Biometrical Letters
Vol. 44 (2007), No. 1, 1-15

Multivariate analysis of variance in a modified model of split-plot design with hierarchical classification for investigation of changes in the weed community

Agnieszka Kubik-Komar, Mirosława Wesołowska-Janczarek

Department of Applied Mathematics, University of Agriculture, Akademicka 13, 20-950 Lublin, Poland, agnieszka.kubik@ar.lublin.pl, miroslawa.wesolowska@ar.lublin.pl

SUMMARY

The objective of the paper is to present a modification of a well-known model with the split plot design by adding new parameters and applying a hierarchical classification as well as a multivariate analysis of variance. The issue is illustrated by the data gathered in the course of a three-year experiment in which changes in a weed community under the tillage systems and doses of herbicides were studied. In the multivariate analysis of variance the weed species were treated as variables and the relative abundance of species as the studied feature. The added parameters describe the effects of the times of observations and the relevant interaction effects. The data were analyzed for each year separately. Herbicides were applied between the first and second observation of the quantity of weeds; thus there were no effects of the herbicide doses for the first time of observation in the first year. Due to this, appropriate restrictions were added to the model. Additionally, in the case of rejection of the null hypothesis the analysis was continued by applying the multiple comparison formula for some chosen sources of variation for the data from the last year of the experiment.

Tillage systems, as well as terms of observations and interaction of these factors, caused significant changes in the relative abundance of the weed community. Moreover, in 1999 and 2000 significant differences occurred in doses of herbicides. There were no significant changes observed in the weed community caused by the interaction of tillage systems, terms of observations and doses of herbicides.

Key words: herbicide doses, relative abundance index, Roy's test function, simultaneous multivariate multiple comparisons, tillage system.

1. Introduction

In most plant community experiments, dry weight or diversity indices are counted (Pawłowski and Wesołowski, 1982; Jędruszczak *et al.* 1997; 2000), which determines the application of a univariate analysis of variance.

In 1997-2000 in the Department of Soil Tillage and Plant Cultivation at the Agricultural University in Lublin, an experiment was conducted in which the influence of experimental factors on production and weed infestation of winter wheat cultivated in a short-time monoculture was analyzed. The numerical data from this experiment were analyzed by univariate standard methods and presented in a paper by Antoszek (2000).

Changes in the weed community with regard to species relationships and the use of the multivariate methods to estimate them (Morrison, 1990) seemed to be promising. In the paper a proposition of a multivariate model, which provides the possibility of estimating changes in the weed community under experimental factors by verifying general and particular hypotheses, is presented. An additional factor, time of observation, determined the modification of the well-known model for the split-plot design, under which the experiment was carried out. Not only was a new parameter added to the model, but also, since time of observation influenced the estimation of one of the experimental factors, the hierarchical classification was applied (Kuna-Broniowska, 2000). Additionally, for the data for the first year, some non-standard restrictions on the model parameters were used. The necessary modification of the split-plot design model and additional restrictions made it impossible to use the known analysis of variance formulas for three factorial designs.

The results of multiple comparisons provided detailed information about the species for which the studied feature changed significantly and about the experimental factor levels at which these changes appeared.

2. Material and methods

The experiment was carried out by a split-plot method, arranged in 4 blocks (r=4) in the 1997/1998-1999/2000 seasons. In the model, the tillage system (A_1 - conventional, A_2 - reduced with disk harrow, A_3 - reduced with cultivator, A_4 - zero) was the treatment randomized on the main plots, and doses of herbicides (B_1-100%, B_2-75%, B_3-50%, B_4-25%, B_5-0% of permissible dose) on the sub-plots. The changes in the weed community developing in winter wheat cultivated in a short-lived monoculture were analyzed. Weeds were counted three times a year: T_1 (in spring, before the application of herbicides), T_2 (about 15 days after the application) and T_3 (in summer before the wheat harvest) (Antoszek, 2000).

To avoid the problem of the non-uniform distribution of weeds, relative abundance (Ra) was calculated (Conn and Delapp, 1980) according to the formula

$$Ra = \frac{rd + rf}{2} \cdot 100\ \% , \qquad (2.1)$$

where the relative density (*rd*) was calculated as the number of individual occurrences for a given species within four samples of the subplot divided by the total number of weeds from these samples; relative frequency (*rf*) was calculated as the ratio of the number of samples in which the species was present to the number of samples with weed species per subplot.

In spite of the fact that the experiment was conducted in a split-plot design, an additional source of weed variances should have been taken into account, namely the time of observation. Moreover, herbicides were applied each year at the beginning of summer, so changes in the studied features caused by herbicide doses could be estimated only at particular times of observation, and this is why the mathematical model presented below includes hierarchical classifications of some parameters.

The application of herbicides after the first time of weed estimation also determined non-standard restrictions on the model parameters for the data from the first year of studies caused by the lack of herbicide dose effects (it was before the herbicides were used).

2.1. The model

Let us repeat the notation introduced above: A – main-plot treatment (tillage system; a=4), B – sub-plot treatment (dose of herbicides; b=5), T – additional treatment (time of observation; t=3).

