \beta_j\theta_i\right)^2 .$$

Now we obtain

$$c_j = 1; \; j = 1, 2, \dots, J \quad \text{and} \quad d_j = \frac{\sum_{i=1}^{b}\theta_i\left(Y_{ij} - Y_{i1}\right)}{\sum_{i=1}^{b}\theta_i^2}; \; j = 2, \dots, J$$

as well as

$$\beta_1 = \frac{\dfrac{\sum_{i=1}^{b}\theta_i T_i - \sum_{i=1}^{b}\theta_i^2\sum_{j=1}^{J} d_j}{J\sum_{i=1}^{b}\theta_i} - \dfrac{\sum_{i=1}^{b} T_i - \sum_{i=1}^{b}\theta_i\sum_{j=1}^{J} d_j}{Jb}}{\dfrac{\sum_{i=1}^{b}\theta_i^2}{\sum_{i=1}^{b}\theta_i} - \dfrac{\sum_{i=1}^{b}\theta_i}{b}} ,$$

so that, as for the previous test, the first part of the iteration is much easier to carry out than in the general case. This is also true for the second part of the iteration, since now we obtain

$$\theta_i = \frac{\sum_{j=1}^{J}\left(Y_{ij} - \alpha\right)\beta_j}{\sum_{j=1}^{J}\beta_j^2} \; ; \; i = 1,\dots,b .$$

Finally, we have the test for

$$H_0^{'''}:\begin{cases}\alpha_1=\ldots=\alpha_J=\alpha\\ \\ \beta_1=\ldots=\beta_J=\beta\end{cases}.$$

Then we have to minimize

$$S_\omega\left(\alpha\mathbf{1}^J,\beta\mathbf{1}^J,\boldsymbol{\theta}^b\right)=\sum_{i=1}^{b}\sum_{j=1}^{J}\left(Y_{ij}-\alpha-\beta\theta_i\right)^2 .$$

Proceeding as before, we obtain in this case

$$\begin{cases}\alpha=\dfrac{\sum_{i=1}^{b}\theta_i^2\sum_{i=1}^{b}T_i-\sum_{i=1}^{b}\theta_i\sum_{i=1}^{b}\theta_iT_i}{J\left[b\sum_{i=1}^{b}\theta_i^2-\left(\sum_{i=1}^{b}\theta_i\right)^2\right]}\\ \\ \beta=\dfrac{b\sum_{i=1}^{b}\theta_iT_i-\sum_{i=1}^{b}\theta_i\sum_{i=1}^{b}T_i}{J\left[b\sum_{i=1}^{b}\theta_i^2-\left(\sum_{i=1}^{b}\theta_i\right)^2\right]}\end{cases},$$

which simplifies considerably the first part of the iteration. In the second part, it is straightforward to obtain

$$\theta_i=\frac{1}{J\beta}\sum_{j=1}^{J}\left(Y_{ij}-\alpha\right);\ i=1,\ldots,b .$$

5. Example

In this application, we use data obtained in 17 experiments in α-designs carried out by the Research Center for Cultivar Testing at Słupia Wielka (Poland) in the years 1997 and 1998.

In these experiments, carried out in resolvable block designs, cultivars of winter rye were compared. In each design there were 4 superblocks, each with 5 blocks of 4 plots. Each cultivar occurred on one plot per superblock. The 20 cultivars are listed in Table 1.

Table 1. Names of the cultivars studied

Names of cultivars			
01AMILO	1WARKO	1MARDER	ADAR
02ZDUNO	1SMH 1094	1SMH 1195	1SMH 1295
03NAD 195	ESPRIT	RAH 496	CHD 296
04 CHD 396	1RAH 596	WID 196	RAH 697
05RAPID	RAH 797	RAH 897	URSUS

The locations and years of the different experiments are listed in Table 2.

