

Multi-treatment regression analysis: the unbalanced case

Elsa Estevão Moreira and João Tiago Mexia

Abstract—Under multi-treatment regression analysis, instead of a sample for each treatment of a linear model, there is a linear regression in the same variables. Then, instead of the action of the treatments on the sample mean values, the action on regression coefficients is studied. When data is unbalanced, the regression matrices differs between regressions. This problem is solved through the use of a block-wise diagonal covariance matrix in the ANOVA procedures. The methodology was then applied to data obtained from experiments of electrochemical removal of 3 heavy metals from contaminated wood. First, polynomial regressions of the 4th and 3rd were fitted to each metal concentration in the electrolytes through time. Then the unbalanced case of multi-treatment regression analysis was applied aiming to choose the best treatment in jointly removing the 3 metals. Results pointed to the choice of treatment 1 as the most efficient.

Keywords—ANOVA, F tests, multiple regression, Scheffé multiple comparison, unbalanced data.

I. INTRODUCTION

The present paper is weighted either toward the methodology developed either for its application using real data situation. The methodology developed, concerning the multi-treatment regression analysis for the unbalanced case, is exposed and its utility is demonstrated simultaneously through the application. Some emphasis is given to the case study, since it can interest from the point of view of modeling and problem solving. However, in this paper the case study should be viewed as a mean to explain and demonstrate the applicability of the method.

In a multi-treatment regression model, for each treatment - combination of factors levels - of a linear base model, instead of a random sample, there is a multiple regression in the same variables, both controlled and dependent [2], [6], [7]. The linear base model can be for instance an one-way, a two-way layout ANOVA, a factorial model with fixed effects, a cross-nested model, etc [1], [5], [6]. The multiple regressions will correspond to the treatments of this linear base model. Then, instead of the influence of the treatments on the sample mean values of the samples, the influence on the regression coefficients is studied. ANOVA algorithms and multiple comparison methods are adapted to perform the comparison between the coefficients of different regressions.

In a previous paper, a similar case study was treated using the regular case of multi-treatment regression analysis [7]. In that case study the data was balanced, i.e., the number of observations in each regression assigned to a treatment was the same. However, in the present paper we present the solution for the case of having different observations for regression, that is, different model matrices.

II. MULTI-TREATMENT REGRESSION ANALYZES: THE UNBALANCED CASE

Suppose that there are L treatments in a linear base model, thus L multiple regressions all with k controlled variables

$$\mathbf{y}_l = \mathbf{X}_l \boldsymbol{\beta}_l + \mathbf{e}_l, l = 1, \dots, L$$

where $\mathbf{y}_l = \begin{bmatrix} y_{l1} \\ \vdots \\ y_{ln_l} \end{bmatrix}$ is the vector of observations with n_l components and mean vector $\boldsymbol{\mu}_l = \mathbf{X}_l \boldsymbol{\beta}_l, l = 1, \dots, L,$

$$\mathbf{X}_l = \begin{bmatrix} 1 & x_{11}^{(l)} & \cdots & x_{1k}^{(l)} \\ 1 & x_{21}^{(l)} & \cdots & x_{2k}^{(l)} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n_l 1}^{(l)} & \cdots & x_{n_l k}^{(l)} \end{bmatrix}_{n_l \times (k+1)}$$

is the model matrix for the l -th regression with linearly independent column vectors, thus $\text{Rank}(\mathbf{X}_l) = k + 1.$

$\boldsymbol{\beta}_l = \begin{bmatrix} \beta_{l0} \\ \vdots \\ \beta_{lk} \end{bmatrix}$ is the vector of coefficients, and $\mathbf{e}_l =$

$\begin{bmatrix} e_{l1} \\ \vdots \\ e_{ln_l} \end{bmatrix}$ is the vector of the random error for the l -th regression, $l = 1, \dots, L.$ We assume as usually, that $\mathbf{e}_l, l = 1, \dots, L$ is normally distributed with null mean value and covariance matrix $\sigma^2 \mathbf{I}_{n_l}$ and in order to perform a multi-treatment regression analysis, we also have to assume that there are independence and equality of variances σ^2 between the different regressions - homoscedasticity. This assumption is reasonable in this kind of analysis, since the regressions are always of the same type, thus the variances should be similar across the regressions even with different sample sizes.

In the regular case of multi-treatment regression, the model matrix \mathbf{X} is the same for all regressions, here, as presented, we can have different model matrices in the regressions.

The estimators for the vectors of coefficients are normal and given by

$$\begin{aligned} \tilde{\boldsymbol{\beta}}_l &= \left(\mathbf{X}_l^T \mathbf{X}_l \right)^{-1} \mathbf{X}_l^T \mathbf{y}_l \\ &\sim N(\boldsymbol{\beta}_l, \sigma^2 \left(\mathbf{X}_l^T \mathbf{X}_l \right)^{-1}), l = 1, \dots, L \end{aligned} \quad (1)$$

and are also independent from the sum of square for errors of the regression $SSE_l = \mathbf{y}_l^T \mathbf{y}_l - \mathbf{y}_l^T \mathbf{X}_l \tilde{\boldsymbol{\beta}}_l \sim \sigma^2 \chi_{n_l - k - 1}^2, l = 1, \dots, L$ [5].

