

A note on testing hypotheses concerning interaction with special reference to a graphical presentation in the space of canonical variates

Dariusz Kayzer

Department of Mathematical and Statistical Methods, Poznan University of Life Sciences,
Wojska Polskiego 28, 60-637 Poznan, Poland, e-mail: dkayzer@up.poznan.pl

SUMMARY

Results of ecological studies that involve the use of multivariate analysis of variance techniques for testing various hypotheses, interesting from the point of view of comparing the linear functions of parameters, were considered. For testing the most interesting hypotheses on a variety of interaction effects and on contrasts of class means, the application of a multivariate test statistic is recommended. Canonical variate analysis is used for graphical presentation of the results of multidimensional experiments. In this paper it is shown how a generalized form of canonical variate analysis can be useful to reveal which parametric functions of a multivariate analysis of variance model are responsible for rejecting the linear hypothesis. As an example, an analysis was made of an ecological study of trace element accumulation in plants of Italian ryegrass as a method of biomonitoring of air pollution.

Key words: multivariate analysis of variance, contrasts, canonical variate analysis, trace elements, biomonitoring, air pollution

1. Introduction

When applying multivariate analysis of variance to ecological experiments, one is interested in testing various linear functions of parameters. If a hypothesis is rejected, then one may wish to identify which components of the hypothesis are responsible for its rejection. As an example, an ecological study of the effect of air pollution by trace elements on plants exposed in various environmental conditions was analysed. The response of trace element concentration is compared here using the tools of multivariate analysis of variance.

The experiment considered here was carried out during the growing season of the year 2011. Seeds of Italian ryegrass (*Lolium multiflorum* L.) were sown in 5 pots filled with a standard mixture of soil and peat. The plants were watered with deionized water to avoid the additional application of heavy metals. Five sites were selected for the investigations, located in the city of Poznań and surrounding areas. The sites varied in terms of air quality characteristics – there were two city sites (no. 1 and 2), one site in a suburban area (no. 4, in Luboń near Wielkopolska National Park), one site in an agricultural area (no. 5, Tarnowo Podgórne) and one site located in a landscape park (no. 3, Turew). The plants were exposed for 28 days. Two exposure series were carried out during the growing season in the following periods: 16 May to 12 June and 11 July to 7 August. Five pots with plants were exposed at each site. Similar sets of plants (no. 6: control) were cultivated in greenhouse conditions with a very low probability of atmospheric heavy metal pollution. Collected samples of plants were dried and milled, and then an Elan DRC II ICP-MS instrument (Perkin Elmer SCIEX, Ontario, Canada) was used to determine As, Cd, Cr, Ni and Pb levels.

In the present paper we would like to show how a multivariate test statistic can be split into component elements by the application of interaction effects or contrasts of class means. The canonical variate analysis used here makes it possible to present graphically in Euclidean space the comparison of trace element concentrations in terms of interaction effects.

2. Materials and methods

2.1. Mathematical model of observations

Let the model for the observation y_{ikjm} concerning the i -th trace element ($i = 1, \dots, I$; here $I = 5$), the k -th exposure site ($k = 1, \dots, K$; here $K = 6$), the j -th series ($j = 1, \dots, J$; here $J = 2$), and the m -th replication ($m = 1, \dots, M$; here $M = 5$) be of the form:

$$y_{ikjm} = \mu_i + \xi_{ik}^1 + \xi_{ij}^2 + \xi_{ijk}^{12} + e_{ikjm} \quad (1)$$

where for the i -th selected (one of five) trace element, μ_i is the grand mean, ξ_{ik}^1 is the k -th exposure site effect, ξ_{ij}^2 is the j -th exposure series effect, ξ_{ijk}^{12} is the k,j -th effect of exposure site \times series interaction, and e_{ikjm} is the random error, which follows an independent distribution with zero expected value and with variance σ_i^2 .

Due to the potential correlations between trace element concentrations, it is appropriate to use tools of multivariate analysis for the determination of differences between effects of sites and series. Here it is assumed that $[y_{1kjm}, \dots, y_{Ikjm}]'$ for any k, j and m is an independent vector random variable with dispersion matrix Σ_{kjm} . For the estimation of parameters and analysis of variance it is additionally assumed that the dispersion matrices for all vectors $[y_{1kjm}, \dots, y_{Ikjm}]'$ are equal, which means that $\Sigma_{kjm} = \Sigma$ for any k, j and m . The multivariate linear model can then be written in the form:

$$Y = \mathbf{1}_N \boldsymbol{\mu}' + X_1 \mathbf{E}_1 + X_2 \mathbf{E}_2 + X_{12} \mathbf{E}_{12} + U \quad (2)$$

where

$$Y = \begin{bmatrix} y_{1111} & \dots & y_{I111} \\ \dots & & \dots \\ y_{1KJM} & \dots & y_{IKJM} \end{bmatrix} \text{ is the } N \times I \text{ matrix of observations } (N=KJM),$$

