Biometrical Letters
Vol. 45 (2008), No. 2, 9-22

Analysis of a with-covariates model for the Environmental Indexes

Iola Pinto[1], Célia Nunes[2], João Tiago Mexia[3]

[1]Instituto Superior de Engenharia de Lisboa, Rua Conselheiro Emdio Navarro 1,
1959-007 Lisboa, Portugal: e-mail: ipinto@deetc.isel.ipl.pt
[2]Universidade da Beira Interior, Covilhã, Portugal
[3]Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, Monte da Caparica,
Portugal

Summary

In the context of the use of Joint Regression Analysis (JRA) for cultivars comparison, as in Pinto (2006), the environmental index is a variable which measures the productivity per block in field experiments. Initially, randomized blocks were used, and their mean yields (classical environmental indexes) were taken as measuring their environmental indexes. Later on, following Patterson and Williams (1976), α - designs, which have incomplete blocks, superseded randomized blocks. Then the environmental indexes for each block, couldn't be measured by the respective mean yields, which would lead to biased estimates. This problem was solved by the introduction of L_2 environmental indexes see Mexia et al. (1999). In that paper the authors, through the application of a Zig-Zag algorithm, showed how to adjust simultaneously the regression coefficients and the environmental indexes. This study presents an analysis of a with-covariates model for the Environmental Indexes (see Searle, 1987) using the adjusted L_2 environmental indexes as observations of the dependent variable. Concerning the explicative variables, we consider a general mean, the main effects of two qualitative factors, year and location and as covariates the "classical environmental indexes". The presence of multi-collinearity led to the use of principal components, see Judge et al. (1988), to perform the adjustment. Our results are applied to the data of a Wheat Plant Breeding Program in Portugal (1986-2000), kindly supplied by the Portuguese Plant Breeding Station. The sole significant result was for the covariate, which validates the linearity assumed in JRA.

Key words: linear models, covariates, dummy variables, multicollinearity, principal components, L_2 environmental indexes, JRA.

1. Introduction

This study is a follow-up of an analysis using JRA of the Portuguese Wheat Plant Breeding Program (1986-2000) data.

JRA is (see Aastveit and Mejza, 1992) a technique for the analysis of genotype × environment interaction. It has been widely used for cultivar comparison and selection. For each cultivar a linear regression of yields on environmental index is adjusted. This controlled variable measures the productivity of the pairs (location, year).

For some years after its invention JRA was applied to networks of experiments designed as randomized blocks. Following Gusmão (1985) and (1986a) the environmental indexes for the different blocks were measured by the respective mean yields (classical indexes). Later on, the field experiments were mainly of α type designs, thus with incomplete blocks see Patterson and Williams (1976). Then, the environmental index for each block could not be measured by the respective mean yield. This problem was solved by the introduction of L_2 environmental indexes see Mexia et al. (1999). The application of JRA with L_2 environmental indexes consists in the joint adjustment of two vectors: the vector of regression coefficients and that of environmental indexes. To perform the adjustment, we minimize the quadratic goal function

$$S(\boldsymbol{\alpha}_C, \boldsymbol{\beta}_C, \boldsymbol{x}_b) \sum_{j=1}^{C} [\sum_{i=1}^{b} p_{ij}(y_{ij} - \alpha_j - \beta_j x_i)^2], \tag{1}$$

where (α_j, β_j) are the intercept and the slope for the regression line to be adjusted for cultivar j, with, $j = 1, ..., C$, $\boldsymbol{x}_b = [x_1, x_2, ..., x_b]'$ is the environmental indexes vector, p_{ij} are the weights of the j^{th} cultivar in the i^{th} block, assuming the values 1 [0] when the j^{th} cultivar is present [absent] in the i^{th} block and y_{ij} is the yield of the j^{th} cultivar in the i^{th} block if $p_{ij} = 1$ and any value if $p_{ij} = 0$.