A single observation for the h-th variable in the newly proposed model is described as a linear function

$$y_{ijksh} = \mu_h + \rho_{ih} + \alpha_{jh} + \tau_{kh} + (\alpha\tau)_{jkh} + e_{ijkh} + \beta_{s(k)h} + (\alpha\beta)_{js(k)h} + e_{ijksh}, \qquad (2.1.1)$$

where μ_h denotes general mean (h=1,...,p), ρ_{ih} – the effect of the i-th block (i=1,...,r), α_{jh} – the effect of the j-th level of factor A (j=1,...,a), τ_{kh} – the effect of the k-th level of factor T (k=1,...,t), $\beta_{s(k)h}$ – the effect of s-th level of factor B inside the k-th level of T (s=1,...,b), $(\alpha\tau)_{jkh}$, $(\alpha\beta)_{js(k)h}$ – relevant interaction effects, and e_{ijkh}, e_{ijksh} – random effects of experimental errors.

Taking into consideration p variables – species of weeds – a single observation is a row vector consisting of p values of the studied feature (Ra).

In the matrix notation, the model (2.1.1) can be written as follows:

$$\mathbf{Y}=\mathbf{X}_M\boldsymbol{\mu}+\mathbf{X}_R\boldsymbol{\rho}+\mathbf{X}_A\boldsymbol{\alpha}+\mathbf{X}_T\boldsymbol{\tau}+\mathbf{X}_{AT}(\boldsymbol{\alpha\tau})+\mathbf{X}_1\mathbf{e}_1+\mathbf{X}_{B(T)}\boldsymbol{\beta}+\mathbf{X}_{AB(T)}(\boldsymbol{\alpha\beta})+\mathbf{X}_2\mathbf{e}_1, \quad (2.1.2)$$

where N=rabt is the number of observations for each of p variables, $\mathbf{Y}$ is the Nxp matrix of observations, $\boldsymbol{\mu}$ is the p vector of general means, $\boldsymbol{\rho}$ is the rxp matrix of the fixed effects of blocks, $\boldsymbol{\alpha}$ is the axp matrix of the fixed effects of treatment A, $\boldsymbol{\tau}$ is the txp matrix of the fixed effects of treatment T, $\boldsymbol{\alpha\tau}$ is the atxp matrix of the interaction effects between A and T, $\mathbf{e}_1$ is the r(a+t+at)xp matrix of the random errors for the main plots, $\boldsymbol{\beta}$ is the btxp matrix of the fixed effects of treatment B, $\boldsymbol{\alpha\beta}$ is the atbxp matrix of the interaction effects between treatments A and B, and $\mathbf{e}_2$ is the Nxp matrix of the random errors for the sub-plots. The matrices

$$\mathbf{X}_M=\mathbf{1}_N,\ \mathbf{X}_R=\mathbf{I}_r\otimes\mathbf{1}_{atb},\ \mathbf{X}_A=\mathbf{1}_r\otimes\mathbf{I}_a\otimes\mathbf{1}_{tb},\ \mathbf{X}_T=\mathbf{1}_{ra}\otimes\mathbf{I}_t\otimes\mathbf{1}_b,\ \mathbf{X}_{AT}=\mathbf{1}_r\otimes\mathbf{I}_{at}\otimes\mathbf{1}_b,\ \mathbf{X}_{B(T)}=\mathbf{1}_{ra}\otimes\mathbf{I}_{tb},$$

$$\mathbf{X}_{AB(T)}=\mathbf{1}_r\otimes\mathbf{I}_{atb},\ \mathbf{X}_1=[\mathbf{I}_{ra}\otimes\mathbf{1}_{tb} \,\vdots\, \mathbf{I}_r\otimes\mathbf{1}_a\otimes\mathbf{I}_t\otimes\mathbf{1}_b \,\vdots\, \mathbf{I}_{rat}\otimes\mathbf{1}_b],\ \mathbf{X}_2=\mathbf{I}_N$$

are the relevant design matrices. $\otimes$ denotes Kroneker's product of matrices, $\mathbf{I}_k$ is the kxk identity matrix and $\mathbf{1}_k$ is the k vector of ones.

The multivariate linear model (2.1.2) can also be written in the following form

$$\mathbf{Y}=\mathbf{X}\boldsymbol{\Theta}+\mathbf{U}, \quad (2.1.3)$$

where $\mathbf{Y}$ is the Nxp matrix of observations, $\boldsymbol{\Theta}$ is the qxp matrix of the unknown parameters of the model, $\mathbf{X}=[\mathbf{X}_M \,\vdots\, \mathbf{X}_R \,\vdots\, \mathbf{X}_A \,\vdots\, \mathbf{X}_T \,\vdots\, \mathbf{X}_{AT} \,\vdots\, \mathbf{X}_{B(T)} \,\vdots\, \mathbf{X}_{AB(T)}]$ is the Nxq design matrix,

$$\mathbf{U}=[\mathbf{X}_1 \,\vdots\, \mathbf{X}_2]\begin{bmatrix}\mathbf{e}_1\\ \mathbf{e}_2\end{bmatrix}$$

is the Nxp matrix and q=1+r+a+t+at+bt+abt is the number of parameters in the model for each of p variables.

We assumed that: 1) the rows of experimental error matrices $\mathbf{e}_1$ and $\mathbf{e}_2$ are independent and normally distributed with the null mean vectors and $\boldsymbol{\Sigma}_1$, $\boldsymbol{\Sigma}_2$ the pxp covariance matrices, respectively, 2) the number of observations (N) for each of p variables is not less than p+q, 3) $\mathbf{Y}\sim\mathbf{N}_p(\mathbf{X}\boldsymbol{\Theta},\mathbf{V})$, where the covariance matrix $\mathbf{V}$ is equal to $(\mathbf{I}_{ra}\otimes\mathbf{J}_{tb}+\mathbf{I}_r\otimes\mathbf{J}_a\otimes\mathbf{I}_t\otimes\mathbf{J}_b+\mathbf{I}_{rat}\otimes\mathbf{J}_b)\otimes\boldsymbol{\Sigma}_1+\mathbf{I}_N\otimes\boldsymbol{\Sigma}_2$, and $\mathbf{J}_k$ is the kxk matrix of ones.

