Colloquium Biometricum 39
2009, 91–100

APPLICATION OF VARIANCE ANALYSIS FOR SPLIT-PLOT WITH REPEATED MEASURES DESIGN IN ESTIMATION OF *APERA SPICA VENTI* CHANGES UNDER IMPACT OF TILLAGE SYSTEMS AND HERBICIDES DOSES

Agnieszka Kubik-Komar, Izabela Kuna-Broniowska

Department of Applied Mathematics and Computer Science
University of Life Science
Akademicka 13, 20-950 Lublin
agnieszka.kubik@up.lublin.pl
izabela.kuna@up.lublin.pl

Summary

In this paper the analysis of variance for fixed split-plot with repeated measures model is presented. The method based on univariate and multivariate analysis of orthonormal contrast, making calculations independent of the covariance matrix structure, was chosen. In this way the impact of tillage systems as well as doses of herbicides on relative abundance (Ra) Apera spica venti evaluated three times a year was estimated. Average values of studied feature differed significantly under experimental factors, time of weeds evaluation and, in addition, doses of herbicides and time interaction. The time trends analysis revealed the significant differences of mean Ra in every three time and significant changes of linear trends between herbicide subplots.

Key words and phrases: contrast analysis, herbicide doses, relative abundance, repeated measures design, split-plot design, tillage systems

Classification AMS 2000: 62-07

1. Introduction

Repeated measures can be definite as measures over time or space on the same object or experimental unit (Moser *et al.*, 1990), so these measures are correlated.

The analysis of repeated measures is the analysis of the impact of repeated measures factor and factors applied in experiment as well as their interaction on studied feature. The levels of experimental factors are distributed at random on the experimental units but the factor of repeated measures is fixed in advance. In most cases the lack of random nature of repeated measures effects that the assumptions needed to obtain correct results in ANOVA are not held.

There are some methods of analysis this type of design. The data analysis on each level of repeated measures factor separately is the less effective one. We lose too much information about the changes in time or space of studied feature. However this method might be the initial part of data analysis.

The application of analysis of variance for split-plot design, where time/space is (confounded with experimental factors) a split-plot factor (Linnel Nemec, 1996). However, the analysis of this design is correct when the assumptions about covariance matrix are held. There is a compound symmetry condition assuming constant variances along the diagonal or less restrictive form of this condition called sphericity, which refers to the equality of variances of the differences between time/space factor levels (Fidel, 1998).

The next method of the analysis, in which the structure of covariance matrix is of no importance, is multivariate analysis with application of contrasts analysis. The repeated measures are treated as elements of multivariate vector of observations, and experimental data are transformed by orthogonal contrasts before MANOVA is used. The detail description of this method with the examples of its application for the randomized block and split-plot experiments are presented in the Gumpertz and Brownie (1993) paper.

The mixed model approach to this subject is preferred in the recent papers (Li *et al.*, 2004, Blouin *et al*. 2004). On the contrary to the method described above, the structure of covariance matrix is very important here. In this case two parts of the analysis might be distinguished i. e. the modeling of variance matrix structure and analyzing the trends by estimation and comparing of mean values (Littell *et al.*, 1998).

In this paper a multivariate approach was used to analyze experimental data. It was connected with the character of experiment, in which the relations between results in different time were disrupted by the application of one treatment between first two dates of measures.

The data were collected during field experiment carried out in split-plot design in 1997–2000, the goal of which was the estimation of the yield of winter wheat grown in short-term monoculture. In addition the crop weed infestation was studied and for the illustration of the chosen method the dominant of weed community, *Apera spica venti*, was selected.

2. Material and Methods

The experiment, the part of which results were used in this paper, was carried out by the Department of Soil Tillage and Plant Cultivation, University Life Sciences in Lublin in 1997–2000, and investigated the effect of the tillage system (A) and herbicide doses (B) on yielding and weeds infestation of winter wheat grown in short-term (3-year) monoculture. The studies lasted three years but in this paper the second year results were only used. Then the effect of experimental factors was stated and the effect of monoculture not as strong as in the last year of the experiment. Thus the conditions for studying the influence of treatments on dominant species were optimal.

The experiment was conducted in split-plot design with four replication and with four tillage system (A_1 – conventional, A_2 – reduced with disk harrow, A_3 – reduced with cultivator, A_4 – direct sowing) as 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 statistical analysis was focused on weed infestation evaluated three times: T_1 – before herbicides application (30^{th} – 31st of March), T_2 – about 15 days after the last herbicide application (11^{th} of May) and T_3 – before winter wheat harvest (15^{th}–16^{th} of July).