We point out that, in a breeding program in which in every year all cultivars used are present in all locations, the blocks are yearly complete, which enables us to treat their block average yields as "classical indexes".

Since L_2 environmental indexes measure productivity, it is worthwhile to model them in order to assess the statistical significance of years, locations and classical indexes. So, after adjusting the goal function given by (1), an analysis of a with-covariates model (see Searle, 1987 and Scheffé,

1959) is performed, with the adjusted L_2 environmental indexes as dependent variable. In our case it is reasonable to consider the "classical indexes" as covariates, since the observations of the dependent variable arise from the adjustment performed by the Zig-Zag algorithm (see Mexia et al., 1999), which uses the "classical indexes" as a starting point. Namely, we wanted to see if, beyond the impact of the covariates, there were systematic effects of years and locations.

Let v_{ijk} be the k^{th} observation of the response variable, $k = 1, ..., K$, the adjusted environmental index, for the i^{th} level of the first factor (year) and j^{th} level of the second factor (location). The model equation is given by

$$v_{ijk} = \mu + \alpha_i + \beta_j + \gamma z_{ijk} + e_{ijk}, \tag{2}$$

where μ is the general mean, α_i, $i = 1, ...I$, is the effect of the i^{th} level of the first factor, β_j, $j = 1, ..., J$, is the effect of the j^{th} level of the second factor, z_{ijk} is the k^{th} covariate observation corresponding to the i^{th} year and j^{th} location, γ is the sole slope and e_{ijk} is the k^{th} random residual error variable corresponding to the i^{th} year and j^{th} location.

We will apply our results to the Portuguese Wheat Plant Breeding Program (1986-2000) data, kindly supplied by the Portuguese Plant Breeding Station.

2. Model

We now briefly present the statistical methods used.

2.1. The with-covariates model

Assuming there are b adjusted environmental indexes and r factor effects, in matrix notation the previous model may be written as

$$\boldsymbol{V}_b = \boldsymbol{U}\boldsymbol{\eta}_r + \boldsymbol{Z}\gamma + \boldsymbol{e}_b, \tag{3}$$

where

- $\boldsymbol{V}_b$ is the vector of the adjusted L_2 environmental indexes;
- $\boldsymbol{U}$ is an incidence matrix for the factor effects which will be the components of the vector $\boldsymbol{\eta}_r$;

- ○ $\boldsymbol{Z}$ is the vector of classical indexes, thus corresponding to the covariates, and γ is the slope;
- ○ $\boldsymbol{e}_b$ is the errors vector.

Following Pinto (2006), it is assumed that, $\boldsymbol{e}_b \sim N(\boldsymbol{0}_b, \sigma^2 I_b)$, i.e. that the error vector is normal with null mean vector and covariance matrix $\sigma^2 I_b$.

The matrix $\boldsymbol{U}$ has r columns, thus can be written

$$\boldsymbol{U} = [\boldsymbol{u}_1, ..., \boldsymbol{u}_r], \tag{4}$$

where $\boldsymbol{u}_1 = \boldsymbol{1}_b$, since it corresponds to the first component of the vector of parameters to be estimated, the general mean value $\eta_1 = \mu$. Next, we would have the column vectors $\boldsymbol{u}_2, ..., \boldsymbol{u}_{I+1}$ associated to the I years. Since every observation of the dependent variable corresponds to one and only one pair (location, year), we would have

$$\sum_{j=2}^{I+1} \boldsymbol{u}_j = \boldsymbol{1}_b = \boldsymbol{u}_1, \tag{5}$$

which led to the existence of multi-collinearity.