$\boldsymbol{\mu} = [\mu_1 \quad \dots \quad \mu_I]'$ is the $I \times 1$ vector of general means,

$$\mathbf{E}_1 = \begin{bmatrix} \xi_{11}^1 & \dots & \xi_{I1}^1 \\ \dots & & \dots \\ \xi_{1K}^1 & \dots & \xi_{IK}^1 \end{bmatrix} \text{ is the } K \times I \text{ matrix of exposure site effects,}$$

$$\mathbf{E}_2 = \begin{bmatrix} \xi_{11}^2 & \dots & \xi_{I1}^2 \\ \dots & & \dots \\ \xi_{1J}^2 & \dots & \xi_{IJ}^2 \end{bmatrix} \text{ is the } J \times I \text{ matrix of exposure series effects,}$$

$$\mathbf{E}_{12} = \begin{bmatrix} \xi_{111}^{12} & \dots & \xi_{I11}^{12} \\ \dots & & \dots \\ \xi_{1KJ}^{12} & \dots & \xi_{IKJ}^{12} \end{bmatrix} \text{ is the } KJ \times I \text{ matrix of sites } \times \text{ series interaction}$$

parameters, $X_1 = I_K \otimes \mathbf{1}_{JM}$, $X_2 = \mathbf{1}_J \otimes I_K \otimes \mathbf{1}_M$, $X_{12} = I_{KJ} \otimes \mathbf{1}_M$, are design matrices (the symbol \otimes denotes the Kronecker product of matrices),

$$\mathbf{U} = \begin{bmatrix} e_{1111} & \cdots & e_{I111} \\ \cdots & & \cdots \\ e_{1KJM} & \cdots & e_{IKJM} \end{bmatrix} \text{ is the } N \times I \text{ matrix of errors.}$$

Finally, the considered model can be described in the form:

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\Xi} + \mathbf{U} \quad (3)$$

where $\mathbf{X} = [\mathbf{1}_N : \mathbf{X}_1 : \mathbf{X}_2 : \mathbf{X}_{12}]$ and $\boldsymbol{\Xi}' = [\boldsymbol{\mu} : \boldsymbol{\Xi}'_1 : \boldsymbol{\Xi}'_2 : \boldsymbol{\Xi}'_{12}]$. When estimating the matrix $\boldsymbol{\Xi}$ or testing a hypothesis concerning it, it is worth noting that the considered design is orthogonal relative to the model assumptions. This allows estimation and hypothesis testing concerning a group of parameters to be performed independently of the values of parameters of other groups (Seber 1980, Seber 1984). In this paper we assume that the random terms of the model follow the assumptions of a normal distribution.

2.2. Testing interaction effects

Multifactorial experiments are usually performed to make inferences concerning interaction effects. We can formalize these questions in the form of the hypothesis $H_{0;12}: \boldsymbol{\Xi}_{12} = \mathbf{0}$, that there is no site-by-series interaction. In this paper we pay special attention to the issue of structuring interactions.

To obtain the best linear unbiased estimator (BLUE) of the sites \times series interaction matrix $\boldsymbol{\Xi}_{12}$, one uses the formula:

$$\hat{\boldsymbol{\Xi}}_{12} = (\mathbf{C}_1 \otimes \mathbf{C}_2)(\mathbf{X}'_{12}\mathbf{X}_{12})^{-1}\mathbf{X}'_{12}\mathbf{Y} = (\mathbf{C}_1 \otimes \mathbf{C}_2)[\mathbf{I}_{KJ} \otimes M^{-1}\mathbf{1}'_M]\mathbf{Y} \quad (4)$$

with $\mathbf{C}_1 = \mathbf{I}_K - \frac{1}{K}\mathbf{1}_K\mathbf{1}'_K$ and $\mathbf{C}_2 = \mathbf{I}_J - \frac{1}{J}\mathbf{1}_J\mathbf{1}'_J$. Consequently, when testing the hypothesis $H_{0;12}$, one can apply the Lawley–Hotelling trace statistic in the form:

$$T_{12}^2 = (N - r)(\mathbf{E}^{-1}\mathbf{H}_{12}) \quad (5)$$

where $\mathbf{E} = \mathbf{Y}'(\mathbf{I}_N - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}')\mathbf{Y}$ and $\mathbf{H}_{12} = \mathbf{Y}'\mathbf{X}_{12}(\mathbf{X}'_{12}\mathbf{X}_{12})^{-1}(\mathbf{C}'_1 \otimes \mathbf{C}'_2)[(\mathbf{C}_1 \otimes \mathbf{C}_2)(\mathbf{X}'_{12}\mathbf{X}_{12})^{-1}(\mathbf{C}'_1 \otimes \mathbf{C}'_2)]^{-1}(\mathbf{C}_1 \otimes \mathbf{C}_2)(\mathbf{X}'_{12}\mathbf{X}_{12})^{-1}\mathbf{X}'_{12}\mathbf{Y}$, for which the relevant degrees of freedom are $m_E = N - r$, r being the rank of \mathbf{X} , and $m_{H_{12}} = (K - 1)(J - 1)$, respectively. Note, however, that because \mathbf{C}_1 and \mathbf{C}_2 are idempotent matrices, the matrix \mathbf{H}_{12} reduces here to

$\mathbf{H}_{12} = \mathbf{M}\mathbf{Y}'(\mathbf{I}_{KJ} \otimes \mathbf{M}^{-1} \mathbf{1}'_M)(\mathbf{C}_1 \otimes \mathbf{C}_2)(\mathbf{I}_{KJ} \otimes \mathbf{M}^{-1} \mathbf{1}'_M)\mathbf{Y}$. It can also be written as $\mathbf{H}_{12} = \mathbf{M}\widehat{\boldsymbol{\Xi}}'_{12}\widehat{\boldsymbol{\Xi}}_{12}$; this can be seen to coincide with the relevant results in Table 9.9 of Seber (1984). Thus, when applying the McKeon (1974) approximation of the T_{12}^2 distribution by an F -distribution, the hypothesis $H_{0;12}$ is rejected, provided that $F_{12} = \frac{1}{c} \text{tr}(\mathbf{E}^{-1}\mathbf{H}_{12}) > F_{a,b}(\alpha)$, where: $a = m_{H_{12}}I$, $b = 4 + (a + 2)/(B - 1)$, $c = a(b - 2)/(b(m_E - I - 1))$ and