Table 2. Locations and years in which experiments took place

Trial	Location / Experimental station	Year
1	Lubinicko	1998
2	Pokój	1997/98
3	Dukla	1998
4	Uhnin	1998
5	Ruska wieś	1997
6	Łopuszna	1998
7	Kawęczyn	1998
8	Rychliki	1998
9	Głodowo	1998
10	Rarwino	1998
11	Masłowice	1998
12	Lubliniec Nowy	1998
13	Krościna Mala	1998
14	Seroczyn	1998
15	Cicibór	1998
16	Kochcice	1998
17	Sulejów	1998

To adjust the linear regressions we used the zigzag algorithm. The adjusted regressions are presented in Figure 1. The final results are presented in Table 3.

In comparing the cultivars that integrate the upper contour (dominants) with the remaining ones, we begin by using one-sided *t* tests. The results are given in Table 4.

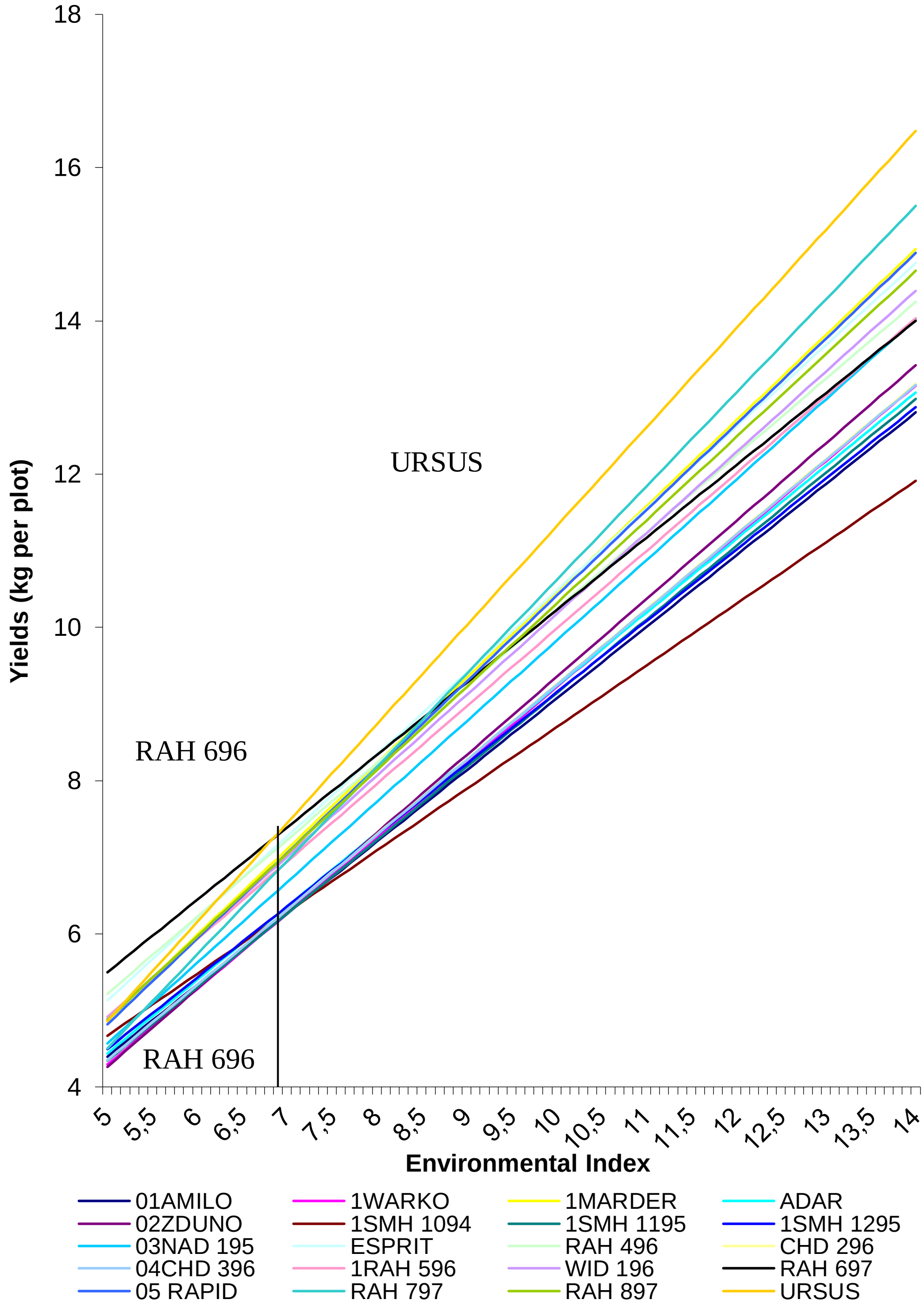

Figure 1. Adjusted regressions, using L_2 environmental indices for the 17 experiments

Table 3. Adjusted coefficients and R^2

Cultivar	$\tilde{\alpha}$	$\tilde{\beta}$	R^2
URSUS	-1.59	1.29	0.96
RAH 797	-1.60	1.22	0.97
05RAPID	-0.78	1.12	0.97
1MARDER	-0.73	1.12	0.94
RAH 897	-0.55	1.09	0.95
ESPRIT	-0.22	1.07	0.92
WID 196	-0.38	1.06	0.96
03NAD 195	-0.68	1.05	0.93
02ZDUNO	-0.82	1.02	0.97
1RAH 596	-0.15	1.01	0.95
RAH 496	0.20	1.00	0.95
1WARKO	-0.63	0.99	0.96
CHD 296	-0.55	0.98	0.93
04CHD 396	-0.54	0.98	0.95
1SMH 1195	-0.45	0.96	0.93
ADAR	-0.35	0.96	0.96