2.2. Estimators of the model parameters

The design matrix **X** is not of full rank, so in order to obtain the unique estimators the restrictions on the model parameters should be added. Here, for the h-th variable, a part of the restrictions takes a standard form

$$\sum_i \rho_{ih}=0,\quad \sum_j \alpha_{jh}=0,\ \sum_k \tau_{kh}=0,\ \underset{j}{\forall}\sum_k (\alpha\tau)_{jkh}=0,$$
$$\underset{k}{\forall}\sum_j (\alpha\tau)_{jkh}=0,\quad \underset{j,k}{\forall}\sum_s (\alpha\beta)_{js(k)h}=0,\quad \underset{s,k}{\forall}\sum_j (\alpha\beta)_{js(k)h}=0, \tag{2.2.1}$$

and since there is no effect of the B factor for the first level of T for the first year of study

$$\left.\begin{array}{l}\beta_{s(1)h}=0,\quad \underset{k\neq 1}{\forall}\sum_s \beta_{s(k)h}=0,\\ (\alpha\beta)_{s(1)h}=0,\quad \underset{j,k\neq 1}{\forall}\sum_s (\alpha\beta)_{js(k)h}=0,\quad \underset{s,k\neq 1}{\forall}\sum_j (\alpha\beta)_{js(k)h}=0\end{array}\right\}\ \text{for the first year} \tag{2.2.2}$$

$$\left.\underset{k}{\forall}\sum_s \beta_{s(k)h}=0,\quad \underset{j,k}{\forall}\sum_s (\alpha\beta)_{js(k)h}=0,\quad \underset{s,k}{\forall}\sum_j (\alpha\beta)_{js(k)h}=0\right\}\ \text{for the next years}$$

In the matrix notation the restrictions (2.2.1) and (2.2.2) can be written using the following formula

$$\mathbf{Z}=\mathrm{diag}(\mathbf{Z}_R,\mathbf{Z}_A,\mathbf{Z}_T,\mathbf{Z}_{AT},\mathbf{Z}_{B(T)},\mathbf{Z}_{AB(T)}) \tag{2.2.3}$$

Sub-block matrices in (2.2.3) take the following forms

$$\mathbf{Z}_R=\mathbf{1}'_r,\ \mathbf{Z}_A=\mathbf{1}'_a,\ \mathbf{Z}_T=\mathbf{1}'_t,\ \mathbf{Z}_{AT}=\begin{bmatrix}\mathbf{I}_a\otimes\mathbf{1}'_t\\ \mathbf{1}'_a\otimes\mathbf{I}_t\end{bmatrix}$$

$$\left.\begin{array}{l}\mathbf{Z}_{B(T)}=\begin{bmatrix}\mathbf{I}_b & \mathbf{0}_{b\times(t-1)b}\\ \mathbf{0}_{(t-1)\times b} & \mathbf{I}_{t-1}\otimes\mathbf{1}'_b\end{bmatrix}\\ \mathbf{Z}_{AB(T)}=\begin{bmatrix}\mathbf{I}_a\otimes\left[\mathbf{I}_b\ \vdots\ \mathbf{0}_{b\times b(t-1)}\right]\\ \mathbf{I}_a\otimes\left[\mathbf{0}_{(t-1)\times b}\ \vdots\ \mathbf{I}_{t-1}\otimes\mathbf{1}'_b\right]\\ \mathbf{1}'_a\otimes\left[\mathbf{0}_{b(t-1)\times b}\ \vdots\ \mathbf{I}_{b(t-1)}\right]\end{bmatrix}\end{array}\right\}\ \text{for the first year} \tag{2.2.4}$$

$$\left.\mathbf{Z}_{B(T)}=\left[\mathbf{I}_t\otimes\mathbf{1}'_b\right],\ \mathbf{Z}_{AB(T)}=\begin{bmatrix}\mathbf{I}_{at}\otimes\mathbf{1}'_b\\ \mathbf{1}'_{at}\otimes\mathbf{I}_b\end{bmatrix}\right\}\ \text{for the next years}$$

where $\mathbf{0}_{ixj}$ denotes the ixj matrix of nulls.

The location of sub-blocks (2.2.4) in the **Z** matrix enables us to choose relevant parameters of the $\mathbf{\Theta}$ matrix. The other elements of **Z** are nulls. Thus the restrictions (2.2.1) and (2.2.2) for p variables can be written as: $\mathbf{Z\Theta} = \mathbf{0}$. Since the rank of $[X^{'} \vdots Z^{'}]^{'}$ is equal to q, the estimator of the unknown parameters can be calculated according to the formula

$$\hat{\mathbf{\Theta}} = \mathbf{GX'Y}, \tag{2.2.5}$$

where $\mathbf{G}=(\mathbf{X}'\mathbf{X}+\mathbf{Z}'\mathbf{Z})^{-1}$ is one of the generalized inverse matrices of $\mathbf{X}'\mathbf{X}$.