$$B = \frac{(m_E + m_{H_{12}} - I - 1)(m_E - 1)}{(m_E - I - 3)(m_E - I)}.$$

When the hypothesis $H_{0;12}$ is rejected, one is usually interested in finding which of the interaction parameters are responsible for this rejection. Note that the I parameters of the interaction between the k -th site and j -th series form the kj -th row of the matrix \mathbf{H}_{12} . Then the BLUE has the form $\widehat{\boldsymbol{\Xi}}_{12(kj)} = (\mathbf{c}'_{1(k)} \otimes \mathbf{c}'_{2(j)})(\mathbf{X}'_{12}\mathbf{X}_{12})^{-1}\mathbf{X}'_{12}\mathbf{Y} = (\mathbf{c}'_{1(k)} \otimes \mathbf{c}'_{2(j)})(\mathbf{I}_{KJ} \otimes \mathbf{M}^{-1} \mathbf{1}'_M)\mathbf{Y}$, where $\mathbf{c}'_{1(k)}$ is the k -th row of \mathbf{C}_1 and $\mathbf{c}'_{2(j)}$ is the j -th row of \mathbf{C}_2 . The Lawley–Hotelling trace statistic for testing the hypothesis of no interaction between the k -th site and j -th series is of the form

$$T_{12(kj)}^2(N - r)M(\mathbf{c}'_{1(k)}\mathbf{c}_{1(k)} \otimes \mathbf{c}'_{2(j)}\mathbf{c}_{2(j)})\widehat{\boldsymbol{\Xi}}_{12(kj)}\mathbf{E}^{-1}\widehat{\boldsymbol{\Xi}}'_{12(kj)} \\ (T_{12(kj)}^2(m_E - I - 1)/(m_E I) \sim F_{I, m_E - I - 1}),$$

where, in fact, $\mathbf{c}'_{1(k)}\mathbf{c}_{1(k)} \otimes \mathbf{c}'_{2(j)}\mathbf{c}_{2(j)} = (1 - K^{-1})(1 - J^{-1})$. The relevant statistic for testing the hypothesis of no interaction between the k -th site and j -th series with respect to one among the I variables (trace elements) is obtainable from $T_{12(kj)}^2$ in an obvious way.

2.3. Canonical variate analysis

To identify the structure of trace element concentration in Italian ryegrass leaves, it is useful to present the position of points representing sites \times series interaction effects in the space of the two first canonical variates. Canonical variate analysis is a method which makes possible a graphical presentation of the results of multidimensional experiments (Lejeune and Caliński 2000). According to the method, we have to transform the matrix $\widehat{\boldsymbol{\Xi}}_{12}$ into a set of new variables, which

carry information on the structure of the interaction, but allow representation of the objects in a multivariate Euclidean space (Kayzer et al. 2009, Borowiak et al. 2011, Budka et al. 2011, Kayzer et al. 2015). Following the transformation, the matrix $\widehat{\mathbf{E}}_{12}$ can be expressed in the form:

$$\widehat{\mathbf{E}}_{12} = \sum_{h=1}^{\nu} \lambda_{12h}^{-1/2} \boldsymbol{\psi}_{12h} \boldsymbol{\varphi}'_{12h}, \quad (6)$$

where $\nu = \min(I, KJ - 1)$ and the vectors $\boldsymbol{\psi}_{12h}$ and $\boldsymbol{\varphi}_{12h}$ and scalars λ_{12h} are determined from equations of the form:

$$(N - r) \widehat{\mathbf{E}}_{12} \mathbf{E}^{-1} \widehat{\mathbf{E}}'_{12} [(\mathbf{C}_1 \otimes \mathbf{C}_2) (\mathbf{X}'_{12} \mathbf{X}_{12})^{-1} (\mathbf{C}'_1 \otimes \mathbf{C}'_2)]^{-1} \boldsymbol{\psi}_{12h} = \lambda_{12h} \boldsymbol{\psi}_{12h}, \quad (7)$$

$$(N - r) \mathbf{H}_{12} \mathbf{E}^{-1} \boldsymbol{\varphi}_{12h} = \lambda_{12h} \boldsymbol{\varphi}_{12h}. \quad (8)$$

The vectors $\boldsymbol{\psi}_{12h}$ and $\boldsymbol{\varphi}_{12h}$ are normalized in the following way:

$$\boldsymbol{\psi}'_{12h} [(\mathbf{C}_1 \otimes \mathbf{C}_2) (\mathbf{X}'_{12} \mathbf{X}_{12})^{-1} (\mathbf{C}'_1 \otimes \mathbf{C}'_2)]^{-1} \boldsymbol{\psi}_{12h'} = \begin{cases} \lambda_{12h}, & \text{if } h = h' \\ 0, & \text{if } h \neq h' \end{cases}, \quad (9)$$

$$(N - r) \boldsymbol{\varphi}'_{12h} \mathbf{E}^{-1} \boldsymbol{\varphi}_{12h'} = \begin{cases} \lambda_{12h}, & \text{if } h = h' \\ 0, & \text{if } h \neq h' \end{cases}. \quad (10)$$

In the case of the sites \times series interaction effects, the vector $\boldsymbol{\psi}_{12h}$ is called the h -th canonical coordinate, and the vector $\lambda_{12h}^{-1/2} \boldsymbol{\varphi}_{12h}$ is called the h -th dual canonical coordinate.