We point out that the coefficients corresponding to the year effects $\boldsymbol{\eta}_2, ..., \boldsymbol{\eta}_{I+1}$ have null sum, and so

$$\eta_{I+1} = -\sum_{j=2}^{I} \eta_j, \tag{6}$$

which enables us to overcome the observed multi-collinearity. To make clear the mechanism used, let the l^{th} adjusted environmental index correspond to the last year, thus we would have

$$u_{l,2} = \cdots = u_{l,I} = 0; \qquad\qquad u_{l,I+1} = 1, \tag{7}$$

so that

$$\sum_{j=2}^{I+1} \eta_j u_{l,j} = \eta_{I+1} u_{l,I+1}, \tag{8}$$

and, by (6)

$$\eta_{I+1} u_{l,I+1} = -\sum_{j=2}^{I} \eta_j u_{l,I+1}. \tag{9}$$

We can reproduce this result eliminating the $(I+1)^{th}$ column from the matrix $\boldsymbol{U}$ and in the l^{th} row putting

$$u_{l,2} = \cdots = u_{l,I} = -1 \tag{10}$$

whenever initially we had $u_{l,I+1} = 1$.

Likewise, we may discard the last column which would be associated with the last location and, whenever initially $u_{l,I+J+1} = 1$, taking

$$u_{l,1+(I-1)+1} = \cdots = u_{l,1+(I-1)+(J-1)} = -1. \tag{11}$$

Proceeding in this way, we reduce the multi-collinearity problems and the number of parameters to be estimated from $r = I + J + 1$ to $r = I + J - 1$.

The column matrix $\boldsymbol{Z}$ contains the "classical environmental indexes", while γ is the slope.

From now on, we will consider a unique model matrix $\boldsymbol{X}$, constituted by the columns of matrices $\boldsymbol{U}$ and $\boldsymbol{Z}$, whereas the coefficients vector $\boldsymbol{\delta}_h = [\boldsymbol{\eta}_r{}', \boldsymbol{\gamma}]'$, with $h = r + 1$. We can rewrite the model as

$$\boldsymbol{V}_b = \boldsymbol{X}\boldsymbol{\delta}_h + \boldsymbol{e}_b. \tag{12}$$

Even with the previous method of building the incidence matrix $\boldsymbol{U}$, in the application we considered we encountered multi-collinearity problems. To carry out the adjustment, we used principal components derived (see Judge et al., 1988, p. 865) from spectral analysis of matrix $\boldsymbol{X}'\boldsymbol{X}$. Thus, with $\lambda_1, ..., \lambda_g$, the non-zero, $g < h$, eigenvalues of matrix $\boldsymbol{X}'\boldsymbol{X}$, and $\boldsymbol{\alpha}_{1,h}, ..., \boldsymbol{\alpha}_{g,h}$ the corresponding eigenvectors, we will replace the model matrix, $\boldsymbol{X}$ by

$$\dot{\boldsymbol{X}} = \boldsymbol{X}\boldsymbol{K} = \begin{bmatrix} \dot{x}_{1,1} & \cdots & \dot{x}_{1,j} & \cdots & \dot{x}_{1,g} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ \dot{x}_{b,1} & \cdots & \dot{x}_{b,j} & \cdots & \dot{x}_{b,g} \end{bmatrix} = \begin{bmatrix} \dot{\boldsymbol{x}}_{1,b}, ..., \dot{\boldsymbol{x}}_{g,b} \end{bmatrix} \tag{13}$$

where $\boldsymbol{K} = [\boldsymbol{\alpha}_{1,h}, ..., \boldsymbol{\alpha}_{g,h}]$.

It is important to point out that the columns of $\dot{\boldsymbol{X}}$ contain the values of the principal components. These will be linear combinations of the initial controlled variables, obtained, using as coefficient vectors, the eigenvectors of $\boldsymbol{X}'\boldsymbol{X}$ associated with the non-zero eigenvalues.