2.3. Hypotheses and tests

In order to find out which of the experimental factors effected significant changes in the studied feature for each variable, the null hypothesis H_0: $\mathbf{K\Theta}=\mathbf{0}$ should be verified against the alternative one H_1: $\mathbf{K\Theta}\neq\mathbf{0}$ under the assumption that rank(**K**) $\leq$ min(rank(**X**);d) for the dxq matrix **K**.

In this model verification of hypothesis H_0 is reduced to verifying the following hypotheses related to the sources of variation: H_{0A} – assuming equality of the A treatment effects (α) i.e. assuming the equality of the vectors of the tillage systems effects, H_{0T} – assuming equality of the T treatment effects (τ) i.e. assuming the equality of the vectors of the terms of observations effects, $H_{0B(T)}$ – assuming equality of the B(T) treatment effects (β) i.e. assuming the equality of the vectors of the doses of herbicides effects in relevant times of observation, H_{0AT} – assuming equality of the AxT interaction effects ($\alpha\tau$) i.e. assuming the equality of the vectors of interaction effects between the tillage systems and times of observation, $H_{0AB(T)}$ – assuming equality of the AxB interaction effects ($\alpha\beta$) i.e. assuming the equality of the vectors of interaction effects between the tillage systems and the doses of herbicides observed in relevant times.

A simplified table of the analysis of variance (Table 1) for the model (2.1.2) is presented as follows

Table 1. Table of analysis of variance for the model (2.1.2)

Source of variation	Degrees of freedom	Sum of products matrices
Blocks	r-1	H_R
A factor	a-1	H_A
T factor	t-1	H_T
AxT interaction	(a-1)(t-1)	H_{AT}
Error I	(at-1)(r-1)	E_1
B(T) factor	t(b-1)	$H_{B(T)}$
AxB(T) interaction	t(a-1)(b-1)	$H_{AB(T)}$
Error II	at(r-1)(b-1)	E_2
Total	ratb-1	

The sum of product matrices connected with the relevant null hypothesis can be obtained in the following way

$$\begin{aligned}
\mathbf{H}_R &= \hat{\mathbf{\Theta}}'\mathbf{K}_R'[\mathbf{K}_R\mathbf{G}\mathbf{K}_R']^{-1}\mathbf{K}_R\hat{\mathbf{\Theta}},\\
\mathbf{H}_A &= \hat{\mathbf{\Theta}}'\mathbf{K}_A'[\mathbf{K}_A\mathbf{G}\mathbf{K}_A']^{-1}\mathbf{K}_A\hat{\mathbf{\Theta}},\\
\mathbf{H}_T &= \hat{\mathbf{\Theta}}'\mathbf{K}_T'[\mathbf{K}_T\mathbf{G}\mathbf{K}_T']^{-1}\mathbf{K}_T\hat{\mathbf{\Theta}}\\
\mathbf{H}_{AT} &= \hat{\mathbf{\Theta}}'\mathbf{K}_{AT}'[\mathbf{K}_{AT}\mathbf{G}\mathbf{K}_{AT}']^{-1}\mathbf{K}_{AT}\hat{\mathbf{\Theta}}\\
\mathbf{H}_{B(T)} &= \hat{\mathbf{\Theta}}'\mathbf{K}_{B(T)}'[\mathbf{K}_{B(T)}\mathbf{G}\mathbf{K}_{B(T)}']^{-1}\mathbf{K}_{B(T)}\hat{\mathbf{\Theta}}\\
\mathbf{H}_{AB(T)} &= \hat{\mathbf{\Theta}}'\mathbf{K}_{AB(T)}'[\mathbf{K}_{AB(T)}\mathbf{G}\mathbf{K}_{AB(T)}']^{-1}\mathbf{K}_{AB(T)}\hat{\mathbf{\Theta}}
\end{aligned} \tag{2.3.1}$$

where:

$$\begin{aligned}
\mathbf{K}_R &= [\mathbf{0}_{(r-1)x1} \vdots \mathbf{I}_{r-1} \vdots -\mathbf{1}_{r-1} \vdots \mathbf{0}_{(r-1)x(q-r-1)}],\\
\mathbf{K}_A &= [\mathbf{0}_{(a-1)x(1+r)} \vdots \mathbf{I}_{a-1} \vdots -\mathbf{1}_{a-1} \vdots \mathbf{0}_{(a-1)x(q-a-r-1)}],\\
\mathbf{K}_T &= [\mathbf{0}_{(t-1)x(1+r+a)} \vdots \mathbf{I}_{t-1} \vdots -\mathbf{1}_{t-1} \vdots \mathbf{0}_{(t-1)x(q-t-a-r-1)}],\\
\mathbf{K}_{AT} &= [\mathbf{0}_{(at-1)x(1+r+a+t)} \vdots \mathbf{I}_{at-1} \vdots -\mathbf{1}_{at-1} \vdots \mathbf{0}_{(at-1)xbt(a+1)}],\\
\mathbf{K}_{B(T)} &= [\mathbf{0}_{t(b-1)x(q-bt(a+1))} \vdots \mathbf{I}_t \otimes [\mathbf{I}_{b-1} \vdots -\mathbf{1}_{b-1}] \vdots \mathbf{0}_{t(b-1)xabt}],\\
\mathbf{K}_{AB(T)} &= [\mathbf{0}_{at(b-1)x(q-abt)} \vdots \mathbf{I}_{at} \otimes [\mathbf{I}_{b-1} \vdots -\mathbf{1}_{b-1}]]
\end{aligned} \tag{2.3.2}$$

are contrast matrices for the null hypotheses connected with the sources of variation. The rows of these matrices contain 1 and –1, corresponding to the comparisons of the effects of experimental factors. The other elements of the rows are equal to zero.