2.4. Testing contrasts among means of classes

In many cases the experimenter is interested in testing hypotheses concerning both interaction and main effects. This question can be expressed by the hypothesis $H_{0;12}^*: \mathbf{C}_3 \mathbf{E}_{12}^* = \mathbf{0}$, where $\mathbf{C}_3 = \mathbf{I}_{KJ} - \frac{1}{KJ} \mathbf{1}_{KJ} \mathbf{1}'_{KJ}$. The best linear unbiased estimator of the matrix \mathbf{E}_{12}^* is obtained using the formula $\widehat{\mathbf{E}}_{12}^* = (\mathbf{X}'_{12} \mathbf{X}_{12})^{-1} \mathbf{X}'_{12} \mathbf{Y}$ (Anderson 2003, Morrison 1967). Also, the matrix \mathbf{C}_3 is used to analyse the differences between effects of experimental objects. In this case, this concerns the trace element concentrations in Italian ryegrass for a particular site exposed in certain exposure series. An object effect is determined as differences between values of trace element concentration in the plant for a specified experimental object and the mean for all objects. Note, however,

that the matrix E may be calculated from the form $E = Y'(I_N - X_{12}(X'_{12}X_{12})^{-1}X'_{12})Y$. If the hypothesis $H_{0;12}^*$ is rejected, this may be either due to the fact that the site factor causes some interaction with the series factor, or because there are some significant main effects of the former factor, or both.

2.5. Testing differences between sites and control

Let us consider another hypothesis $H_{0;12}^{**} : C_4\hat{\mathcal{E}}_{12}^* = \mathbf{0}$, where $C_4 = \frac{1}{j}\mathbf{1}'_j \otimes [I_{K-1} \quad -\mathbf{1}_{K-1}]$. The elements of the matrix $C_4\hat{\mathcal{E}}_{12}^*$ are mean differences in trace element concentration in leaves between the individual exposure sites and the control.

In addition, the experimenter may be interested in testing the hypotheses $H_{0;12(1)}^* : C_{5(1)}\hat{\mathcal{E}}_{12}^* = \mathbf{0}$, where $C_{5(1)} = [I_{K-1} \quad -\mathbf{1}_{K-1} \quad \mathbf{0}_{K-1} \quad \vdots \quad \mathbf{0}_{K-1}]$. Using the matrix $C_{5(1)}$, we would like to make inferences on the differences between the effects of individual exposure sites and the control for the first exposure series. The inference will concern the j -th series when we use the matrix $C_{5(j)} = [\mathbf{0}_{K-1} \quad \vdots \quad \mathbf{0}_{K-1} \quad I_{K-1} \quad -\mathbf{1}_{K-1} \quad \mathbf{0}_{K-1} \quad \vdots \quad \mathbf{0}_{K-1}]$ ($(j-1)K$ Columns $\mathbf{0}_{K-1}$, ..., $(j-j)K$ Columns $\mathbf{0}_{K-1}$). Moreover, the elements of the k -th column of matrix $C_{5(j)}\hat{\mathcal{E}}_{12}^*$ represent mean differences in the concentration of the k -th trace element in leaves. The relevant statistics for testing the hypotheses $H_{0;12}^*$, $H_{0;12}^{**}$, $H_{0;12(j)}^*$ and canonical variate analysis decompositions of the matrices $C_3\hat{\mathcal{E}}_{12}^*$, $C_4\hat{\mathcal{E}}_{12}^*$ and $C_{5(j)}\hat{\mathcal{E}}_{12}^*$ are obtainable from the statistic T_{12}^2 (5) and equations (6)–(10) in an obvious way.

3. Results and discussion

Effects of the various sources of variation were examined using the Lawley–Hotelling trace test statistic T^2 and the approximation F . Results of the testing of hypotheses, according to the formula (5), are presented in Table 1. It can be concluded that the trace element content varies due to sites, series and interaction between sites and series, because the corresponding empirical significance levels are lower than 0.05.

Table 1. MANOVA of all parameters

Source of variability	Hotelling–Lawley statistic T^2	Approx. F statistic	Numerator degrees of freedom	Denominator degrees of freedom	Empirical significance level
sites	297.1	55.3	5	54	< 0.01
series	181.3	8.23	20	105	< 0.01
interaction	105.9	4.81	20	105	< 0.01

Comparison of the results for site \times series interaction parameters and tests of the hypotheses concerning those effects (Table 2) showed that the effects of the trace element concentrations in Italian ryegrass leaves in the plants exposed at sites no. 6, 4 and 3 were significantly different from zero. The first series at site no. 4 and the second series at site no. 3 had a similar interaction effect (Figure 1). These effects were characterized by high relative concentrations of chromium in the plants in comparison to the other effects (Table 2). Additionally, the results based on canonical variate analysis, presented in Figure 1, showed that the interaction effects for the control plants exposed in both series were the most different from the results from other exposure sites. This was caused by the accumulation of arsenic and lead in the leaves of the plants.