Matrices $\boldsymbol{X}$ and $\dot{\boldsymbol{X}}$ will have the same range space, which allows us to consider the model as

$$\boldsymbol{V}_b = \dot{\boldsymbol{X}}\boldsymbol{\gamma}_g + \boldsymbol{e}_b, \tag{14}$$

with $rank(\dot{\boldsymbol{X}}) = g$. We now obtain the unbiased estimator

$$\tilde{\boldsymbol{\gamma}}_g = (\dot{\boldsymbol{X}}'\dot{\boldsymbol{X}})^{-1}\dot{\boldsymbol{X}}'\boldsymbol{V}_b, \tag{15}$$

which once $\boldsymbol{V}_b$ is assumed to be normal, has normal distribution with covariance matrix

$$Cov(\tilde{\gamma}_g) = \sigma^2(\dot{\boldsymbol{X}}'\dot{\boldsymbol{X}})^{-1} = \sigma^2(\boldsymbol{K}'\boldsymbol{X}'\boldsymbol{X}\boldsymbol{K})^{-1}\sigma^2 diag(\lambda_1, ..., \lambda_g)^{-1}. \tag{16}$$

The adjusted vector will be

$$\tilde{\boldsymbol{V}_b} = \sum_{j=1}^{g}(\tilde{\gamma}_j \cdot \dot{\boldsymbol{x}}_{jb}), \tag{17}$$

with $\tilde{\boldsymbol{\gamma}}_g = (\tilde{\gamma}_1, ..., \tilde{\gamma}_g)$. With $\boldsymbol{x}_{1,b}, ..., \boldsymbol{x}_{h,b}$ the column vectors of matrix $\boldsymbol{X}$ and $\alpha_{1,j}, ..., \alpha_{h,j}$ the components of the $\boldsymbol{\alpha}_{j,h}, j = 1, ..., g$, we have, by (13),

$$\dot{\boldsymbol{x}}_{j,b} = \sum_{l=1}^{h} \alpha_{l,j}\boldsymbol{x}_{l,b}, \qquad j = 1, ..., g, \tag{18}$$

which replaced in (17), gives:

$$\tilde{\boldsymbol{V}_b} = \sum_{j=1}^{g}(\tilde{\gamma}_j \cdot \dot{\boldsymbol{x}}_{j,b}) = \sum_{j=1}^{g}(\tilde{\gamma}_j \cdot \sum_{l=1}^{h} \alpha_{l,j}\boldsymbol{x}_{l,b}) \sum_{l=1}^{h}(\sum_{j=1}^{g} \alpha_{l,j}\tilde{\gamma}_j)\boldsymbol{x}_{l,b} = \sum_{l=1}^{h}(\tilde{\delta}_l\boldsymbol{x}_{l,b}), \tag{19}$$

with $\tilde{\delta}_l, l = 1, ..., h$, the components of the vector we wanted to estimate. We now have the estimated vector

$$\tilde{\boldsymbol{\delta}}_{\boldsymbol{h}} = \boldsymbol{K} \cdot \tilde{\gamma}_g, \tag{20}$$

which (see Seber, 1980, p.5) will be normal, with mean vector $\boldsymbol{\delta}_{\boldsymbol{h}} = \boldsymbol{K} \cdot \boldsymbol{\gamma}_g$ and covariance matrix

$$Cov(\tilde{\boldsymbol{\delta}}_{\boldsymbol{h}}) = \sigma^2\boldsymbol{C}, \tag{21}$$

where $\boldsymbol{C} = \boldsymbol{K}\boldsymbol{D}\boldsymbol{K}' = [c_{l,j}]$, with $\boldsymbol{D} = diag(\lambda_1^{-1}, ..., \lambda_g^{-1})$.

If $\boldsymbol{V}_b$ is independent from $S \sim \sigma^2\chi^2_p$, for testing the hypothesis

$$H_{l,o} : \delta_l = \delta_{l,o}, \qquad l = 1, ..., h \tag{22}$$

we can use the test statistics

$$t(\delta_{l,o}) = \frac{\tilde{\delta}_l - \delta_{l,o}}{\sqrt{c_{l,l}\frac{S}{p}}}, \qquad l = 1, ..., h \tag{23}$$

which, when the corresponding hypothesis holds, have central t distribution with p degrees of freedom.