Now, let \mathbf{a} be a vector of coefficients for a linear combination of the regression coefficients $y_l^* = \mathbf{a}^T \tilde{\boldsymbol{\beta}}_l$ with mean value

$\mu_l^* = \mathbf{a}^T \beta_l, l = 1, \dots, L$. The vector of the linear combinations of the regression coefficients for the L regressions $\mathbf{y}^* = \begin{bmatrix} y_1^* \\ \vdots \\ y_L^* \end{bmatrix}$ with mean vector $\mu^* = \begin{bmatrix} \mu_1^* \\ \vdots \\ \mu_L^* \end{bmatrix}$ will be the vector of observations used in the base model. Then, it is easy to prove that

$$\begin{cases} \mu^* = [\mathbf{a}^T \beta_1 \dots \mathbf{a}^T \beta_L]^T = [\mathbf{I}_L \otimes \mathbf{a}^T][\beta_1^T \dots \beta_L^T]^T \\ \mathbf{y}^* = [\mathbf{a}^T \tilde{\beta}_1 \dots \mathbf{a}^T \tilde{\beta}_L]^T = [\mathbf{I}_L \otimes \mathbf{a}^T][\tilde{\beta}_1^T \dots \tilde{\beta}_L^T]^T \end{cases}$$

[6], and

$$\text{COV}(\mathbf{y}^*) = [\mathbf{I}_L \otimes \mathbf{a}^T] \text{COV}([\tilde{\beta}_1^T \dots \tilde{\beta}_L^T]^T) [\mathbf{I}_L \otimes \mathbf{a}^T]^T$$

where \otimes is the Kronecker product of matrices [9]. Now, on account of (1)

$$\text{COV}([\tilde{\beta}_1^T \dots \tilde{\beta}_L^T]^T) = \sigma^2 \mathbf{D} \left((\mathbf{X}_1^T \mathbf{X}_1)^{-1}, \dots, (\mathbf{X}_L^T \mathbf{X}_L)^{-1} \right)$$

where $\mathbf{D} \left((\mathbf{X}_1^T \mathbf{X}_1)^{-1}, \dots, (\mathbf{X}_L^T \mathbf{X}_L)^{-1} \right)$ is a matrix block-wise diagonal [6]. So,

$$\mathbf{y}^* \sim N(\mu^*, \sigma^2 \mathbf{W}^*)$$

with

$$\mathbf{W}^* = [\mathbf{I}_L \otimes \mathbf{a}^T] \mathbf{D} \left((\mathbf{X}_1^T \mathbf{X}_1)^{-1}, \dots, (\mathbf{X}_L^T \mathbf{X}_L)^{-1} \right) [\mathbf{I}_L \otimes \mathbf{a}^T]^T.$$

In a multi-treatment regression analysis the usual aim is to compare the k homologue coefficients of the L regressions, choosing the following \mathbf{a} vectors:

$$\mathbf{a}^T = [1 \ 0 \ \dots \ 0] \Rightarrow y_l^* = \mathbf{a}^T \tilde{\beta}_l = \tilde{\beta}_{l,1} \Rightarrow \mathbf{y}^* = \begin{bmatrix} \tilde{\beta}_{1,1} \\ \vdots \\ \tilde{\beta}_{L,1} \end{bmatrix},$$

for $l = 1, \dots, L$;

\vdots

$$\mathbf{a}^T = [0 \ 0 \ \dots \ 1] \Rightarrow y_l^* = \mathbf{a}^T \tilde{\beta}_l = \tilde{\beta}_{l,k} \Rightarrow \mathbf{y}^* = \begin{bmatrix} \tilde{\beta}_{1,k} \\ \vdots \\ \tilde{\beta}_{L,k} \end{bmatrix},$$

for $l = 1, \dots, L$.

A. Test of hypotheses

We may want to test hypotheses about the influence of the treatments on the linear combinations of regression coefficients

$$H_0(j) : \mathbf{A}_j \mu^* = \mathbf{0}, j = 1, \dots, m$$

where \mathbf{A}_j is a matrix of contrasts with null sums for the elements in the different rows $j = 1, \dots, m$, with m the number of hypotheses to test regarding the parameters of the linear base model [1], [6]. These matrices are used to correctly formulate the null hypothesis. An example of this matrix for one-way ANOVA base model with L treatments is

$$\mathbf{A}_1 = \begin{bmatrix} 1 & -1 & 0 & \dots & 0 \\ 1 & 0 & -1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & \dots & 0 & -1 \end{bmatrix}_{(L-1) \times L}, \quad \text{with}$$

$$\text{Rank}(\mathbf{A}_1) = L - 1.$$

Note that testing $H_0 : \mathbf{A}_1 \mu^* = \mathbf{0}$ is equivalent to test $H_0 : \beta_{1,i} = \dots = \beta_{L,i}, i = 1, \dots, k$. For a two-way base model and more complex linear models a general formulation for $\mathbf{A}_j, j = 1, \dots, m$ can be found in [6].