Relative abundance (Ra) of chosen species was the studied feature, which values were calculated according to the formulae:

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

where *rd* is a relative density calculated as the number of individual occurencies for a given species within four samples of the subplot divided by the total number of weeds from these samples and *rf* – relative frequency of occurrence the chosen species in weed community calculated as a proportion of the number of samples in which the species was present to the number of samples with weed species per subplot (Derksen *et al.*, 1993).

The single observation for split-plot with repeated measures might be described by the following linear function:

$$\begin{aligned} y_{ijkl} &= \mu + \rho_i + \alpha_j + \varepsilon_{ij} + \beta_k + (\alpha\beta)_{jk} + \delta_{ijk} + \tau_l + \\ &+ (\rho\tau)_{il} + (\alpha\tau)_{jl} + \theta_{ijl} + (\beta\tau)_{kl} + (\alpha\beta\tau)_{jkl} + \xi_{ijkl} \end{aligned} \tag{2.2}$$

where μ – general mean of studied feature, ρ_i – effect of ith block (i=1,...,r), α_j – effect of jth tillage system (j=1,...,a), β_k – effect of kth herbicide dose (k=1,...,b), τ_l – effect of lth date of weeds estimation (l=1,...,t), $(\rho\tau)_{il}, (\alpha\tau)_{jl}, (\beta\tau)_{kl}, (\alpha\beta\tau)_{jkl}$ – interaction effects, ε_{ij}, δ_{ijk} – errors of experimental factors, θ_{ijl}, ξ_{ijkl} – errors connected with repeated measures.

The effects of random errors for main plots (tillage system) and subplots (herbicide doses) are uncorrelated and normal distributed with mean value equal to 0 and variances equal to σ_1^2 and σ_2^2, respectively. The errors of repeated measures on the same main plot as well as on the same subplot are correlated:

$$\operatorname{cov}(\theta_{ijl}, \theta_{i'j'l'}) = \begin{cases} 0 & dla \quad ij \neq i'j' \\ \sigma_{\theta ll'} & dla \quad ij = i'j' \end{cases} \qquad \operatorname{cov}(\xi_{ijkl}, \xi_{i'j'k'l'}) = \begin{cases} 0 & dla \quad ijk \neq i'j'k' \\ \sigma_{\xi ll'} & dla \quad ijk = i'j'k' \end{cases}. \tag{2.3}$$

In vector notation the formulae (2.2) might be described as follows:

$$\mathbf{y}_{ijk} = \boldsymbol{\mu} + \boldsymbol{\rho}_i + \boldsymbol{\alpha}_j + \boldsymbol{\varepsilon}_{ij} + \boldsymbol{\beta}_k + (\boldsymbol{\alpha\beta})_{jk} + \boldsymbol{\delta}_{ijk} \tag{2.4}$$

where the vector of observations and factor vectors are presented in the following form:

$$\begin{aligned} &\mathbf{y}_{ijk} = \begin{bmatrix} y_{ijk1} & y_{ijk2} & \ldots & y_{ijkt} \end{bmatrix} \\ &\boldsymbol{\mu} = \begin{bmatrix} \mu + \tau_1 & \mu + \tau_2 & \ldots & \mu + \tau_t \end{bmatrix} \\ &\boldsymbol{\rho}_i = \begin{bmatrix} \rho_i + (\rho\tau)_{i1} & \rho_i + (\rho\tau)_{i2} & \ldots & \rho_i + (\rho\tau)_{it} \end{bmatrix} \\ &\ldots \\ &\boldsymbol{\varepsilon}_{ij} = \begin{bmatrix} \varepsilon_{ij} + \theta_{ij1} & \varepsilon_{ij} + \theta_{ij2} & \ldots & \varepsilon_{ij} + \theta_{ijt} \end{bmatrix} \\ &\boldsymbol{\delta}_{ijk} = \begin{bmatrix} \delta_{ijk} + \xi_{ijk1} & \delta_{ijk} + \xi_{ijk2} & \ldots & \delta_{ijk} + \xi_{ijkt} \end{bmatrix}. \end{aligned} \tag{2.5}$$

Let $\mathbf{\Sigma}_\varepsilon$ and $\mathbf{\Sigma}_\delta$ denote the covariance matrices of $\boldsymbol{\varepsilon}_{ij}$ and $\boldsymbol{\delta}_{ijk}$ respectively. The covariance matrix of observation vector is then described by the following formulae:

$$\mathbf{V} = \mathbf{I}_{ra} \otimes \left(\mathbf{1}_b \mathbf{1}_b^{'} \otimes \mathbf{\Sigma}_\varepsilon + \mathbf{I}_b \otimes \mathbf{\Sigma}_\delta\right), \tag{2.6}$$

where the symbols **I** and **1** mean the identity matrix and the vectors of ones respectively.