In the application we are to present the components of $\boldsymbol{V_b}$ were the mean values of samples of 4 observations. Each of these samples corresponded to a pair (location, year). Thus, S was the sum of sums of squares of residues for the different samples.

3. Application

The existing data corresponds to the yields observed from a set of experiments integrated in a Wheat Breeding Program that was carried out by the Portuguese Plant Breeding Station. There are cultivars that were discarded from the plan and others that were introduced into the plan over the years. The years in which the different locations were used are given in Table 1. In Table 2 the rows correspond to the cultivars, while the columns correspond to the years. The presence or absence of a cultivar in the program in a given year is indicated by 1 [0].

Using the yields of the cultivars from the various environments (location, year), the "classical environmental indexes" were given by the mean yields for each environment. We point out that each one of these "classical indexes" represents the contribution of the cultivars presented in the corresponding environment. Using those "classical indexes" as a starting point, the Zig-Zag algorithm was used for adjustment of the goal function (1), giving the adjusted vector of regression coefficients and the adjusted L_2 environmental indexes vector see Pinto (2006). Both adjusted environmental indexes and classical indexes are shown in Table 3.

We used the results in Table 3 to adjust the model presented in subsection 2.1. In order to make clear how the model was applied to the data, we state that:

- ∘ the adjusted environmental indexes are the observations of the vector of dependent variables, $\boldsymbol{V_b}$, with $b = 41$;
- ∘ there are eleven years, $I = 11$ and sixteen locations, $J = 16$;
- ∘ the matrix $\boldsymbol{U}$ has $r = 1 + (I - 1) + (J - 1) = 26$ columns and $b = 41$ rows, as presented in Table 4;

Table 1. Years and locations for the field trials.

Experiment	Locations	Years
1	Almeirim	1986/87
2	Évora	1986/88/89
3	Coruche	1986/87
4	Mirandela	1986/89/92
5	Comenda	1986/95/97/99
6	Fundão	1987/88/89
7	E.N.M.P.	1988/89/90/91/92
8	Lamaçais	1988
9	Beja	1988/89/90/91/95/97
10	Benavila	1990/92
11	Elvas	1990
12	Revilheira	1990/95/97/99
13	Santarém	1991
14	Abrantes	1991
15	V.F.Xira	1992/99
16	M.Alhos	2000

- the vector of coefficients to be estimated, $\boldsymbol{\eta}_r$, has the first component, $\eta_1 = \mu$, corresponding to the general mean, $I - 1 = 10$ components, $\eta_2, ..., \eta_{11}$ corresponding to year effects, and $J - 1 = 15$ components, $\eta_{12}, ..., \eta_{26}$ corresponding to location effects;
- $\boldsymbol{Z}$ is a vector with $b = 41$ components, the classical indexes presented in Table 3;
- γ is the sole slope to be adjusted.

Since it is not possible to present the complete matrix $\boldsymbol{U}$ in one table, we write it as $\boldsymbol{U} = [\boldsymbol{U}_1, \boldsymbol{U}_2]$, and present the sub-matrices in Tables 4.1 and 4.2. The first sub-matrix $\boldsymbol{U}_1$ will contain the columns corresponding to the general mean and to years, while the sub-matrix $\boldsymbol{U}_2$ contain the columns corresponding to the locations. The locations will be ranked according to the indexes given in Table 1. The first column in Table 4.1 and 4.2 indicates the pair (location, year) for the corresponding field experiment, thus the first one was carried out in 1986 at Almeirim.

In Table 5 the adjusted coefficients are presented, as well as the t tests for nullity. The significance level of these tests is indicated as * 5%, ** 1% and *** 0.1%.

Table 2. Presence and absence of cultivars during the plan.