Now, let $\eta_j = \mathbf{A}_j \mu^*$ and $\tilde{\eta}_j = \mathbf{A}_j \mathbf{y}^*$, then

$$\tilde{\eta}_j \sim N(\eta_j, \sigma^2 \mathbf{W}_j^*), j = 1, \dots, m$$

where

$$\mathbf{W}_j^* = \mathbf{A}_j [\mathbf{I}_L \otimes \mathbf{a}^T] \mathbf{D} \left[(\mathbf{X}_1^T \mathbf{X}_1)^{-1}, \dots, (\mathbf{X}_L^T \mathbf{X}_L)^{-1} \right] [\mathbf{I}_L \otimes \mathbf{a}^T]^T \mathbf{A}_j^T$$

and we have the sum of squares for the treatments

$$SST_j = \tilde{\eta}_j^T \mathbf{W}_j^{*+} \tilde{\eta}_j \sim \sigma^2 \chi_{r_j, \delta_j}^2$$

where $r_j = \text{Rank}(\mathbf{W}_j^*)$ and $\delta_j = \frac{1}{\sigma^2} \eta_j^T \mathbf{W}_j^{*+} \eta_j, j = 1, \dots, m$, where \mathbf{W}_j^{*+} is the generalized inverse of Moore-PenRose [9].

Since $SST_j, j = 1, \dots, m$ is independent from each sum of squares for the errors of the L regressions, is also independent from the sum of those sums

$$SSSE = \sum_{l=1}^L SSE_l \sim \sigma^2 \chi_g^2$$

with $g = \sum_{l=1}^L (n_l - k - 1)$. Then, we can use the F test statistics

$$\mathcal{F}_j = \frac{g}{r_j} \frac{SST_j}{SSSE} \sim F(z|r_j, g, \delta_j), j = 1, \dots, m$$

to test the $H_0(j), j = 1, \dots, m$ [2], [3], [6], [8]. When $H_0(j)$ holds then, $\delta_j = 0, j = 1, \dots, m$ and

$$\mathcal{F}_j = \frac{g}{r_j} \frac{SST_j}{SSSE} \sim F(z|r_j, g)$$

i.e., \mathcal{F}_j has central F distribution with r_j and g degrees of freedom. Then $H_0(j)$ is rejected at the significance level q test, if $\mathcal{F}_j > F_{1-q, r_j, g}, j = 1, \dots, m$, with $F_{1-q, r_j, g}$ the $(1 - q)$ -th quantile for an F distribution with r_j and g degrees of freedom, $j = 1, \dots, m$.

B. Scheffé multiple comparison method

Using the Scheffé theorem [8], we have

$$\Pr \left[\bigcap_{\mathbf{d}} \left(|\mathbf{d}^T \eta_j - \mathbf{d}^T \tilde{\eta}_j| \leq \sqrt{r_j F_{1-q, r_j, g} \mathbf{d}^T \mathbf{W}_j^* \mathbf{d} \frac{SSSE}{g}} \right) \right] = 1 - q$$

where $\bigcap_{\mathbf{d}}$ indicates that all vectors $\mathbf{d} \in \mathbb{R}^n$ are considered.

So, the simultaneous confidence intervals with joint confidence level q for all the $\mathbf{d}^T \eta_j$, are given by the inequalities

$$|\mathbf{d}^T \eta_j - \mathbf{d}^T \tilde{\eta}_j| \leq \sqrt{r_j F_{1-q, r_j, g} \mathbf{d}^T \mathbf{W}_j^* \mathbf{d} \frac{SSSE}{g}} \quad (2)$$

[4], [8]. When $\mathbf{d}^T \eta_j = 0$, the inequality (2) can not be satisfied if

$$|0 - \mathbf{d}^T \tilde{\eta}_j| > \sqrt{r_j F_{1-q, r_j, g} \mathbf{d}^T \mathbf{W}_j^* \mathbf{d} \frac{SSSE}{g}} \quad (3)$$

then, we can conclude that $\mathbf{d}^T \eta_j$ is significantly different from 0 at significance level q [4]. For instance, taking

$$\mathbf{d} = [0 \ \dots \ 0 \ 1 \ 0 \ \dots \ 0 \ -1 \ 0 \ \dots \ 0]^T$$

TABLE I
EXPERIMENTAL CONDITIONS (TREATMENTS)

| Treatments | Assisting agent | Initial current (mA) | Duration (days) |
|------------|---------------------------------------|----------------------|-----------------|
| 1 | 2.5% oxalic acid | 40 | 14 |
| 2 | 2.5% oxalic acid | 60 | 14 |
| 3 | 2.5% oxalic acid | 120 | 14 |
| 4 | 5.0% formic acid: 2.5% oxalic acid | 40 | 14 |

where the values 1 and -1 correspond to the l -th l' -th vector components, we get

$$\mathbf{d}^T \boldsymbol{\eta}_j = \eta_l - \eta_{l'}$$

and we can detect which pairs of components of $\boldsymbol{\eta}_j, j = 1, \dots, m$ differ significantly [4], [8].

III. CASE STUDY

Electrodialytic remediation is a method that uses a direct electric current as cleaning agent. Chromated copper arsenate (CCA) is the most common formulation that has been used to preserve wood. In spite of CCA usefulness, due to its strong fixation in wood, chromium and arsenic are hazardous to human health and present a potential threat to the environment. The movement of charged particles in an electrical field is applied to CCA-treated wood waste, to assist Cu (copper), Cr (Chromium) and As (Arsenic) removal.