Table 2. Estimates of interaction effect ($\hat{\mathcal{E}}_{12}$) in terms of trace element content of Italian ryegrass

site	series	Cr [$\mu\text{g}/\text{kg}$]	Ni [$\mu\text{g}/\text{kg}$]	As [$\mu\text{g}/\text{kg}$]	Cd [$\mu\text{g}/\text{kg}$]	Pb [$\mu\text{g}/\text{kg}$]	$T_{12(kj)}^2$
1	1 st	0.085	-0.017	-0.004	-0.032	-0.042	2.65
2	1 st	-0.177	-0.209**	0.112	-0.029	-0.049	18.6*
3	1 st	-0.643**	-0.115	-0.030	-0.054**	0.014	30.6**
4	1 st	0.755**	0.247	-0.016	0.052*	0.083*	29.8**
5	1 st	-0.094	-0.101	0.036	0.044	0.091	11.4
control	1 st	0.074	0.195	-0.098**	0.019	-0.096*	33.9**
1	2 nd	-0.085	0.017	0.004	0.032	0.042	2.65
2	2 nd	0.177	0.209**	-0.112	0.029	0.049	18.6*
3	2 nd	0.643**	0.115	0.030	0.054**	-0.014	30.6**
4	2 nd	-0.755**	-0.247	0.016	-0.052*	-0.083*	29.8**
5	2 nd	0.094	0.101	-0.036	-0.044	-0.091	11.4
control	2 nd	-0.074	-0.195	0.098**	-0.019	0.096*	33.9**

* significance level $\alpha=0.05$

** significance level $\alpha=0.01$

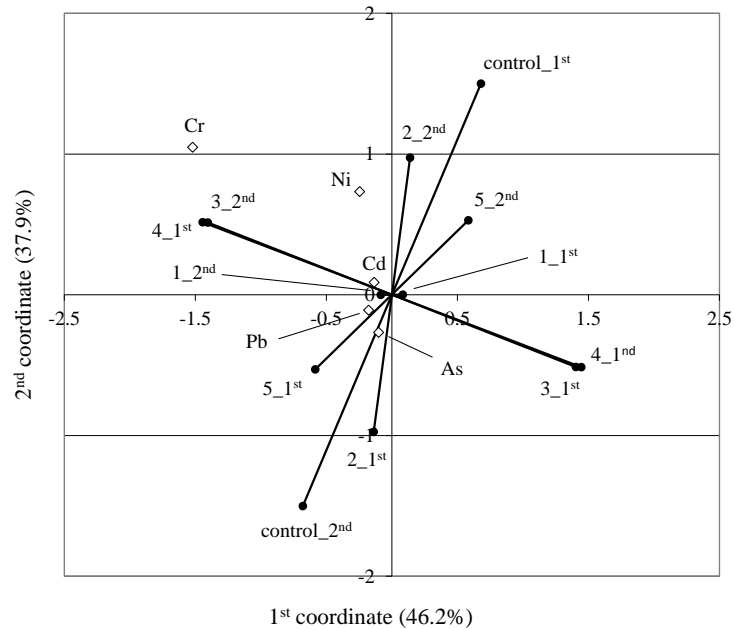
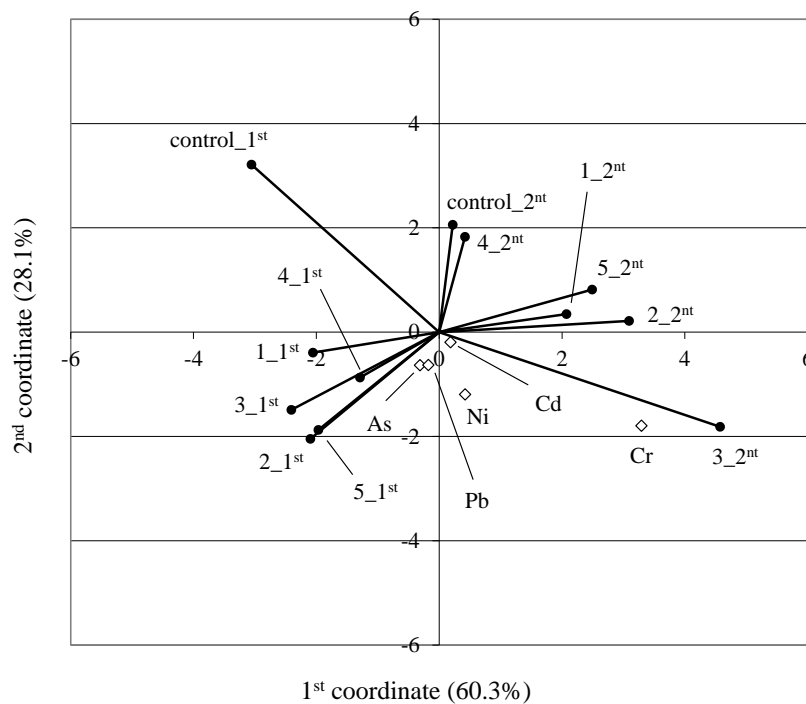


Figure 1. Position of selected experimental objects describing interaction effects in the space of the first two canonical variates, and spacing of the trace elements in the dual space of canonical variates

Values of object means compared with the general means are presented in Table 3. Based on these results it can be concluded that levels of chromium and cadmium accumulation in the leaves of Italian ryegrass were lower in the first series than in the second series (excluding the results obtained for site no. 4 in the second series), while values of arsenic and lead concentration were higher in the first series than in the second series. Furthermore, it is noted that trace element contents were lowest in the plants cultivated in greenhouse conditions. Values of interaction estimates ($C_3\hat{\mathcal{E}}_{12}^*$) after transformation to Euclidean space are presented in Figure 2. It is observed that the main direction of variation is connected with the division into experimental series. Points representing trace element concentrations for the first series are located on one side of the horizontal axis, while those for the second series lie on the other side.