Assuming that the number of repeated measures is equal to three, the matrices $\mathbf{\Sigma}_\varepsilon$ and $\mathbf{\Sigma}_\delta$ take the following forms:

$$\mathbf{\Sigma}_\delta = \begin{bmatrix} \sigma_\delta^2 + \sigma_{\xi 1}^2 & \sigma_\delta^2 + \sigma_{\xi 12} & \sigma_\delta^2 + \sigma_{\xi 13} \\ \sigma_\delta^2 + \sigma_{\xi 12} & \sigma_\delta^2 + \sigma_{\xi 2}^2 & \sigma_\delta^2 + \sigma_{\xi 23} \\ \sigma_\delta^2 + \sigma_{\xi 13} & \sigma_\delta^2 + \sigma_{\xi 23} & \sigma_\delta^2 + \sigma_{\xi 3}^2 \end{bmatrix} \mathbf{\Sigma}_\varepsilon = \begin{bmatrix} \sigma_\varepsilon^2 + \sigma_{\theta 1}^2 & \sigma_\varepsilon^2 + \sigma_{\theta 12} & \sigma_\varepsilon^2 + \sigma_{\theta 13} \\ \sigma_\varepsilon^2 + \sigma_{\theta 12} & \sigma_\varepsilon^2 + \sigma_{\theta 2}^2 & \sigma_\varepsilon^2 + \sigma_{\theta 23} \\ \sigma_\varepsilon^2 + \sigma_{\theta 13} & \sigma_\varepsilon^2 + \sigma_{\theta 23} & \sigma_\varepsilon^2 + \sigma_{\theta 3}^2 \end{bmatrix} \tag{2.7}$$

Let the matrix $\mathbf{\Theta}' = [\boldsymbol{\mu}, \boldsymbol{\rho}_1, \boldsymbol{\rho}_2, ..., (\boldsymbol{\alpha\beta})_{a(b-1)}, (\boldsymbol{\alpha\beta})_{ab}]$ is the matrix of the model (2.4) parameters. Then, in order to check if experimental factors influenced significantly on mean value of studied feature, the null hypothesis H_0: **LΘM=0** should be verified against the alternative one H_1: **LΘM≠0**, where contrast matrix **L** describes linear combination of the experimental factor parameters and **M** describes the linear combination of the parameters connected with repeated measures.

The following orthonormal contrast matrices, connected with considered sources of variation, were used in this example:

$$\mathbf{L}_0 = \frac{1}{\sqrt{q}} \mathbf{1}_q', \ \mathbf{L}_A = \frac{1}{\sqrt{2}} \left[\mathbf{0}_{(a-1)x(r+1)} \ \vdots \mathbf{I}_{a-1} \vdots - \mathbf{1}_{a-1} \vdots \mathbf{0}_{(a-1)x(q-r-a-1)}\right],$$

$$\mathbf{L}_B = \frac{1}{\sqrt{2}} \left[\mathbf{0}_{(b-1)x(a+r+1)} \vdots \mathbf{I}_{b-1} \vdots - \mathbf{1}_{b-1} \vdots \mathbf{0}_{(b-1)x(q-r-a-b-1)}\right],$$

$$\mathbf{L}_{AB} = \frac{1}{\sqrt{2}} \left[\mathbf{0}_{(ab-1)x(r+a+b+1)} \vdots \mathbf{I}_{ab-1} \vdots - \mathbf{1}_{ab-1}\right], \tag{2.8}$$

$$\mathbf{M}_0 = \frac{1}{\sqrt{t}} \mathbf{1}_t, \ \mathbf{M}_T' = \begin{bmatrix} 1/\sqrt{2} & 0 & -1/\sqrt{2} \\ 1/\sqrt{6} & -2/\sqrt{6} & 1/\sqrt{6} \end{bmatrix},$$

where q describing the number of parameters is , in this case, equal to 34 and t is equal to 3. It is worth of noticing that the first column of $\mathbf{M}_T$ specify the linear time trend and the second one – quadratic time trend.