Experiments of electrodialytic remediation were conducted using wood chips from wood treated with CCA. For the statistical analysis of data collected from experiments, the treatments to be compared were defined by: a) the type and percentage of the extracting solution used to saturate the wood waste; b) the initial current passing through the electrodialytic cell and c) duration of the procedure.

Several electrodialytic experiments were carried out in a electrodialytic cell and use different extracting solutions, initial currents, and durations. Experimental conditions, corresponding to each different combination of those factors will define a treatment of the wood chips. During each experiment, samples of the electrolyte solutions were periodically collected and analyzed for Cu, Cr and As determination.

After a first selection, based on another kind of data, just 4 experiments corresponding to the different treatments presented in Table 1 were selected to be analyzed under multi-treatment regression analysis. In Table 2, the collected data is presented and in Figure 1, the same data is presented graphically.

In Table 2, we can observe that the data is unbalanced, i.e., different number of observations per treatment are available. In the treatment 1, 2 and 3, one observation per day of experiment was collected. However, in treatment 2 and 3 something went wrong in the last day of experiments and the observation could not be collected correctly, thus it had to be eliminated. For treatment 4, more than one observation per day was collected. Our main concern was to compare the time evolution of the 4 experiments without having to throw away some of the available data.

TABLE II
EXPERIMENTAL DATA: CONCENTRATION OF EACH HEAVY METAL IN THE ELECTROLYTE SOLUTIONS COLLECTED DURING THE TIME OF THE EXPERIMENTS (14 DAYS)

| days | Treat. 1 | | | Treat. 2 | | | Treat. 3 | | | Treat. 4 | | |
|------|----------|------|-------|----------|------|-------|----------|------|-------|----------|------|------|
| | Cu | Cr | As | Cu | Cr | As | Cu | Cr | As | days | Cu | Cr |
| 0 | 0.00 | 0.00 | 0.11 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.15 | 0.0 | 0.00 | 0.00 |
| 1 | 2.99 | 2.94 | 15.99 | 1.69 | 2.94 | 11.04 | 1.59 | 2.79 | 10.47 | 0.5 | 0.00 | 0.00 |
| 2 | 2.61 | 4.80 | 21.14 | 2.14 | 4.75 | 12.35 | 2.13 | 3.88 | 12.65 | 0.6 | 0.03 | 0.04 |
| 3 | 3.11 | 5.73 | 22.40 | 2.54 | 5.54 | 16.51 | 2.57 | 4.22 | 18.76 | 0.7 | 0.31 | 0.24 |
| 4 | 3.76 | 6.19 | 25.70 | 2.81 | 6.03 | 20.16 | 2.95 | 4.37 | 19.65 | 0.8 | 0.66 | 0.39 |
| 5 | 4.18 | 6.21 | 30.55 | 4.13 | 6.32 | 30.09 | 3.43 | 4.37 | 20.67 | 1.5 | 0.42 | 2.56 |
| 6 | 4.60 | 6.23 | 35.39 | 4.32 | 6.10 | 32.72 | 3.90 | 4.38 | 20.64 | 2.5 | 2.49 | 3.85 |
| 7 | 4.54 | 6.03 | 38.06 | 4.57 | 5.95 | 29.53 | 4.44 | 4.31 | 20.91 | 3.4 | 2.29 | 4.16 |
| 8 | 4.59 | 5.77 | 37.75 | 4.78 | 5.90 | 31.22 | 5.21 | 4.43 | 20.80 | 3.7 | 2.42 | 4.14 |
| 9 | 4.61 | 5.37 | 38.55 | 4.89 | 5.63 | 34.73 | 5.44 | 4.58 | 21.45 | 4.4 | 3.09 | 3.91 |
| 10 | 4.60 | 5.17 | 37.21 | 5.04 | 5.51 | 35.07 | 5.70 | 4.55 | 21.38 | 4.7 | 2.77 | 3.78 |
| 11 | 4.39 | 4.81 | 36.84 | 5.27 | 5.68 | 33.56 | 5.97 | 4.52 | 21.45 | 5.4 | 4.03 | 3.36 |
| 12 | 4.18 | 4.45 | 36.48 | 5.25 | 5.58 | 34.69 | 6.46 | 4.52 | 20.88 | 6.4 | 5.45 | 2.81 |
| 13 | 2.22 | 4.98 | 36.95 | 5.08 | 5.21 | 35.91 | 6.99 | 4.89 | 22.50 | 7.4 | 6.85 | 2.32 |
| 14 | 1.99 | 4.74 | 37.46 | | | | | | | 8.0 | 6.89 | 1.92 |
| | | | | | | | | | | 9.4 | 6.93 | 1.52 |
| | | | | | | | | | | 10.4 | 7.00 | 1.25 |
| | | | | | | | | | | 11.4 | 7.28 | 1.09 |
| | | | | | | | | | | 12.4 | 7.56 | 0.91 |
| | | | | | | | | | | 13.4 | 7.37 | 0.79 |
| | | | | | | | | | | 14.4 | 6.10 | 0.70 |

Fig. 1. Concentration of Cu, Cr and As in the electrolytes through time

A multi-treatment regression analysis, the unbalanced case, was then used to compare the 4 different treatments with the aim to find the most efficient treatment for removal all heavy metals in the shortest possible time.