The sum of product matrices for errors was calculated by applying projection operators (Mikos, 1973; Kuczyński 1976)

$$\mathbf{E}_1 = \mathbf{Y'P}_1\mathbf{Y}, \quad \mathbf{E}_2 = \mathbf{Y'P}_2\mathbf{Y} \tag{2.3.3}$$

where $\mathbf{P}_1$: $\boldsymbol{R}^N \rightarrow \boldsymbol{R}(\mathbf{X}_1)$ takes the following form

$$\mathbf{P}_1 = \frac{1}{tb}(\mathbf{I}_r - \frac{1}{r}\mathbf{J}_r) \otimes (\mathbf{I}_a - \frac{1}{a}\mathbf{J}_a) \otimes \mathbf{J}_{tb} + \frac{1}{ab}(\mathbf{I}_r - \frac{1}{r}\mathbf{J}_r) \otimes \mathbf{J}_a \otimes (\mathbf{I}_t - \frac{1}{t}\mathbf{J}_t) \otimes \mathbf{J}_b +$$
$$+ \frac{1}{b}(\mathbf{I}_r - \frac{1}{r}\mathbf{J}_r) \otimes (\mathbf{I}_a - \frac{1}{a}\mathbf{J}_a) \otimes (\mathbf{I}_t - \frac{1}{t}\mathbf{J}_t) \otimes \mathbf{J}_b, \tag{2.3.4}$$

and $\mathbf{P}_2$: $\boldsymbol{R}^N \rightarrow \boldsymbol{R}(\mathbf{X}_2)$ is obtained in the following way

$$\mathbf{P}_2 = \frac{1}{at}(\mathbf{I}_r - \frac{1}{r}\mathbf{J}_r) \otimes \mathbf{J}_{at} \otimes (\mathbf{I}_b - \frac{1}{b}\mathbf{J}_b) + \frac{1}{t}(\mathbf{I}_r - \frac{1}{r}\mathbf{J}_r) \otimes (\mathbf{I}_a - \frac{1}{a}\mathbf{J}_a) \otimes \mathbf{J}_t \otimes (\mathbf{I}_b - \frac{1}{b}\mathbf{J}_b) +$$
$$+ \frac{1}{a}(\mathbf{I}_r - \frac{1}{r}\mathbf{J}_r) \otimes \mathbf{J}_a \otimes (\mathbf{I}_t - \frac{1}{t}\mathbf{J}_t) \otimes (\mathbf{I}_b - \frac{1}{b}\mathbf{J}_b) +$$
$$+ (\mathbf{I}_r - \frac{1}{r}\mathbf{J}_r) \otimes (\mathbf{I}_a - \frac{1}{a}\mathbf{J}_a) \otimes (\mathbf{I}_t - \frac{1}{t}\mathbf{J}_t) \otimes (\mathbf{I}_b - \frac{1}{b}\mathbf{J}_b) \tag{2.3.5}$$

The matrices $\mathbf{P}_1$ and $\mathbf{P}_2$ are projection operators of N-dimensional Euclid's space $\boldsymbol{R}^N$ on the column space of matrix $\mathbf{X}_1$ and column space of matrix $\mathbf{X}_2$, respectively, which was written in the short form $\mathbf{P}_1$: $\boldsymbol{R}^N \rightarrow \boldsymbol{R}(\mathbf{X}_1)$ and $\mathbf{P}_2$: $\boldsymbol{R}^N \rightarrow \boldsymbol{R}(\mathbf{X}_2)$.

Roy's test function W_s is used to verify the null hypothesis. This test is based on the maximum eigenvalue λ_s of the matrix type $\mathbf{HE}^{-1}$

$$W_s = \frac{\lambda_s}{1 + \lambda_s}. \tag{2.3.6}$$

The obtained value should be compared to the 100α-percentage critical value $W_{\alpha;s,m,n}$ from Pillai's tables or Heck's nomograms. The parameters of the W_s distribution are calculated according to the formulas

$$s = \min(r(\mathbf{K}), p), \quad m = \frac{|r(\mathbf{K}) - p| - 1}{2}, \quad n = \frac{N - r(\mathbf{X}) - p - 1}{2}, \tag{2.3.7}$$

where $r(\mathbf{K})$ is the rank of the matrix $\mathbf{K}$.

In order to verify the null hypotheses connected with individual sources of variation the maximum eigenvalues of the following matrices should be calculated: $\mathbf{H}_A\mathbf{E}_1^{-1}$, $\mathbf{H}_T\mathbf{E}_1^{-1}$, $\mathbf{H}_{AT}\mathbf{E}_1^{-1}$, $\mathbf{HB}_{(T)}\mathbf{E}_2^{-1}$, $\mathbf{H}_{AB(T)}\mathbf{E}_2^{-1}$.