RAH 697	0.77	0.95	0.91
01AMILO	-0.27	0.93	0.93
1SMH 1295	-0.16	0.93	0.96
1SMH 1094	0.65	0.80	0.90

Table 4. Dominant and dominated cultivars

Dominant cultivars	Range of dominance	Dominated cultivars at the 5% level
RAH 697	$\left[5.42\ ;\ 6.84\right]$	RAH 797, 05RAPID, 1MARDER, RAH 897, WID 196, 03NAD 195, 02ZDUNO, 1RAH 596, 1WARKO, CHD 296, 04CHD 396, 1SMH 1195, ADAR, 01AMILO, 1SMH 1295, 1SMH 1094
URSUS	$\left[6.84\ ;\ 13.47\right]$	RAH 797, 05RAPID, 1MARDER, RAH 897, WID 196, 03NAD 195, 02ZDUNO, 1RAH 596, 1WARKO, CHD 296, 04CHD 396, 1SMH 1195, ADAR, 01AMILO, 1SMH 1295, 1SMH 1094

We can see easily that besides the dominant cultivars (URSUS and RAH 697) only the cultivars ESPRIT and RAH 496 are not significantly dominated at the 5% level, in the range $\left[5.42\ ;\ 13.47\right]$. If we work at the 1% level, we will also have to consider the cultivar 1MARDER as non-significantly dominated. Thus, we obtain the efficiency ratios

$$r_1 = \frac{\text{Number of dominant cultivars}}{\text{Number of cultivars}} = 0.1$$

and $$r_2 = \frac{\text{Number of non dominant cultivars}}{\text{Number of cultivars}} = 0.20$$

(0.25 if we work at the 1% level).

If we wanted to use more robust methods, we could use the Scheffé and Bonferroni multiple comparison methods. In this case we have the results given in Table 5.