A. Modeling

Before proceeding to the multi-treatment regression analysis, in order to perform the comparison between the 4 treatments, polynomial regressions were fitted to the data in Figure 1. Using standard statistical techniques, namely the least squares method and significance tests for regression coefficients [5], a 4th-degree polynomial regression

$$y = \beta_1 t + \beta_2 t^2 + \beta_3 t^3 + \beta_4 t^4 + e$$

was fitted for Cu and Cr, while for As a 3rd-degree polynomial regression

$$y = \beta_1 t + \beta_2 t^2 + \beta_3 t^3 + e$$

was fitted, where t is time in days, y the concentration of the heavy metals in the electrolytes, $\beta_i, i = 1, 2, 3, 4$ are the regression coefficients and e error term for which it is assumed the normality with zero mean value and variance σ^2 (equality of variances). The adjusted polynomials are homogeneous since at time 0 no metal would have been removed. The estimates of the regression coefficients for the 4 treatments and the values of the R^2 are presented in Table 3.

We also decided to obtain the instantaneous removal speed given by the 1st derivatives of the polynomials. In the case of 4th-degree polynomial, the expression of the removal speed are given by

$$y = \beta_1 + 2\beta_2 t + 3\beta_3 t^2 + 4\beta_4 t^3 + e.$$

This removal speed is important, since in our aim the duration of the removal process should be the shortest possible and the first impulse for heavy metals mobilization to the electrolytes is crucial to this duration. Moreover, the coefficient of lower degree of t , β_1 , has a high weight in the regression (Table 3),

Fig. 2. 4th degree polynomials fitted to Cu concentrations and corresponding removal speed

Fig. 3. 4th degree polynomials fitted to Cr concentrations and corresponding removal speed

thus it is relevant for the determination of the initial speed of heavy metals mobilization to the electrolytes.

In Figures 2, 3 and 4 are graphically presented the 4th and 3rd degree polynomials fitted to Cu, Cr and As concentrations respectively, as well as the corresponding curves for the removal speed.

B. Multi-treatment regression analysis of the heavy metals concentrations

In this section, we are going to apply the multi-treatment regression analysis to our problem. As a starting point, there are 4 polynomial regressions of the 4th degree for Cu and Cr and 3rd degree for As, fitted to each metal concentrations in the electrolytes through time. Next, we are going to exemplify the procedure just for the 4th degree polynomials.

Let

$$y_l = X_l \beta_l + e_l, l = 1, 2, 3, 4$$

be the 4 polynomials in matrix notation, where $\beta_l = [\beta_{l,1} \beta_{l,2} \beta_{l,3} \beta_{l,4}]^T$ are vector of coefficients for the l -th polynomial,

$$X_1 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 4 & 8 & 16 \\ 3 & 9 & 27 & 81 \\ 4 & 16 & 64 & 256 \\ 5 & 25 & 125 & 625 \\ 6 & 36 & 216 & 1296 \\ 7 & 49 & 343 & 2401 \\ 8 & 64 & 512 & 4096 \\ 9 & 81 & 729 & 6561 \\ 10 & 100 & 1000 & 10000 \\ 11 & 121 & 1331 & 14641 \\ 12 & 144 & 1728 & 20736 \\ 13 & 169 & 2197 & 28561 \\ 14 & 196 & 2744 & 38416 \end{bmatrix}$$

is the model matrix for treatment 1, in which we had one observation per day during 14 days. The model matrix for treatment 2 and 3 is equal to that for treatment 1 without the

Fig. 4. 3rd degree polynomials fitted to As concentrations and corresponding removal speed

TABLE III
ESTIMATES FOR THE REGRESSION COEFFICIENTS

| | | Treatment | | | |
|----------|-----------------|-----------|--------|--------|---------|
| | | 1 | 2 | 3 | 4 |
| Copper | $\hat{\beta}_4$ | -0.001 | 0.000 | -0.001 | 0.001 |
| | $\hat{\beta}_3$ | 0.036 | 0.013 | 0.030 | -0.023 |
| | $\hat{\beta}_2$ | -0.393 | -0.167 | -0.296 | 0.240 |
| | $\hat{\beta}_1$ | 2.079 | 1.326 | 1.230 | 0.090 |
| | R^2 | 0.760 | 0.960 | 0.900 | 0.970 |
| | SSE | 2.843 | 0.779 | 0.396 | 3.860 |
| Chromium | $\hat{\beta}_4$ | -0.001 | -0.002 | -0.002 | 0.000 |
| | $\hat{\beta}_3$ | 0.041 | 0.057 | 0.061 | 0.021 |
| | $\hat{\beta}_2$ | -0.596 | -0.698 | -0.668 | -0.371 |
| | $\hat{\beta}_1$ | 3.364 | 3.514 | 2.968 | 2.089 |
| | R^2 | 0.970 | 0.990 | 0.810 | 0.860 |
| | SSE | 0.294 | 0.089 | 0.574 | 5.744 |
| Arsenic | $\hat{\beta}_3$ | 0.033 | 0.016 | 0.045 | 0.007 |
| | $\hat{\beta}_2$ | -1.047 | -0.619 | -1.111 | -0.478 |
| | $\hat{\beta}_1$ | 10.950 | 8.009 | 8.702 | 7.518 |
| | R^2 | 0.910 | 0.940 | 0.900 | 0.900 |
| | SSE | 70.619 | 62.812 | 16.166 | 257.269 |