	1986	1987	1988	1989	1990	1991	1992	1995	1997	1999	2000
anza	1	1	1	1	1	1	1	1	1	1	1
lima1	1	1	1	1	1	1	0	0	0	0	0
te8401	1	0	0	0	0	0	0	0	0	0	0
flycatcher	1	1	1	0	0	0	0	0	0	0	0
te8501	1	1	1	0	0	0	0	0	0	0	0
te8502	1	1	1	0	0	0	0	0	0	0	0
te8504	1	1	1	0	0	0	0	0	0	0	0
hahn-s	1	1	0	0	0	0	0	0	0	0	0
sunbird-s	1	0	0	0	0	0	0	0	0	0	0
neelkant-s	1	0	0	0	0	0	0	0	0	0	0
miwivet-s	1	1	0	0	0	0	0	0	0	0	0
te8601	0	1	1	0	0	0	0	0	0	0	0
te8602	0	1	1	0	0	0	0	0	0	0	0
te8603	0	1	1	1	1	0	0	0	0	0	0
te8701	0	0	1	1	1	0	0	0	0	0	0
te8702	0	0	1	1	1	0	0	0	0	0	0
almansor	0	0	0	1	1	1	1	1	1	1	1
liz1	0	0	0	1	0	0	0	0	0	0	0
liz2	0	0	0	1	1	0	0	0	0	0	0
alva	0	0	0	1	0	0	0	0	0	0	0
te8801	0	0	0	1	1	0	0	0	0	0	0
te8802	0	0	0	1	1	1	0	0	0	0	0
te8901	0	0	0	0	1	1	0	0	0	0	0
te8902	0	0	0	0	1	1	0	0	0	0	0
te9001	0	0	0	0	0	1	0	0	0	0	0
te9002	0	0	0	0	0	1	1	0	0	0	0
milan	0	0	0	0	0	1	0	0	0	0	0
te9003	0	0	0	0	0	1	0	0	0	0	0
te8906	0	0	0	0	0	1	0	0	0	0	0
mondego	0	0	0	0	0	0	1	0	0	0	0
te9101	0	0	0	0	0	0	1	0	0	0	0
te9102	0	0	0	0	0	0	1	0	0	0	0
te9111	0	0	0	0	0	0	1	1	0	0	0
te9112	0	0	0	0	0	0	1	1	0	0	0
te9113	0	0	0	0	0	0	1	1	1	0	0
te9114	0	0	0	0	0	0	1	1	1	0	0
te9203	0	0	0	0	0	0	0	1	1	1	1
te9301	0	0	0	0	0	0	0	1	1	0	0
te9302	0	0	0	0	0	0	0	1	1	0	0
te9303	0	0	0	0	0	0	0	1	1	0	0
te9406	0	0	0	0	0	0	0	1	1	1	1
te9503	0	0	0	0	0	0	0	0	1	1	1
te9504	0	0	0	0	0	0	0	0	1	1	1
te9712	0	0	0	0	0	0	0	0	0	1	1
te9713	0	0	0	0	0	0	0	0	0	1	1
te9714	0	0	0	0	0	0	0	0	0	1	1
te9715	0	0	0	0	0	0	0	0	0	1	1
te9716	0	0	0	0	0	0	0	0	0	1	1

Table 3. "Classical" and L_2 estimated environmental Indexes per year and location.