Table 5. Significantly dominated cultivars, using the Scheffé and Bonferroni multiple comparison methods

	Dominant cultivars	RAH 697	URSUS
	Range of dominance	$[5.42\ ;\ 6.84]$	$[6.84\ ;\ 13.47]$
Method		Dominated cultivars at the 5% level	
Scheffé		01AMILO, 1SMH 1295, 1SMH 1094 03NAD 195, 02ZDUNO, 1WARKO, CHD 296, 04CHD 396, 1SMH 1195, ADAR	03NAD 195, 02ZDUNO, 1WARKO, CHD 296, 04CHD 396, 1SMH 1195, ADAR, 01AMILO, 1SMH 1295, 1SMH 1094
Bonferroni		01AMILO, 1SMH 1295, 1SMH 1094 RAH 797, WID 196, 03NAD 195, 02ZDUNO, 1RAH 596, 1WARKO, CHD 296, 04CHD 396, 1SMH 1195, ADAR	RAH 797, WID 196, 03NAD 195, 02ZDUNO, 1RAH 596, 1WARKO, CHD 296, 04CHD 396, 1SMH 1195, ADAR, 01AMILO, 1SMH 1295, 1SMH 1094

While the efficiency ratio r_1 is still 0.1, r_2 increases considerably. This would point to high performance, being required for a new cultivar to be recommended, and robust decisions when downgrading previously recommended cultivars; see Pereira & Mexia (2003). This may be quite acceptable, due to the costs involved in obtaining cultivars suitable for use.

6. Conclusion

The use of L_2 environmental indices led to a theoretical framework for JRA, integrating it into the statistical inference for normal models. Thus, it was possible to:

- obtain maximum likelihood estimators for the environmental indices and the coefficients of the regressions;
- perform likelihood ratio tests, namely linearity tests and tests on the regression coefficients.

Acknowledgements

The Research Centre for Cultivar Testing (Słupia Wielka, Poland) is thanked for supplying the data used in this paper.
This work was partially supported by the FCT-Fundação para a Ciência e a Tecnologia, Portugal, under project UIDB/04674/2020 (CIMA).

REFERENCES

Aastveit A.H. Mejza S. (1992): A selected bibliography on statistical methods for the analysis of genotype × environment interaction. Biuletyn Oceny Odmian 24-25: 83–97.

Baeta J., Gusmão L., Mexia J.T., Costa-Rodrigues L. (1990): Análise de variância de ensaios comparativos de cultivares de triticale, delineados em Blocos Completos Causalizados. I. Identificação dos factores de restrição. Melhoramento 32: 189–197.

Bazarra M.S., Sherali H.D., Shetty C.M. (1992): Nonlinear programming, Theory and Algorithms. 2^{nd} ed., John Wiley and Sons, New York.

Becker H.C., Leon J. (1988): Stability analysis in plant breeding. Plant Breeding 101: 1–23.

Digby P.G.N. (1979): Modified joint regression analysis for incomplete variety × environment data. J. Agric. Sci., Camb. 93: 81–86.

Eberlhart S.A., Rusell W.A. (1966): Stability parameters for comparing varieties. Crop. Sci. 6: 36–40.

Finlay K.W., Wilkinson G.N. (1963): The analysis of adaptation in a plant-breeding programme. Aust. J. Agric. Res. 14: 742–754.

Fisz M. (1963): Probability theory and mathematical statistics. 3^{rd} ed. John Wiley and Sons, New York.

Gusmão L. (1985): An adequate design for regression analysis of yield trials. Theor. Appl. Genet. 71: 314–319.

Gusmão L. (1986a); Inadequacy of blocking in cultivar yield trials. Theor. Appl. Genet. 72: 98–104.

Gusmão L. (1986b): A interacção genótipo × ambiente e a comparação de cultivares de cereais. PhD thesis. Instituto Superior de Agronomia. Universidade Técnica de Lisboa, Lisboa.

Gusmão L. (1988): Assessing small grain cultivars for yields in variable environments through Joint Regression Analysis. Rachis 7: 22–25.

Gusmão L., Mexia J.T., Gomes M.L. (1989): Mapping of equipotential zones for cultivar yield pattern evolution. Plant Breeding 103: 293–298.

Gusmão L. (1990): Avaliação de cultivares (uma perspectiva experimental diferente). Revista de Ciências Agrárias - nº2, vol. 13: 11–18.