When the null hypothesis is rejected it is possible to obtain additional information using multiple comparisons by applying Roy's confidence bounds formula (Kuna-Broniowska, 2000)

$$\sum_{h=1}^{p}\sum_{i=1}^{q_i} a_h k_i \bar{y}_{ih} - \sqrt{\frac{W_\alpha}{1-W_\alpha}\sum_{i=1}^{q_i}\frac{k_i^2}{n_{ih}}\mathbf{a}'\mathbf{E}_m\mathbf{a}} \le \sum_{h=1}^{p}\sum_{i=1}^{q_i} a_h k_i \theta_{hi} \le \sum_{h=1}^{p}\sum_{i=1}^{q_i} a_h k_i \bar{y}_{ih} + \sqrt{\frac{W_\alpha}{1-W_\alpha}\sum_{i=1}^{q_i}\frac{k_i^2}{n_{ih}}\mathbf{a}'\mathbf{E}_m\mathbf{a}}, \quad (2.3.8)$$

where θ_{hi} is the effect of the studied treatment for the h-th variable, W_α is the 100α-percentage critical value from Pillai's tables or Heck's nomograms, $\bar{y}_{ih}$ is the mean value of the studied feature in treatment subclasses for the h-th variable, n_{ih} denotes the number of observations on the basis of which the mean value $\bar{y}_{ih}$ was counted, a_h is a coordinate of vector **a** choosing the h-th variable, k_i is a coordinate of contrast coefficient vector **k** choosing the parameter θ_{hi} from the matrix $\mathbf{\Theta}$, $\mathbf{E}_m$ denotes sums and products of the experimental errors matrix (m=1,2). Since the multivariate comparisons were made within the particular p variables, the vector **a** for the h-th variable is a p-dimensional vector of zeros with one in the h-th position.

In the paper we did not determine confidence bounds, but used part of the formula (2.3.8) to obtain the lowest significant differences denoted by

$$\mathrm{L}=\sqrt{\frac{W_\alpha}{1-W_\alpha}\sum_{i=1}^{q_i}\frac{k_i^2}{n_{ih}}\mathbf{a}'\mathbf{E}_m\mathbf{a}} \quad (2.3.9)$$

which enabled us to find out if there were significant differences between the Ra mean values of weed species related to experimental factor levels.

3. Results

The data were analyzed separately for each year of the experiment. In order to make an assumption about the truthfulness of the normal distribution of the data, a transformation was made, according to the formula

$$z = \arcsin\sqrt{\frac{Ra}{100\%}} \quad (3.1)$$

This is the most common transformation for percentage (Trętowski and Wójcik, 1988; Derksen, 1995). In spite of this, there were still a lot of species which had to be restricted because of the assumptions of the multivariate analysis of variance (Morrison, 1990; Krzyśko, 2000). Therefore, among the 46 weed species which infested the winter wheat canopy, only 31 species from the first year of the investigation and 20 species per year from 1999 and 2000 were analyzed (Table 2). The calculations were made using the *Maple* application and our own implementation of the formulas presented above.

Table 2. List of weed species and mean value of the index Ra in 1998, 1999 and 2000

Latin name of species	Bayer code	1998	1999	2000
Apera spica-venti[2]	APESV	18.12	28.81	38.84
Galium aparine[2]	GALAP	9.85	9.53	5.61
Matricaria maritima L. subsp. Inodora[2]	MATIN	9.42	10.46	8.71
Stellaria media[2]	STEME	14.21	17.91	8.58
Galeopsis tetrahit[2]	GAETE	6.40	1.75	0.65
Galinsoga parviflora[1]	GASPA	0.33	0.01	0.02
Cirsium arvense[2]	CIRAR	2.63	1.16	0.73
Equisetum arvense[2]	EQUAR	0.62	0.74	0.51
Polygonum lapathifolium[1]	POLLL	1.66	0.08	0.01
Chenopodium album[2]	CHEAL	1.89	0.57	0.38
Taraxacum officinale[1]	TAROF	0.65	0.07	0.06
Gnaphalium uliginosum[1]	GNAUL	0.26	0.03	0.01
Veronica arvensis[2]	VERAR	6.62	3.07	5.16
Myosotis arvensis[2]	MYOAR	5.80	8.81	9.09
Fallopia convonvulus[2]	POLCO	0.98	0.24	0.04
Plantago intermedia[1]	PLAPA	0.48	0.00	0.01
Capsella bursa-pastoris[2]	CAPBP	7.62	3.45	4.01
Lapsana communis[2]	LAPCO	1.09	1.28	2.28
Viola arvensis[2]	VIOAR	1.05	2.63	5.61
Poa annua[1]	POAAN	0.42	0.03	0.03
Lamium amplexicaule[2]	LAMAM	5.64	6.05	4.77
Chamomilla suaveolens[1]	MATMT	0.17	0.19	0.20
Lamium purpureum[2]	LAMPU	1.68	1.79	1.49
Agropyron repens[2]	AGRRE	0.16	0.13	0.26