Following the method presented by Gumpertz and Brownie (1993) we do not construct the sums of squares and cross products (SSCP) matrices directly on experimental data but on contrast vectors $\mathbf{z}_{0}$,..., $\mathbf{z}_{(t-1)}$. The matrix $\mathbf{Z}$, consisting of these vectors, can be expressed as $\mathbf{Z}=\mathbf{Y}\cdot\mathbf{M}$, where $\mathbf{Y}$ is a matrix of observations with $\mathbf{y}_{ijk}$ describing by (2.4) as rows and $\mathbf{M}=[\mathbf{M}_0 \vdots \mathbf{M}_T]$.

The standard F test was used for verifying the null hypothesis for $\mathbf{z}_0$, while to estimate the significance of repeated measures the test statistic based on Wilks Λ (Morrison, 1990) was applied. The form of this function is depended on the minimum of t-1 and degrees of freedom for the hypothesis. The formulas for test statistic as well as for critical value might be found in Morrison (1990) or Gumpertz and Brownie (1993).

3. Results

The results of between-plot analysis, based on $\mathbf{z}_0$ contrast, estimating the treatments impact on mean value of chosen species are presented in Table 1.

Table 1. The results of the between-plot analysis

Source of variation	SS	df	MS	F	p-value
R (blocks)	1471.099	3	490.366	11.512	0.000
A (tillage systems)	1490.426	3	496.809	11.663	0.000
E_1=RxA	383.374	9	42.597		
B (herbicide doses)	9725.778	4	2431.444	33.983	0.000
AxB	908.320	12	75.693	1.058	0.415
E_2=RxB(A)	3434.368	48	71.549		

These results indicate significant differences of mean Ra between tillage system as well as herbicide subplots. There are no significant differences of this feature with respect to the interaction of experimental factors (AxB).

Table 2. The results of the within –plot multivariate analysis of variance

Source of variation	SSCP	df	Test statistic	p-value
T (time)	$\begin{bmatrix} 5355.753 & 9428.756 \\ 9428.756 & 16599.240 \end{bmatrix}$	1	487.563	0.0001
RxT	$\begin{bmatrix} 155.215 & -132.727 \\ -132.727 & 367.682 \end{bmatrix}$	3	2.464	0.0700
AxT	$\begin{bmatrix} 72.399 & 109.764 \\ 109.764 & 262.030 \end{bmatrix}$	3	1.952	0.1335
E_1=RxAxT	$\begin{bmatrix} 101.705 & -15.865 \\ -15.865 & 286.614 \end{bmatrix}$	9		
BxT	$\begin{bmatrix} 4132.027 & 1.306 \\ 1.306 & 160.004 \end{bmatrix}$	4	13.389	0.0001
BxAxT	$\begin{bmatrix} 312.705 & 149.452 \\ 149.452 & 626.920 \end{bmatrix}$	12	1.272	0.2057
E_2=RxB(A)xT	$\begin{bmatrix} 1326.092 & -140.560 \\ -140.560 & 1543.094 \end{bmatrix}$	48		

The results of within-plot multivariate analysis of variance (Table 2) indicate significant changes of studied feature during the time as well as significant influence of the interaction of herbicide doses and time on mean Ra of *Apera spica venti*. In order to check the character of these changes the trend analysis of time was done. Trends over 4 month period was examined and, as there are 3 nearly equally spaced measurement times, the time effect was partitioned into linear and quadratic contrasts (Table 3).

Table 3. ANOVA results for time trends

Source of variation	SS	df	MS	F	p-value
time linear					
T	5355.753	1	5355.753	473.939	<0.01
RxT	155.215	3	51.738	4.578	0.032
AxT	72.399	3	24.133	2.136	0.166
E1	101.705	9	11.301		
BxT	4132.027	4	1033.007	12.464	<0.01
BxAxT	312.705	12	26.059	0.314	0.976
E_2	1326.092	16	82.881		
time quadratic					
T	16599.240	1	16599.240	521.235	<0.01
RxT	367.682	3	122.561	3.849	0.0504
AxT	262.030	3	87.343	2.743	0.105
E1	286.614	9	31.846		
BxT	160.004	4	40.001	0.415	0.795
BxAxT	626.920	12	52.243	0.542	0.856
E_2	1543.094	16	96.443		

Both trends – linear and quadratic were statistically significant, which means that the average of studied feature were significantly different in every three dates of weeds estimation.