Table 3. Estimates of interaction ($C_3\hat{\varepsilon}_{12}^*$) in terms of trace element content of Italian ryegrass

site	series	Cr [$\mu\text{g}/\text{kg}$]	Ni [$\mu\text{g}/\text{kg}$]	As [$\mu\text{g}/\text{kg}$]	Cd [$\mu\text{g}/\text{kg}$]	Pb [$\mu\text{g}/\text{kg}$]	Test statistic (T^2)
1	1 st	-0.567	0.150	0.094*	-0.048	0.120*	25.27**
2	1 st	-0.578	0.099	0.318**	0.020	0.082	57.60**
3	1 st	-0.425	0.253	0.224**	-0.055	0.165**	45.08**
4	1 st	-0.269	-0.040	0.067	-0.042	0.174**	22.29*
5	1 st	-0.628*	0.014	0.175**	0.044	0.259**	44.18**
control	1 st	-1.915**	-0.347	-0.100*	-0.111**	-0.222**	117.1**
1	2 nd	0.725*	0.141	-0.157**	0.081*	0.012	28.15**
2	2 nd	1.236**	0.473*	-0.165**	0.142**	-0.012	63.87**
3	2 nd	2.321**	0.440	0.024	0.117**	-0.055	136.9**
4	2 nd	-0.319	-0.576*	-0.160**	-0.083*	-0.184**	26.28**
5	2 nd	1.020**	0.173	-0.156**	0.020	-0.116	38.18**
control	2 nd	-0.601*	-0.780**	-0.164**	-0.085*	-0.222**	32.55**

* significance level $\alpha=0.05$ ** significance level $\alpha=0.01$ **Figure 2.** Position of selected experimental objects describing class means in the space of the first two canonical variates, and spacing of the trace elements in the dual space of canonical variates

Testing of differences in trace element concentrations in Italian ryegrass leaves between the mean values for individual exposure sites and the mean values for the control site revealed statistical differences (Table 4). The lowest levels of all trace elements were recorded at site no. 4. The others sites had similar values of trace element concentrations. The position of exposure series in the canonical space was mainly determined by variability in chromium and nickel concentration (Figure 3).

Table 4. Estimates of differences of trace element concentrations in Italian ryegrass leaves between mean values for individual exposure sites and the mean value for all exposure sites ($C_4\hat{\mathcal{E}}_{12}^*$) (calculated for both experimental series)

site	Cr [$\mu\text{g}/\text{kg}$]	Ni [$\mu\text{g}/\text{kg}$]	As [$\mu\text{g}/\text{kg}$]	Cd [$\mu\text{g}/\text{kg}$]	Pb [$\mu\text{g}/\text{kg}$]	Test statistic (T^2)
1	1.337**	0.709*	0.101	0.114**	0.288**	50.7**
2	1.587**	0.849**	0.208**	0.179**	0.257**	86.2**
3	2.206**	0.910**	0.256**	0.129**	0.277**	124.4**
4	0.964*	0.255	0.085*	0.036	0.217**	36.8**
5	1.454**	0.657**	0.142**	0.130**	0.294**	133.9**

* significance level $\alpha=0.05$

** significance level $\alpha=0.01$

Another way of testing differences between the effects of the sites and the control is to consider particular series separately. For the first series, differences in trace element concentrations between individual exposure sites and the control were found for every exposure site (Table 5). The highest relative values were observed at site no. 4 for chromium and at sites no. 2 and 3 for arsenic. Note that there were practically no differences between the effects of the sites in relation to nickel, cadmium and lead concentrations. These results are graphically presented in Figure 4, and confirm the absence of differences between sites.

In analysis of the differences in trace element concentration effects between individual exposure sites and the control for the second series, no difference was found in the case of site no. 4 (Table 6). The highest relative values were observed at site no. 3 for chromium and arsenic concentration in leaves. However, the sites

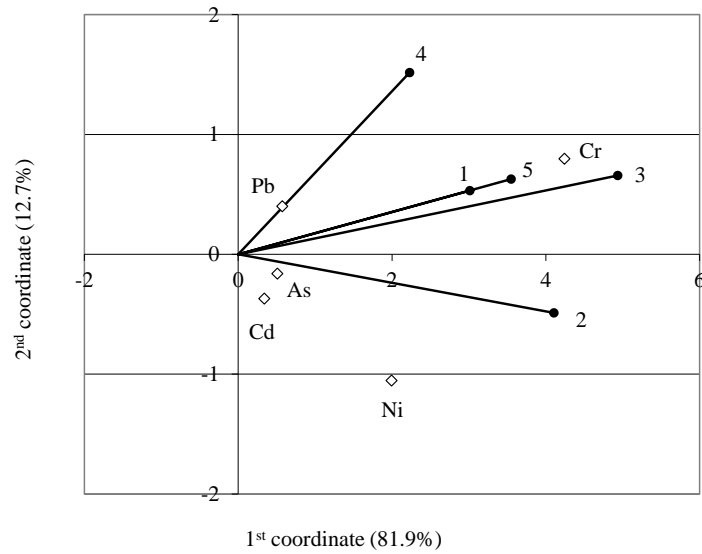


Figure 3. Position of exposure sites determined by trace element concentration in Italian ryegrass in relation to the control site in the first two canonical variates, and spacing of the heavy metal contents in the dual space (calculated for both experimental series)

Table 5. Estimates of differences of trace element concentrations in Italian ryegrass leaves between mean values for individual exposure sites and the mean value for all exposure sites ($C_{5(1)}\bar{z}_{12}^*$) (calculated for the first experimental series)

site	Cr [$\mu\text{g}/\text{kg}$]	Ni [$\mu\text{g}/\text{kg}$]	As [$\mu\text{g}/\text{kg}$]	Cd [$\mu\text{g}/\text{kg}$]	Pb [$\mu\text{g}/\text{kg}$]	Test statistic (T^2)
1	0.135**	0.0497	0.0194**	0.0063	0.0342**	40.6**
2	0.134**	0.0446	0.0418**	0.0132*	0.0304**	74.1**
3	0.149**	0.0600	0.0324**	0.0056	0.0388**	62.1**
4	0.165**	0.0307	0.0167*	0.0069	0.0396**	66.1**
5	0.129**	0.0361	0.0275**	0.0155**	0.0481**	75.9**

* significance level $\alpha=0.05$

** significance level $\alpha=0.01$

located in Poznań (no. 1 and 2) were characterized by the highest lead concentrations. These results are presented graphically in Figure 5, and confirm the outlying position of site no. 4 in terms of contents of trace elements in leaves. The investigations of differences between sites and control for the individual series showed no similar tendency for both series.