Kang M.S., Gauch H.G. (1996): Genotype by Environmental Interaction. CRC Press, New York.

Lin C.S., Binns M.R., Lefkovitch L.P. (1986): Stability analysis: Where Do We Stand? Crop. Sci. 26: 894–900.

Mexia J.T., Gusmão L., Ferreira M.T., Baeta J. (1990): Homogeneidade numa zona equipotencial de adaptação e suas implicações no delineamento de ensaios de adaptação. Garcia de Orta, Série de Estudos Agronómicos 16: 65–70.

Mexia J.T., Gusmão L., Baeta J. (1991): Analysis of cultivar yield trials designed in Randomize Complete Blocks. Revista de Ciências Agrárias 14: 59–63.

Mexia J.T., Amaro A.P., Gusmão L., Baeta J. (1997): Upper contour of a Joint Regression Analysis, J. Genet. & Breed. 51: 253–255.

Mexia J.T., Pereira D.G., Baeta J. (1999): L_2 environmental indexes. Biometrical Letters 36: 137–143.

Mexia J.T., Pereira D.G., Baeta J. (2001): Weighted linear joint regression analysis. Biometrical Letters 38: 33–40.

Mood A.M., Graybill F.A., Boes D.C. (1974): Introduction to the theory of statistics. McGraw-Hill, Singapore, 3rd ed.

Mooers C.A. (1921): The agronomic placement of varieties. J. Amer. Soc. Agron. 13: 337–352.

Ng M.P., Grunwald G.K. (1997): Nonlinear regression analysis of the joint-regression model. Biometrics 53: 1366–1372.

Ng M.P., Williams E.R. (2001): Joint-regression analysis for incomplete two-way tables. Aust. N. Z. J. Stat. 43(2): 201–206.

Patterson H.D., Williams E.R. (1976): A new class of resolvable incomplete block designs. Biometrika 63: 83–92

Pereira D.G., Mexia J.T. (2002): Multiple comparison in Joint Regression Analysis with special reference to variety selection. Scientific papers of the Agricultural University of Poznan, Agriculture 3: 67–74.

Pereira D.G., Mexia J.T. (2003): The use of Joint Regression Analysis in selecting recommended cultivars. Biuletyn Oceny Odmian (Cultivar Testing Bulletin) 31: 19–25.

Severine T.A. (2000): Likelihood methods in statistics. Oxford University Press, New York.

Westcott B. (1986): Some methods of analysing genotype-environment interaction. Heredity, 56: 243–253.

Yates F., Cochran W.G. (1938): The analysis of groups experiments. J. Agric. Sci., Cambridge, 28: 556–580.

APPENDIX

The goal function $S\left(\boldsymbol{\alpha}^J, \boldsymbol{\beta}^J, \mathbf{x}^b\right)$ has the derivatives

$$\frac{\partial^2 S}{\partial x_i \partial \beta_j} = -2p_{ij}\left(Y_{ij} - \alpha_j - \beta_j x_i\right) + 2p_{ij}\beta_j x_i \,,$$

thus when $Y_{ij} - \alpha_j - \beta_j x_i \approx 0,\ i = 1,\ldots,b,\ j = 1,\ldots,J$ we have

$$\frac{\partial^2 S}{\partial x_i \partial \beta_j} \approx 2p_{ij}\beta_j x_i,\ i = 1,\ldots,b,\ j = 1,\ldots,J$$