last line, and

$$X_4 = \begin{bmatrix} 0.51 & 0.26 & 0.13 & 0.07 \\ 0.58 & 0.34 & 0.20 & 0.12 \\ 0.68 & 0.47 & 0.32 & 0.22 \\ 0.77 & 0.59 & 0.45 & 0.35 \\ 1.48 & 2.18 & 3.21 & 4.73 \\ 2.46 & 6.07 & 14.97 & 36.90 \\ 3.39 & 11.46 & 38.80 & 131.36 \\ 3.68 & 13.52 & 49.71 & 182.80 \\ 4.42 & 19.51 & 86.18 & 380.64 \\ 4.72 & 22.27 & 105.07 & 495.80 \\ 5.39 & 29.05 & 156.55 & 843.76 \\ 6.39 & 40.85 & 261.12 & 1669.00 \\ 7.39 & 54.67 & 404.20 & 2988.54 \\ 8.00 & 64.00 & 512.00 & 4096.00 \\ 9.42 & 88.67 & 835.01 & 7863.01 \\ 10.43 & 108.72 & 1133.67 & 11820.92 \\ 11.42 & 130.39 & 1488.87 & 17000.99 \\ 12.42 & 154.17 & 1914.32 & 23769.50 \\ 13.43 & 180.23 & 2419.60 & 32483.08 \\ 14.43 & 208.14 & 3002.86 & 43322.56 \end{bmatrix}$$

is the model matrix for treatment 4, in which we had more than one observation per day.

In this 4-treatments regression approach, we want to compare corresponding regression coefficients of the same controlled variable (power of t) between the 4 treatments. In order to do this, we take

- 1) $\mathbf{a}^T = [1 \ 0 \ 0 \ 0]$ to compare the coefficient of t ;
- 2) $\mathbf{a}^T = [0 \ 1 \ 0 \ 0]$ to compare the coefficient of t^2 ;
- 3) $\mathbf{a}^T = [0 \ 0 \ 1 \ 0]$ to compare the coefficient of t^3 ;
- 4) $\mathbf{a}^T = [0 \ 0 \ 0 \ 1]$ to compare the coefficient of t^4 .

The base model for our problem was the simplest one: the one-way ANOVA with fixed effects and 4 treatments. The null hypotheses that we want to test for each metal are

$$H_0 : \eta_1 = \mathbf{A}_1 \mu^* = \mathbf{0}$$

with

$$\mathbf{A}_1 = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 1 & 0 & -1 & 0 \\ 1 & 0 & 0 & -1 \end{bmatrix},$$

which is equivalent to test

- 1) $H_0^1 : \beta_{1,1} = \beta_{2,1} = \beta_{3,1} = \beta_{4,1}$ equality of coefficient for t ;
- 2) $H_0^2 : \beta_{1,2} = \beta_{2,2} = \beta_{3,2} = \beta_{4,2}$ equality of coefficient for t^2 ;
- 3) $H_0^3 : \beta_{1,3} = \beta_{2,3} = \beta_{3,3} = \beta_{4,3}$ equality of coefficient for t^3 ;
- 4) $H_0^4 : \beta_{1,4} = \beta_{2,4} = \beta_{3,4} = \beta_{4,4}$ equality of coefficient for t^4 .

To find significant differences between each pair of treatments, the Scheffé multiple comparison method was applied to the cases for which we had significant F tests. The inequality in equation (3) was verified for the 6 pairwise comparisons

- 1) $\mathbf{d}^T \tilde{\eta}_1 = \beta_{1,i} - \beta_{2,i}$ taking $\mathbf{d}^T = [1 \ -1 \ 0 \ 0]$;
- 2) $\mathbf{d}^T \tilde{\eta}_1 = \beta_{1,i} - \beta_{3,i}$ taking $\mathbf{d}^T = [1 \ 0 \ -1 \ 0]$;
- 3) $\mathbf{d}^T \tilde{\eta}_1 = \beta_{1,i} - \beta_{4,i}$ taking $\mathbf{d}^T = [1 \ 0 \ 0 \ -1]$;
- 4) $\mathbf{d}^T \tilde{\eta}_1 = \beta_{2,i} - \beta_{3,i}$ taking $\mathbf{d}^T = [0 \ 1 \ -1 \ 0]$;
- 5) $\mathbf{d}^T \tilde{\eta}_1 = \beta_{2,i} - \beta_{4,i}$ taking $\mathbf{d}^T = [0 \ 1 \ 0 \ -1]$;
- 6) $\mathbf{d}^T \tilde{\eta}_1 = \beta_{3,i} - \beta_{4,i}$ taking $\mathbf{d}^T = [0 \ 0 \ 1 \ -1]$,

for the regression for the same power of t , $i = 1, \dots, k = 4$.

C. Results and discussion

The results obtained for the F tests at a 5% of significance are presented in Table 4. For the Cu, the the null hypothesis was rejected for all powers of t , for the Cr just for t and t^2 , and finally for As just for t .