There is a significant linear component of the trend across the time for herbicides but this component is not the same for considered doses of herbicides. These significant differences of linear time component of mean Ra might be explained by different situation in B_5 (without herbicides), where mean Ra was greater in T_3 then T_1, on the contrary to other herbicide subplots (Table 4).

The analysis of variance of the quadratic contrast indicates that response over time is curved rather than linear and this nonlinear component is affected by herbicides. The quadratic time trends were the same on every herbicide subplot – mean Ra was the lowest in T_2.

Table 4. Mean Ra on herbicide subplots

Herbicide doses	T_1	T_2	T_3
0%	42.15	27.99	48.14
25%	40.31	19.48	33.50
50%	38.89	11.26	23.12
75%	38.08	12.33	16.14
100%	40.03	11.26	20.71
Average	39.89	16.47	28.32

The results of within plot analysis pointed out the lack of significant differences between Ra for the interaction between time and tillage systems, so the trends over time for this factor were no examined.

4. Conclusions

In the paper complete specification one of the methods of variance analysis and interpretation of results of agricultural experiment conducted in split-plot with repeated measures design was presented. In this method estimation of covariance matrix is omitted, what is an important advantage of this analysis since sometimes the complicated structure of this matrix might pose the problem.

The obtained results enable us to state that there are significant differences of mean relative abundance of studied species between tillage system as well as herbicides doses.

The character of changes of feature in time was curvilinear for the reason of the significance of both time trends – linear and quadratic. It was probably caused by the application of herbicides before the second date of estimation of the weeds infestation.

The significant linear trend of Ra changes on herbicide subplots was caused by the differences in changes of *Apera spica venti* between first and third dates on control (zero) subplot comparing to others. On the other hand decreasing Ra in second date and increasing in third one affected the lack of significant quadratic trends on herbicide subplots.

References

Blouin D. C., E. P. Webster, W. Zhang (2004). Analysis of synergistic and antagonistic effects of herbicides using nonlinear mixed-model methodology. *Weed Technol*. 18, 464–472.

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.

Fidel A. (1998). A Bluffer's Guide to Sphericity. *Newsletter of the Mathematical, Statistical and computing section of the British Psychological Society*, 6 (1), 13–22.

Gumpertz M. L., Brownie C. (1993). Repeated measures in randomized block and split-plot experiments. *Can.J. For Res*. 23, 625–639.

Li H., Wood C. L., Getchell T. V., Getchell M. L., Stromberg A. J. (2004). Analysis of oligonucleotide array experiments with repeated measures using mixed models. *Bioinformatics* 5, 209.

Linnel Nemec A.F. (1996). *Analysis of repeated measures and time series: An introduction with forestry examples*. Biometric Information Handbook No. 6.

Littell R. C., Henry P. R., Ammerman C. B. (1998). Statistical Analysis of Repeated Measures Data Using SAS Procedures. *J. Anim. Sci*. 76, 1216–1231.

Morrison D. F. (1990). *Multivariate statistical methods* (3rd ed.). N.Y. McGraw Hill.

Moser E. B., Saxton A. M., Pezeshki, S. R. (1990). Repeated measures analysis of variance: application to tree research. *Can. J. For. Res*. 20, 524–535.

ZASTOSOWANIE ANALIZY WARIANCJI DLA UKŁADU SPLIT-PLOT Z POWTARZANYMI POMIARAMI W OCENIE ZMIAN *APERA SPICA VENTI* POD WPŁYWEM SPOSOBÓW UPRAWY ROLI I DAWEK HERBICYDÓW

Streszczenie

W niniejszej pracy przedstawiono analizę wariancji dla modelu stałego split-plot z powtarzanymi pomiarami. Wybrano metodę opartą na jedno- i wielowymiarowej analizie kontrastów ortonormalnych uniezależniającą obliczenia od postaci macierzy kowariancji. W ten sposób oszacowano wpływ sposobów uprawy roli i dawek herbicydów na względną obfitość (Ra) dominanta zbiorowiska chwastów pszenicy ozimej, ocenianą w trzech terminach. Średnia wartość badanej cechy różniła się istotnie zarówno ze względu na termin oceny jak i dawkę herbicydów oraz interakcje tych czynników. Analiza trendów czasowych wykazała istotną różnicę średniej Ra we wszystkich trzech terminach jak również istotną zmianę liniowego trendu czasowego pomiędzy poletkami herbicydowymi.

Słowa kluczowe: analiza kontrastów, dawki herbicydów, sposób uprawy roli, układ split-plot, układ z powtarzanymi pomiarami, względna obfitość

Klasyfikacja AMS 2000: 62–07