Echinochloa crus-galli	ECHCG	0.06	0.04	0.01
Galinsoga ciliata	GASQU	0.08	0.01	0.01
Convolvulus arvensis[1]	CONAR	0.03	0.03	0.04
Veronica persica[2]	VERPE	1.15	1.65	2.36
Papaver rhoeas[1]	PAPRH	0.06	0.06	0.04
Melandrium album	MELAL	0.02	0.00	0.00
Vicia hirsuta[1]	VICHI	0.10	0.00	0.08
Stachys palustris	STAPA	0.02	0.00	0.00
Geranium pusillum[2]	GERPU	0.07	0.07	0.25
Polygonum aviculare	POLAV	0.03	0.05	0.01
Sonchus oleraceus	ERICA	0.01	0.00	0.00
Conyza canadensis	SONOL	0.01	0.00	0.00
Spergula arvensis[1]	SPRAR	0.21	0.02	0.00
Thlaspi arvense[2]	THLAR	0.18	0.01	0.28
Anagallis arvensis	ARBTH	0.01	0.00	0.00
Scleranthus annuus	SINAR	0.03	0.00	0.00
Polygonum persicaria	MYSMI	0.04	0.00	0.00
Anthemis arvensis	ANTAR	0.02	0.00	0.00
Myosurus minimus	POLPE	0.04	0.03	0.00
Sinapis arvensis	SCRAN	0.01	0.00	0.00
Arabidopsis thaliana	ANGAR	0.04	0.00	0.00
Fumaria officinalis	FUMOF	0.01	0.00	0.01

[1]species included in the analysis for the data from 1998
[2]species included in the analysis for the data from each year of study

Each year significant changes in relative abundance of the weed community were influenced by the tillage systems, times of observation and the interaction of these two factors. Moreover, in 1999 and 2000 significant differences occurred in doses of herbicides. There were no significant changes in the weed community caused by the interaction of the tillage systems, terms of observations and doses of herbicides (Table 3).

Table 3. Results of multivariate analysis of variance for the experimental data

Sources of variation	1998		1999		2000	
	W_S	$W_{0.05}$	W_S	$W_{0.05}$	W_S	$W_{0.05}$
Tillage systems A	0.837*	0.36	0.710*	0.22	0.571*	0.22
Terms of observations T	0.987*	0.30	0.912*	0.20	0.952*	0.20
Interaction AxT	0.982*	0.38	0.688*	0.29	0.745*	0.29
Doses of herbicides H(T)	0.370	0.38	0.653*	0.32	0.736*	0.32
Interaction AxH(T)	0.492	0.51	0.320	0.36	0.335	0.36

*significant changes in the mean Ra value at the level α=0.05

Since the experimental factors as well as their interactions caused significant changes in the weed community every year, a large number of multiple comparisons could be considered. However, the authors decided to present only some of the results for the last season data, when the experimental factor activity was settled, as an example of the application of formula (2.3.9). As mentioned above, only the smallest significant differences were calculated and used for comparison. In this case confidence bounds, as counted on the transformed data, would not have any practical interpretation.

The results of multivariate comparisons showed that the tillage systems differentiated the mean value of relative abundance for the following species: CAPBP, STEME, CIRAR, MYOAR, APESV, LAPCO and VIOAR (Table 4). The studied feature for the majority of the species significantly differed between the system with disk harrow (A_2) and direct sowing (A_4). The were no species for which the relative abundance significantly differed between the classic tillage system (A_1) and the reduced system with disc harrow (A_2).

Table 4. Multivariate comparisons of mean Ra between tillage systems (2000)

Weed species	L	$\|A_1-A_2\|$	$\|A_1-A_3\|$	$\|A_1-A_4\|$	$\|A_2-A_3\|$	$\|A_2-A_4\|$	$\|A_3-A_4\|$
APESV	0.0411	0.0364	0.0072	0.0100	0.0292	0.0464*	0.0172
GALAP	0.0441	0.0054	0.0022	0.0315	0.0075	0.0262	0.0337
MATIN	0.0498	0.0326	0.0323	0.0275	0.0003	0.0051	0.0048
STEME	0.0507	0.0459	0.0660*	0.0931*	0.0201	0.0472	0.0271
GAETE	0.0447	0.0062	0.0075	0.0177	0.0013	0.0116	0.0102
CIRAR	0.0416	0.0254	0.0236	0.0293	0.0018	0.0547*	0.0529*
EQUAR	0.0354	0.0052	0.0047	0.0137	0.0099	0.0085	0.0185
CHEAL	0.0270	0.0020	0.0041	0.0103	0.0021	0.0123	0.0144
VERAR	0.0458	0.0029	0.0159	0.0265	0.0188	0.0293	0.0105

MYOAR	0.0420	0.0041	0.0542*	0.0289	0.0501*	0.0248	0.0253
POLCO	0.0076	0.0075	0.0025	0.0000	0.0050	0.0075	0.0025
CAPBP	0.0294	0.0157	0.0325*	0.0171	0.0168	0.0328*	0.0496*
LAPCO	0.0451	0.0031	0.0263	0.0443	0.0294	0.0474*	0.0179
VIOAR	0.0615	0.0604	0.0088	0.0188	0.0517	0.0792*	0.0275
LAMAM	0.0269	0.0059	0.0124	0.0037	0.0183	0.0022	0.0161
LAMPU	0.0471	0.0160	0.0001	0.0313	0.0160	0.0153	0.0313
AGRRE	0.0270	0.0172	0.0214	0.0025	0.0043	0.0147	0.0190
VERPE	0.0431	0.0266	0.0287	0.0103	0.0021	0.0163	0.0184
GERPU	0.0251	0.0088	0.0094	0.0218	0.0007	0.0130	0.0123
THLAR	0.0286	0.0115	0.0123	0.0024	0.0008	0.0091	0.0099

* significant differences at the level $\alpha=0.05$

Regarding the times of observations we can conclude that the mean Ra value differed significantly for each of the studied species, especially between the first and third time of weeds estimation (Table 5).