and we may approach the Hessian matrix of S by the matrix

$$W = 2\left[\begin{array}{ccc|cccccc}
\sum_{j=1}^{J} p_{1j}\tilde{\beta}_j^2 & \cdots & 0 & p_{11}\tilde{\beta}_1 & \cdots & p_{1J}\tilde{\beta}_J & p_{11}\tilde{\beta}_1\tilde{x}_1 & \cdots & p_{1J}\tilde{\beta}_J\tilde{x}_1 \\
\vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
0 & \cdots & \sum_{j=1}^{J} p_{bj}\tilde{\beta}_j^2 & p_{b1}\tilde{\beta}_1 & \cdots & p_{bJ}\tilde{\beta}_J & p_{b1}\tilde{\beta}_1\tilde{x}_b & \cdots & p_{bJ}\tilde{\beta}_J\tilde{x}_b \\
\hline
p_{11}\tilde{\beta}_1 & \cdots & p_{b1}\tilde{\beta}_1 & \sum_{i=1}^{b} p_{i1} & \cdots & 0 & \sum_{i=1}^{b} p_{i1}\tilde{x}_i & \cdots & 0 \\
\vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
p_{1J}\tilde{\beta}_J & \cdots & p_{bJ}\tilde{\beta}_J & 0 & \cdots & \sum_{i=1}^{b} p_{iJ} & 0 & \cdots & \sum_{i=1}^{b} p_{iJ}\tilde{x}_i \\
p_{11}\tilde{\beta}_1\tilde{x}_1 & \cdots & p_{b1}\tilde{\beta}_1\tilde{x}_b & \sum_{i=1}^{b} p_{i1}\tilde{x}_i & \cdots & 0 & \sum_{i=1}^{b} p_{i1}\tilde{x}_i^2 & \cdots & 0 \\
\vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
p_{1J}\tilde{\beta}_J\tilde{x}_1 & \cdots & p_{bJ}\tilde{\beta}_J\tilde{x}_b & 0 & \cdots & \sum_{i=1}^{b} p_{iJ}\tilde{x}_i & 0 & \cdots & \sum_{i=1}^{b} p_{iJ}\tilde{x}_i^2
\end{array}\right]$$

Let us now establish

Proposition

The matrix $\mathbf{W}$ is positive definite.

Proof: With $\mathbf{m}^{b+2JT} = \left[\mathbf{u}^{bT}, \quad \mathbf{v}^{JT}, \quad \mathbf{z}^{JT}\right]$ we obtain

$$\mathbf{m}^{b+2JT}\ \mathbf{W}\ \mathbf{m}^{b+2J} = 2\left[\sum_{i=1}^{b}\left(\sum_{j=1}^{J} p_{ij}\tilde{\beta}_j^2\right)u_i^2 + 2\sum_{i=1}^{b}\sum_{j=1}^{J} p_{ij}\tilde{\beta}_j u_i v_j + 2\sum_{i=1}^{b}\sum_{j=1}^{J} p_{ij}\tilde{\beta}_j \tilde{x}_i u_i z_j + \sum_{j=1}^{J}\left(\sum_{i=1}^{b} p_{ij}\right)v_j^2\right]$$

$$+2\left[2\sum_{j=1}^{J}\left(\sum_{i=1}^{b} p_{ij}\tilde{x}_i\right)v_j z_j + \sum_{j=1}^{J}\left(\sum_{i=1}^{b} p_{ij}\tilde{x}_i^2\right)z_j^2\right]$$

$$= 2\sum_{i=1}^{b}\sum_{j=1}^{J} p_{ij}\left(\tilde{\beta}_j^2 u_i^2 + 2\tilde{\beta}_j u_i\left(v_j + z_j\tilde{x}_i\right) + v_j^2 + 2v_j z_j \tilde{x}_i + z_j^2\tilde{x}_i^2\right)$$

$$= 2\sum_{i=1}^{b}\sum_{j=1}^{J} p_{ij}\left[\tilde{\beta}_j u_i + \left(v_j + z_j\tilde{x}_i\right)\right]^2$$ which establishes the thesis. ❑

Thus *S* will then behave, at least approximately, as a convex function. Such functions do not have local minima distinct from their absolute minimum, which is unique (see Bazaraa et al. 1992, p. 113).

If we obtain small enough residuals, we may assume that the convex approximation to *S* holds and accept our estimates as corresponding, at least with good approximation, to the absolute minimum of *S*.

For this reason, good behavior of the zigzag algorithm may be expected.