TABLE IV

RESULTS OF THE F TESTS FOR 5% OF SIGNIFICANCE FOR CU, CR AND AS

| Cu | | | | |
|--------------------------------|---------|------|---------------|-----------------|
| | SQT | d.f. | \mathcal{F} | $F_{0.95;3;44}$ |
| H_0^1 - coefficient of t | 15.342 | 3 | 29.213 | 2.816 |
| H_0^2 - coefficient of t^2 | 11.568 | 3 | 22.027 | |
| H_0^3 - coefficient of t^3 | 6.300 | 3 | 11.997 | |
| H_0^4 - coefficient of t^4 | 3.791 | 3 | 7.219 | |
| SSSE | 7.877 | 44 | | |
| Cr | | | | |
| | SQT | d.f. | \mathcal{F} | $F_{0.95;3;44}$ |
| H_0^1 - coefficient of t | 3.811 | 3 | 8.341 | 2.816 |
| H_0^2 - coefficient of t^2 | 1.328 | 3 | 2.906 | |
| H_0^3 - coefficient of t^3 | 1.241 | 3 | 2.716 | |
| H_0^4 - coefficient of t^4 | 1.228 | 3 | 2.688 | |
| SSSE | 6.701 | 44 | | |
| As | | | | |
| | SQT | d.f. | \mathcal{F} | $F_{0.95;3;48}$ |
| H_0^1 - coefficient of t | 75.855 | 3 | 2.983 | 2.798 |
| H_0^2 - coefficient of t^2 | 65.075 | 3 | 2.559 | |
| H_0^3 - coefficient of t^3 | 63.357 | 3 | 2.492 | |
| SSSE | 257.269 | 48 | | |

The results of the Scheffé pairwise comparison method, also at 5% significance, are presented in Table 5. The results for the Cu can be resumed in following way:

- For β_1 , there are significant differences between treatments 1 and 4, 2 and 4, and 3 and 4;
- For β_2 , there are significant differences between treatments 1 and 4, 2 and 4, and 3 and 4;
- For β_3 , there are significant differences between treatments 1 and 4 and 3 and 4;
- For β_4 , there are significant differences between treatments 1 and 4.

This results enable us to separate significantly, at the 5% level, the treatment 1 from the other three treatments. The maximum value of β_1 is attained for treatment 1, which is equivalent to saying that treatment 1 (2.5% percentage of oxalic acid and 40 mA) is the treatment with the highest initial speed, followed by treatments 2, 3 and 4 (Table 3, Fig. 2, 3, 4). Thus, given the weight of β_1 in the regression and in the removal speed of Cu, it may be concluded that, treatment 1 is the best one, since it has the highest β_1 coefficient, meaning an higher initial speed of mobilization of Cu into the electrolytes.

Regarding Cr, there are significant differences for β_1 , between treatments 1 and 4 and treatments 2 and 4, while for β_2 just between treatments 2 and 4. So, it can be concluded that there are significant differences between treatment 4 and treatments 1 and 2, but treatment 3 is not significantly different from 4. Moreover, the maximum value for β_1 is attained for treatment 2, followed closely by treatment 1, after which comes treatments 3 and 4 (Table 3). Since β_1 has much more weight in the regression than the other coefficients, we can select both treatment 2 and 1 as being the best ones in mobilizing Cr into the electrolytes.

As for As, significant differences for the β_1 coefficient were only found between treatments 1 and 4. However, the maximum value for β_1 is attained for treatment 1, exceeding considerably the values for the other treatments. Thus, treatment 1 can also be selected with some confidence, as being the treatment with the best initial speed of As removal.

The results obtained from this analysis allowed us to conclude, mainly on account of the initial speed of removal, that:

- For Cu, oxalic acid was the best extracting solution and

TABLE V

SCHEFFÉ PAIRWISE COMPARISON - RESULTS FOR CU, CR AND AS

($\alpha = 5\%$)