Table 5. Multivariate comparisons of mean Ra between time of observations (2000)

Species	L	$\lvert T_1-T_2\rvert$	$\lvert T_1-T_3\rvert$	$\lvert T_2-T_3\rvert$
APESV	0.0335	0.0723*	0.1254*	0.0531*
GALAP	0.0360	0.0210	0.0209	0.0419*
MATIN	0.0406	0.0810*	0.0044	0.0854*
STEME	0.0413	0.0425*	0.1913*	0.1488*
GAETE	0.0365	0.0706*	0.0308	0.0398*
CIRAR	0.0339	0.0230	0.0465*	0.0236
EQUAR	0.0289	0.0218	0.0547*	0.0330*
CHEAL	0.0220	0.0611*	0.0555*	0.0056
VERAR	0.0373	0.0153	0.1142*	0.0990*
MYOAR	0.0343	0.0376*	0.0053	0.0429*
POLCO	0.0062	0.0000	0.0075*	0.0075*
CAPBP	0.0240	0.1094*	0.1476*	0.0382*
LAPCO	0.0367	0.0297	0.0462*	0.0165

VIOAR	0.0502	0.0455	0.1014*	0.0558*
LAMAM	0.0219	0.0254*	0.2505*	0.2759*
LAMPU	0.0384	0.0617*	0.1487*	0.0870*
AGRRE	0.0220	0.0000	0.0308*	0.0308*
VERPE	0.0351	0.0305	0.1232*	0.1537*
GERPU	0.0205	0.0227*	0.0164	0.0063
THLAR	0.0233	0.0428*	0.0389*	0.0039

* significant differences at the level $\alpha=0.05$

The effect of herbicide doses was compared in the relevant time. In the spring (T_1) there were no significant changes in the weed community caused by the doses of herbicides. After the herbicides' application (T_2) the studied feature of only two species – MATIN and CAPBP – changed significantly. The highest significant differences for both species were recorded between the plots with 75% of the recommended dose and the ones without herbicides. In addition, the mean Ra value of *Capsella bursa-pastoris* differed significantly also between 100% as well as 50% of permissible dose and no-herbicide plots. For the last time of weeds observation (T_3), among all weeds CAPBP was the only species for which the relative abundance differed significantly. This variety was noticed between the direct sowing and the rest of the herbicide plots. Due to their large size the tables comparing herbicide plots at successive time points are not presented in the paper.

4. Conclusions

A new form of the model (2.1.2) with restrictions $\mathbf{Z\Theta} = \mathbf{0}$ for the matrix $\mathbf{Z}$ defined by the formula (2.2.3) enables us to use the multivariate analysis of variance in order to analyze the experimental data.

The analysis of variance for the model (2.1.2) yields conclusions comparable with practice – the changes observed in the weed community under experimental factors.

The model (2.1.2) enables us to analyze the data from experiments conducted in a similar way.

In the case of rejection of the null hypothesis it is possible to obtain more detailed information by using the lowest significant differences (2.3.9).

REFERENCES

Antoszek R. (2002). The impact of tillage system and herbicide doses on the productivity and features of weeds of winter wheat cultivated in a short-lived monoculture. PhD thesis. University of Agriculture, Lublin (in Polish).

Conn J.S., Delapp J.A. (1980). Weed species shifts with increasing field age in Alaska. Weed Science 31, 520-524.

Derksen D.A., Thomas A.G., Lafond G.P., Loeppky H.A., Swanton C.J. (1995). Impact of post emergence herbicides on weed community diversity within conservation-tillage systems. Weed Research 35, 311-320.

Derksen D.A., Thomas A.G., Lafond G.P., Loeppky H.A., Swanton C.J. (1993). Impact of Agronomic Practices on Weed Communities: Tillage Systems. Weed Science 41, 409-417.

Jędruszczak M., Bujak K., Wesołowski M. (1997). The impact of tillage systems on weed community on loessial soil in the region of Lublin. Bibl. Fragm. Agron. Proceedings 14th ISTRO Conference July 27 – August 1, Puławy-Poland, 2A, 299-302.

Jędruszczak M., Wesołowski M., Antoszek R. (2000). The effect of tillage practices on weeds in winter wheat grown in short-lasted cereal cropping system. ISTRO – 2000 Conf., Forth Worth, Texas 2-7 July, 28.

Kuczyński M. (1976). An estimation of parameters and verification of hypotheses in a covariance analysis for split-plot design. The 6th Methodological Colloquium of Agro-Biometry, 102-119 (in Polish).

Kuna-Broniowska I. (2000). Multiple comparisons of the object means with control for the hierarchical classification in the split plot design. Fragmenta Agronomica 7, 97-110 (in Polish)

Kuna-Broniowska I. (2000). Multivariate analysis of hierarchical classification in split plot design with respect to the control. Int. Agrophysic 14, 297-305.

Mikos H.(1973). Projection operators in analysis of variance. The 3rd Methodological Colloquium of Agro-Biometry, Warsaw, pp.78-142 (in Polish).

Morrison D. F. (1990). Multivariate statistical analysis. PWN Warsaw (in Polish).

Pawłowski F., Wesołowski M. (1982) Numerical force and some biological features of Apera spica-veenti L./P.B. in monoculture of wither wheat. Ann. Univ. Marie Curie-Skłodowska, sec. E, vol. XXXVII, 1, pp.1-8 (in Polish).

Trętowski J., Wójcik A. R. (1988). The methodology of agricultural experiments, Agricultural and Pedagogical Collage, Siedlce (in Polish).