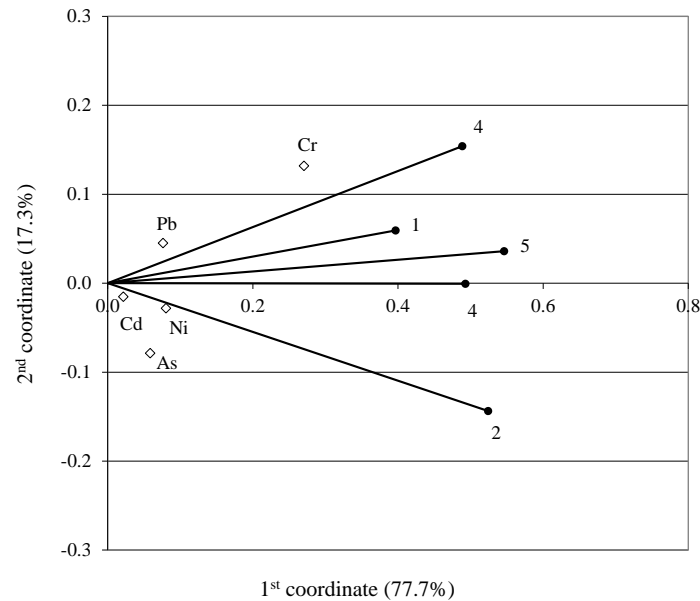


Figure 4. Position of exposure sites determined by trace element concentrations in Italian ryegrass in relation to the control site in the first two canonical variates, and spacing of the heavy metal contents in the dual space (calculated for the first experimental series)

Table 6. Estimates of the differences in trace element concentrations in Italian ryegrass leaves between mean values for individual exposure sites and the mean value for all exposure sites ($C_{5(2)}\hat{\Sigma}_{12}^*$) (calculated for the second experimental series)

Site	Cr [$\mu\text{g}/\text{kg}$]	Ni [$\mu\text{g}/\text{kg}$]	As [$\mu\text{g}/\text{kg}$]	Cd [$\mu\text{g}/\text{kg}$]	Pb [$\mu\text{g}/\text{kg}$]	Test statistic (T^2)
1	0.133**	0.0920**	0.0007	0.0166**	0.0234**	25.0*
2	0.184**	0.1252**	-0.0001	0.0227**	0.0210*	47.4**
3	0.292**	0.1219**	0.0187**	0.0202**	0.0167	89.0**
4	0.028	0.0203	0.0004	0.0002	0.0038	0.524
5	0.162**	0.0953**	0.0008	0.0105*	0.0106	22.1*

* significance level $\alpha=0.05$

** significance level $\alpha=0.01$

The usefulness of the canonical decomposition of the interaction effects matrix in visualizing the empirical departure from differences in observations has been demonstrated. The investigations of the interaction effects between levels of two factors revealed a tendency when considering class (cell) means. Use

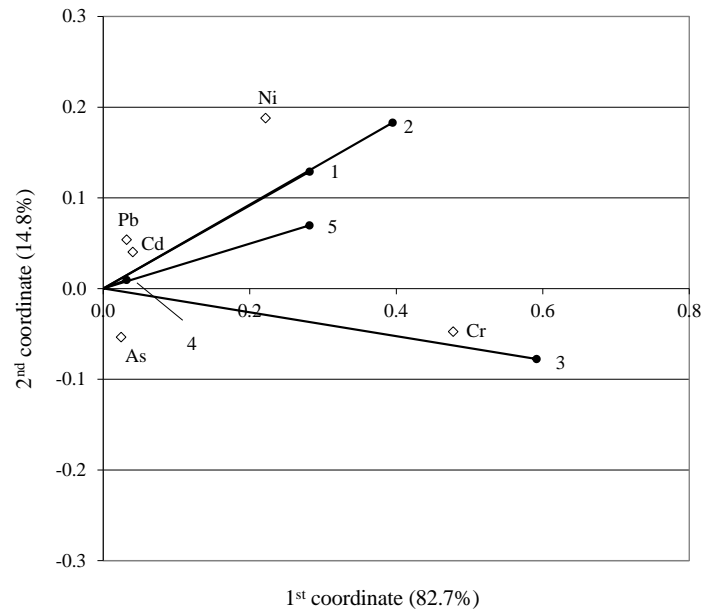


Figure 5. Position of exposure sites determined by trace element concentrations in Italian ryegrass in relation to the control site in the first two canonical variates, and spacing of the heavy metal contents in the dual space (calculated for the second experimental series)

of the canonical variate analysis method enables graphical extension of the interpretation of results obtained during an experiment in the case where multidimensional observations are classified according to linear models. Canonical variate analysis can be applied to other kinds of data matrix, such as matrices of one-way classification (Mucha et al. 2015), two-way classification with one observation per cell (Kayzer et al. 1999, Kayzer et al. 2018), block designs (Lejeune 2000) and other factorial designs (Budka et al. 2015). Canonical variate analysis is similar to another procedure called principal component analysis. Object-oriented averages and the correlation matrix may be used to visualize multidimensional dependences (Niewiadomska et al. 2018, Borowiak et al. 2018). Additionally, the additive main effects and multiplicative interaction model is another approach to evaluate results from experiments where images of experimental objects can be presented in a two-dimensional biplot (Gauch et al.