| Cu | | | | |
|----------------------|-----------------------|-----------------------|-------------------------|------------|
| Treatments | 1 ^o member | 2 ^o member | significantly different | |
| Coefficient of t | | | | |
| 1 | 1 and 2 | 0.75310 | 1.28177 | no |
| 2 | 1 and 3 | 0.84847 | 1.28177 | no |
| 3 | 1 and 4 | 1.98894 | 0.85568 | yes |
| 4 | 2 and 3 | 0.09537 | 1.34997 | no |
| 5 | 2 and 4 | 1.23584 | 0.95482 | yes |
| 6 | 3 and 4 | 1.14047 | 0.95482 | yes |
| Coefficient of t^2 | | | | |
| 1 | 1 and 2 | 0.22593 | 0.50361 | no |
| 2 | 1 and 3 | 0.09646 | 0.50361 | no |
| 3 | 1 and 4 | 0.63271 | 0.32749 | yes |
| 4 | 2 and 3 | 0.12947 | 0.54715 | no |
| 5 | 2 and 4 | 0.40679 | 0.39115 | yes |
| 6 | 3 and 4 | 0.53625 | 0.39115 | yes |
| Coefficient of t^3 | | | | |
| 1 | 1 and 2 | 0.02271 | 0.06178 | no |
| 2 | 1 and 3 | 0.00554 | 0.06178 | no |
| 3 | 1 and 4 | 0.05876 | 0.03949 | yes |
| 4 | 2 and 3 | 0.01717 | 0.06906 | no |
| 5 | 2 and 4 | 0.03605 | 0.05012 | no |
| 6 | 3 and 4 | 0.05323 | 0.05012 | yes |
| Coefficient of t^4 | | | | |
| 1 | 1 and 2 | 0.00083 | 0.00238 | no |
| 2 | 1 and 3 | 0.00024 | 0.00238 | no |
| 3 | 1 and 4 | 0.00182 | 0.00148 | yes |
| 4 | 2 and 3 | 0.00058 | 0.00273 | no |
| 5 | 2 and 4 | 0.00100 | 0.00200 | no |
| 6 | 3 and 4 | 0.00158 | 0.00200 | no |
| Cr | | | | |
| Treatments | 1 ^o member | 2 ^o member | significantly different | |
| Coefficient of t | | | | |
| 1 | 1 and 2 | 0.15034 | 1.19669 | no |
| 2 | 1 and 3 | 0.39549 | 1.19669 | no |
| 3 | 1 and 4 | 1.27483 | 1.02039 | yes |
| 4 | 2 and 3 | 0.54583 | 1.26036 | no |
| 5 | 2 and 4 | 1.42517 | 1.09437 | yes |
| 6 | 3 and 4 | 0.87935 | 1.09437 | no |
| Coefficient of t^2 | | | | |
| 1 | 1 and 2 | 0.10135 | 0.47018 | no |
| 2 | 1 and 3 | 0.07185 | 0.47018 | no |
| 3 | 1 and 4 | 0.22529 | 0.38547 | no |
| 4 | 2 and 3 | 0.02950 | 0.51083 | no |
| 5 | 2 and 4 | 0.32664 | 0.32413 | yes |
| 6 | 3 and 4 | 0.29714 | 0.32413 | no |
| As | | | | |
| Treatments | 1 ^o member | 2 ^o member | significantly different | |
| Coefficient of t | | | | |
| 1 | 1 and 2 | 2.94127 | 4.41590 | no |
| 2 | 1 and 3 | 2.24865 | 4.41590 | no |
| 3 | 1 and 4 | 3.43238 | 3.38553 | yes |
| 4 | 2 and 3 | 0.69262 | 4.64833 | no |
| 5 | 2 and 4 | 0.49111 | 4.05992 | no |
| 6 | 3 and 4 | 1.18374 | 4.05992 | no |

40 mA the best initial current;

- For Cr, oxalic acid was the best extracting solution and 60 mA and 40 mA were the best initial currents;
- For As, oxalic acid was the best extracting solution and 40 mA the best initial current.

Thus, taking in consideration the global results for the three metals, one can conclude that the best extracting solution is in fact oxalic acid 2.5% and the best initial current is 40 mA.

IV. CONCLUSIONS

Regarding the method used, the unbalanced case of the multi-treatment regression approach allows to perform an ANOVA analysis with several factors having for each treatment a regression model of the same type but with different observations, i.e., different model matrices. This difficulty is overcome through the use of block-wise diagonal covariance matrices.

As for the case study, the application of the method allowed us to conclude that the treatment 1 was the most efficient, i.e., in removing the three metals from the wood, the best assisting agent was 2.5% oxalic acid and the best initial current was 40 mA.

ACKNOWLEDGEMENTS

This work was partially supported by Financiamento Base 2010 ISFL-1-297 from FCT/MCTES/PT.

REFERENCES

- [1] A. Khuri, T. Mathew, and B. Sinha. *Statistical Tests for Mixed Linear Models*. John Willey & Sons, New York, 1998.
- [2] J. Mexia. Multi-treatment regression designs. *Research Studies N.1, Mathematics Department, Faculty of Sciences and Technology, Universidade Nova de Lisboa*, 1, 1987.
- [3] J. Mexia. Estimable functions, duality of f tests and scheffé multiple comparison method. *Research Studies N.1, Mathematics Department, Faculty of Sciences and Technology, Universidade Nova de Lisboa*, 1, 1988.
- [4] J. Mexia. Simultaneous confidence intervals: Generalization of the scheffé theorem. *Research Studies N.2, Mathematics Department, Faculty of Sciences and Technology, Universidade Nova de Lisboa*, 1, 1989.
- [5] D. Montgomery. *Design and Analysis of Experiments*. John Willey & Sons, New York, 1997.
- [6] E. Moreira, J. Mexia, R. Zmyslony, and M. Fonseca. L models and multiple regression designs. *Statistical Papers*, 50:pp. 2869–2885, 2009.
- [7] E. Moreira, A. Ribeiro, E. Mateus, J. Mexia, and L. Ottosen. Regressional modeling of electro-dialytic removal of cu, cr and as from cca treated timber waste: Application to sawdust. *Wood Science and Technology*, 39:pp. 291–309, 2005.
- [8] H. Scheffé. *The Analysis of Variance*. John Willey & Sons, New York, 1959.
- [9] J. Schott. *Matrix Analysis for Statistics*. Wiley Series in Probability and Statistics, New York, 1996.