Trial	L_2 Indexes	Classical Indexes
1986, Almeirim	1.853	1.860
1986, Évora	0.699	0.708
1986, Coruche	5.477	5.462
1986, Mirandela	6.234	6.222
1986, Comenda	4.183	4.148
1987, Almeirim	4.843	4.826
1987, Coruche	4.248	4.185
1987, Fundão	3.970	4.190
1988, Évora	1.584	1.592
1988, Fundão	4.441	4.463
1988, ENMP	3.367	3.380
1988, Lamaçais	2.720	2.714
1988, Beja	3.448	3.464
1989, Évora	2.920	2.948
1989, Mirandela	6,025	6.037
1989, Fundão	2.491	2.514
1989, E.N.M.P.	4.154	4.210
1989, Beja	4.300	4.362
1990, E.N.M.P.	2.386	2.366
1990, Beja	2.558	2.550
1990, Benavila	0.945	0.947
1990, Elvas	0.311	0.311
1990, Revilheira	0.351	0.358
1991, E.N.M.P.	4.680	4.674
1991, Beja	3.509	3.512
1991, Santarém	1.837	1.838
1991, Abrantes	2.147	2.158
1992, Mirandela	3.271	3.311
1992, E.N.M.P.	1.783	1.770
1992, Benavila	2.410	2.401
1992, V.F.Xira	4.816	4.789
1995, Comenda	1.859	1.848
1995, Beja	2.112	2.115
1995, Revilheira	1.665	1.661
1997, Comenda	3.314	3.306
1997, Beja	2.823	2.839
1997, Revilheira	1.160	1.166
1999, Comenda	3.828	3.816
1999, Revilheira	3.054	3.066
1999, V.F.Xira	5.271	5.285
2000, M.Alhos	1.542	1.204

Table 4.1. Sub-Matrix $\boldsymbol{U}_1$.

	intercept	1986	1987	1988	1989	1990	1991	1992	1995	1997	1999
1986, Almeirim	1	1	0	0	0	0	0	0	0	0	0
1986, Évora	1	1	0	0	0	0	0	0	0	0	0
1986, Coruche	1	1	0	0	0	0	0	0	0	0	0
1986, Mirandela	1	1	0	0	0	0	0	0	0	0	0
1986, Comenda	1	1	0	0	0	0	0	0	0	0	0
1987, Almeirim	1	0	1	0	0	0	0	0	0	0	0
1987, Coruche	1	0	1	0	0	0	0	0	0	0	0
1987, Fundão	1	0	1	0	0	0	0	0	0	0	0
1988, Évora	1	0	0	1	0	0	0	0	0	0	0
1988, Fundão	1	0	0	1	0	0	0	0	0	0	0
1988, ENMP	1	0	0	1	0	0	0	0	0	0	0
1988, Lamaais	1	0	0	1	0	0	0	0	0	0	0
1988, Beja	1	0	0	1	0	0	0	0	0	0	0
1989, Évora	1	0	0	0	1	0	0	0	0	0	0
1989, Mirandela	1	0	0	0	1	0	0	0	0	0	0
1989, Fundão	1	0	0	0	1	0	0	0	0	0	0
1989, E.N.M.P.	1	0	0	0	1	0	0	0	0	0	0
1989, Beja	1	0	0	0	1	0	0	0	0	0	0
1990, E.N.M.P.	1	0	0	0	0	1	0	0	0	0	0
1990, Beja	1	0	0	0	0	1	0	0	0	0	0
1990, Benavila	1	0	0	0	0	1	0	0	0	0	0
1990, Elvas	1	0	0	0	0	1	0	0	0	0	0
1990, Revilheira	1	0	0	0	0	1	0	0	0	0	0
1991, E.N.M.P.	1	0	0	0	0	0	1	0	0	0	0
1991, Beja	1	0	0	0	0	0	1	0	0	0	0
1991, Santarém	1	0	0	0	0	0	1	0	0	0	0
1991, Abrantes	1	0	0	0	0	0	1	0	0	0	0
1992, Mirandela	1	0	0	0	0	0	0	1	0	0	0
1992, E.N.M.P.	1	0	0	0	0	0	0	1	0	0	0
1992, Benavila	1	0	0	0	0	0	0	1	0	0	0
1992, V.F.Xira	1	0	0	0	0	0	0	1	0	0	0
1995, Comenda	1	0	0	0	0	0	0	0	1	0	0
1995, Beja	1	0	0	0	0	0	0	0	1	0	0
1995, Revilheira	1	0	0	0	0	0	0	0	1	0	0
1997, Comenda	1	0	0	0	0	0	0	0	0	1	0
1997, Beja	1	0	0	0	0	0	0	0	0	1	0
1997, Revilheira	1	0	0	0	0	0	0	0	0	1	0
1999, Comenda	1	0	0	0	0	0	0	0	0	0	1
1999, Revilheira	1	0	0	0	0	0	0	0	0	0	1
1999, V.F.Xira	1	0	0	0	0	0	0	0	0	0	1
2000, M.Alhos	1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1