2008). For instance, results presented for second-order interaction (Neisse et al. 2018) gave the possibility of a similar interpretation as in canonical variate analysis (Caliński et al. 1987).

4. Conclusions

The study described here shows multidimensional analysis of variance using the canonical variate approach to be an appropriate and recommended technique for testing and visualizing the results of environmental experiments concerning interaction effects. It makes it possible to explain the reasons for the rejection of the considered hypothesis. This can lead to much deeper and broader inference from the experiment.

Acknowledgements

The author wishes to acknowledge the support of project grant no. N 0540/B/P01/2011/40 from the National Science Centre for the publication of this work. I am grateful to Prof. Danuta Barańkiewicz and Dr. Anetta Hanć of Adam Mickiewicz University, Department of Trace Element Analysis by Spectroscopic Methods, for performing chemical analyses of plant samples.

REFERENCES

- Anderson T. W. (2003): An Introduction to Multivariate Statistical Analysis. Third Edition, John Wiley and Sons, New York.
- Borowiak K., Budka A., Hanć A., Kayzer D., Lisiak M., Zbierska J., Barańkiewicz D., Iwaniuk D., Łopatka N. (2018): Relations between photosynthetic pigments, macroelement contents and selected trace elements accumulated in *Lolium multiflorum* L. exposed to ambient air conditions. *Acta Biologica Cracoviensia Series Botanica* (60/1): 35-44.
- Borowiak K., Kayzer D., Budka A., Zbierska J. (2011): Study of changes in the degree of tobacco leaf injury caused by tropospheric ozone. *Biometrical Letters* 48(1): 55-66.
- Borowiak K., Kayzer D., Budka A., Zbierska J., Drzewiecka K., Bandurska H., Goliński P. (2012): Cumulative tropospheric ozone effect on visible tobacco leaf injury. *Fresenius Environmental Bulletin* 21(2a): 509-517.

- Budka A., Borowiak K., Kayzer D., Hanć A., Zbierska J., Barańkiewicz D., Wolna-Maruwka A., Lisiak M., Lisiak M. (2015): Nickel and chromium concentrations in Italian ryegrass exposed to ambient air in urban, suburban and rural areas. *Atmospheric Pollution Research* 6(6): 1123-1131.
- Caliński T., Czajka S., Kaczmarek Z. (1987): A model for the analysis of a series of experiments repeated at several places over a period of years. *Cultivar Testing Bulletin* 12(17/18): 7-71.
- Gauch H. G., Piepho H. P., Annicchiarico P. (2008): Statistical Analysis of Yield Trials by AMMI and GGE: Further Considerations. *Crop Science* 48: 866-889.
- Lejeune M., Caliński T. (2000): Canonical analysis applied to multivariate analysis of variance. *Journal of Multivariate Analysis* 72: 100-119.
- Kayzer D., Borowiak K., Budka A., Zbierska J. (2009): Study of interaction in bioindication research on tobacco plant injuries caused by ground level ozone. *Environmetrics* 20: 666-675.
- Kayzer D., Borowiak K., Budka A., Zbierska J., Lisiak M. (2015): An application of canonical variate analysis in profile comparison of dry matter content of white clover between ozone-sensitive and ozone-resistant clones exposed in ambient air conditions. *Biometrical Letters* 52(1): 23-35.
- Kayzer D., Frankowski P., Zbierska J., Staniszewski R. (2018): Evaluation of trophic parameters in newly built reservoir using canonical variates analysis, ITM Web of Conferences 23: 00019 DOI: <https://doi.org/10.1051/itmconf/20182300019>.
- McKeon J. J. (1974): F approximations to the distribution of Hotelling's T_0^2 . *Biometrika* 61: 381-383.
- Morrison D. F. (1967): *Multivariate Statistical Methods*. New York.
- Mucha J., Budka A., Kayzer D., Guzicka M., Łakomy P., Zadworny M. (2015): The origin of reactive oxygen during interaction of *Pinus sylvestris* root and *Heterobasidion annosum* s.l. – the linkage with the iron. *European Journal of Plant Pathology* 143: 277-290.
- Niewiadomska A., Sulewska H., Wolna-Maruwka A., Ratajczak K., Głuchowska K., Waraczewska Z., Budka A. (2018): An assessment of the influence of Co-inoculation with endophytic bacteria and rhizobia, and the influence of PRP SOL and PRP EBV fertilisers on the microbial parameters of soil and nitrogenase activity in yellow lupine (*Lupinus luteus* L.) cultivation. *Polish Journal of Environmental Study* 27(6): 2687-2702.
- Neisse A. C., Kirch J. L., Hongyu K. (2018): AMMI and GGE Biplot for genotype \times environment interaction: a medoid-based hierarchical cluster analysis approach for high-dimensional data. *Biometrical Letters* 55(2): 97-121.
- Seber G. A. F. (1980): *The Linear Hypothesis: A General Theory*. Charles Griffin. London.
- Seber G. A. F. (1984): *Multivariate Observations*. Wiley. New York.