Table 4.2. Sub-Matrix $\boldsymbol{U}_2$.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1986, Almeirim	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1986, Évora	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
1986, Coruche	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
1986, Mirandela	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
1986, Comenda	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
1987, Almeirim	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1987, Coruche	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
1987, Fundão	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
1988, Évora	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
1988, Fundão	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
1988, ENMP	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
1988, Lamaçais	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
1988, Beja	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
1989, Évora	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
1989, Mirandela	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
1989, Fundão	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
1989, E.N.M.P.	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
1989, Beja	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
1990, E.N.M.P.	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
1990, Beja	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
1990, Benavila	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
1990, Elvas	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
1990, Revilheira	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
1991, E.N.M.P.	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
1991, Beja	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
1991, Santarém	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
1991, Abrantes	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
1992, Mirandela	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
1992, E.N.M.P.	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
1992, Benavila	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
1992, V.F.Xira	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
1995, Comenda	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
1995, Beja	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
1995, Revilheira	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
1997, Comenda	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
1997, Beja	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
1997, Revilheira	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
1999, Comenda	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
1999, Revilheira	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
1999, V.F.Xira	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
2000, M.Alhos	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1

Table 5. Adjusted Coefficients and t-tests.

Variables	Adjusted Coefficients	t-tests
intercept	0.0153	0.1336
1986	-0.0265	-0.2990
1987	-0.0766	-0.5704
1988	0.0064	0.0722
1989	-0.0120	-0.2315
1990	0.0062	0.0625
1991	0.0137	0.1173
1992	-0.0007	-0.0066
1995	0.0001	0.0011
1997	-0.0092	-0.0938
1999	-0.0217	-0.2060
2000	0.1282	0,8680
Almeirim	0.0389	0.2884
Évora	-0.0180	-0.1493
Coruche	0.0718	0.5084
Mirandela	-0.0168	-0.1232
Comenda	0.0132	0.1318
Fundão	-0.0763	-0.7408
E.N.M.P.	-0.0250	-0.3181
Lamaçais	-0.0175	-0.1036
Beja	-0.0359	-0.4824
Benavila	-0.0156	-0.1278
Elvas	-0.0216	-0.1245
Revilheira	-0.0156	-0.1566
Santarém	-0.0317	-0.1643
Abrantes	-0.0419	-0.2227
V.F. Xira	-0.0015	-0.0090
M.Alhos	0.1935	1,0930
Covariate	1.0008	24.5501***

4. Conclusions

The only significant result was for the covariate, so there is no additional influence from either years or locations. This point is important since it shows the basic linearity of the problem. Thus if the linear structure assumed did not completely describe what happens, the residues would show departures from linearity. These departures would very probably be connected with years and locations. Thus this study provides an additional validation of the basic JRA model.

Acknowledgments

The authors wish to thank Professor Tadeusz Caliński for his suggestion of using the "classical indexes" as covariate. This suggestion led us to find the linear structure of the problem.

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.

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.

Judge G.G., Hill R.C., Griffiths W.E., Lütkepohl H.L., Lee Tsoung-Chao (1988): Introduction to the theory and practise of econometrics. John Wiley and Sons, Inc. $2^{nd} edition$.

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

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

Pinto I. (2006): AJoint Regression Analysis and Plant Breeding Programs. New University of Lisbon, Technology and Sciences Faculty.

Seber G.A.F. (1980): A General theory. 2^{nd} ed., Charles Griffin and Co - London.

Scheffé H. (1959): The Analysis of Variance. John Willey and Sons - New York.

Searle S.R. (1987): Linear Models for unbalanced Data. John Wiley and Sons, Inc–New York.