THE METHODOLOGICAL ASPECTS OF USING MULTIFACTORAL ANALYSIS OF VARIANCE IN THE EXAMINATION OF EXPLOITATION OF ENGINE SETS

Zbigniew Burski*, Joanna Tarasińska*, Romuald Sadkevič**

* The Agricultural University of Lublin, Poland ** The Higher School of Agriculture in Vilnius, Lithuania

INTRODUCTION

In the vibroacoustic diagnostics of the exploitation of engine sets many dimensional and non-dimensional symptoms that characterise their technical state are used to evaluate the occurring dynamic processes. Some of them may characterise amplitude spectrum of the dynamic process and indicate the dominant energy band. In the first stage, the diagnosis can indicate the development of the defect, an unbalance, misalignment, clearance, etc. [4, 5, 6].

However, a mere qualification of qualitative changes in the development of the failure, for more and more structurally complex machines and in the endorsed work conditions, is not sufficient. A quantitative evaluation of the phenomenon is required [1, 2, 3, 9]. Methods of mathematical statistics play a crucial role here, as they can prove a statistically significant influence of some factors on the examined phenomenon [7, 8, 10].

The first and necessary requirement here is the proper choice of examination methodology and measuring apparatus, so that a repetition of results may be obtained.

THE AIM OF THE PAPER

The aim of this paper is to present the methodology of using multifactoral analysis of variance in the examination of the influence of different exploitation factors on a tractor's engine performance in the conditions of research on experimental simulation. The presented methodology is illustrated with an example of vibroacoustic testing.

THE CHOICE OF THE VARIABLES FOR THE STATISTIC ANALYSIS

During the measurement of an engine's vibration level, the levels of vibrations with chosen frequencies are measured, giving the amplitude spectrum. Such spectra are recorded at different levels of the examined environmental and exploitation factors affecting the spectrum's shape.

Let v_i denote the vibration level at the *i*-nd of the considered frequencies $f(i = 1,...,k \ f_1 < f_2 < ... < f_k)$. The variables representing (characterising) the spectrum in the following statistical analysis can be:

- the vibration level v_{i_0} with steady frequency f_{i_0} ($i_0 = 1, K, k$),

- the arithmetic mean of vibrations,

$$\overline{v} = \frac{1}{k} \sum_{i=1}^{k} v_i \tag{1}$$

- a better measure of an average level of vibrations than \overline{v} is a weighed mean, analogous to the integral mean for the continuous spectrum v(f), counted by the equation:

$$v_m = \frac{1}{2(f_k - f_1)} \sum_{i=1}^{k-1} (v_i + v_{i+1})(f_{i+1} - f_i)$$
(2)

- the 'gravity centre of an amplitude spectrum', i.e. the mean vibration frequency, counted by the equation:

$$\bar{f} = \frac{\sum_{i=1}^{k-1} (v_i f_i + v_{i+1} f_{i+1})(f_{i+1} - f_i)}{\sum_{i=1}^{k-1} (v_i + v_{i+1})(f_{i+1} - f_i)}$$
(3)

- the maximum vibration level $v_{Max} = \max_{i=1,K,k} (v_i)$.

Obviously, each of the random variables mentioned here takes into account only one of the aspects of the amplitude spectrum, not the whole of it.

PEAK, LIN and C can also be examined, where: LIN is the linear vibration (PEAK)

level, PEAK is the peak value, and C is the peak coefficient $\left(\frac{PEAK}{LIN}\right)$.

THE REPETITION OF MEASUREMENTS

The measure of the repetition of the measurements x_1 , K, x_n of the same random variable X can be the coefficient of variation given by the equation:

$$\gamma = \frac{\sigma}{\overline{x}} = \frac{\sqrt{\frac{1}{n-1}\sum_{i=1}^{n} (x_i - \overline{x})^2}}{\overline{x}}$$
(4)

where:

 $\overline{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$ is the average value of the measurements and $\hat{\sigma}$ their standard deviation.

During the measurement of a spectrum of vibration the spectrum is recorded at different levels of the considered factors affecting its shape. The more measurements we have performed, the better is the statistical analysis of such a spectrum (by the analysis of the variables characteristic of the spectrum, e. g. those mentioned in the chapter "The Choice of the Variables for the Statistic Analysis"). However, it is often hard to obtain many spectrum's measurements in the determined experimental conditions. If there are several exploitation factors that we want to consider and each of them can be found at different levels, then, by the recording of only one spectrum in all the possible combinations of levels we receive several dozen or more spectra. In such type of research we often have to be satisfied with only one measurement at the determined levels of the factors. That is why it is important to make sure whether the measurement is 'representative' (repetitive).

In order to check it, 9 measurements of a vibration spectrum were performed in the following exploitation conditions [3]:

- worn out TPC (kinematic pair piston rings bush with limiting wear),
- small chamber (experimental compression chamber with big gas force $-P_g$),
- measure plane lateral to the engine axis,
- measurement point N reverse to the direction of rotation

The vibration level with 44 frequencies from 1 Hz to 20000 Hz was recorded. Fig. 1 shows examples of the recorded spectra for one-cylinder crankpiston system of an engine S-4002/3.



Fig. 1. The measurements of the repetition 1-9 of a vibration spectrum. The conditions of the measurements: worn out TPC, small chamber, lateral plane, point N

47

Let v_{ij} be the *j*-th vibration measurement with the *i*-th frequency (i = 1, ..., 4; j = 1, ..., 9). The coefficients of variation of measurement with the *i*-th frequency were counted:

$$\gamma_{i} = \frac{\sqrt{\frac{1}{8} \sum_{j=1}^{9} (v_{ij} - \overline{v}_{i})^{2}}}{\overline{v}_{i}}$$
(5)

where:

$$\overline{v}_i = \frac{1}{9} \sum_{j=1}^9 v_{ij}$$
 .

As we can see in figure 1, for the low frequencies range (1-5 Hz) the coefficients of variation are very high, up to 40%. Thus, in this range the results of the measurements cannot be recognised as repetitive.

Table 1 presents the mean values, standard deviations and coefficients of variation for the proposed variables characterising the spectrum in the range 6.3-20000 Hz.

Table 1. The mean values, standard deviations and coefficients of variation for some of the variables characterising the spectrum

Variable	v _i with frequency 1000 Hz	\overline{v}	v_m	\overline{f}	v_{Max}
Mean value	124.02	129.35 dB	138.1635 dB	1014135 Hz	145.9635 dB
Standard deviation	5.563	4.266 dB	3.92835 dB	47.9435 Hz	3.76135 dB
Coefficient of variation	4.49%	3.30%	2.84%	0.47%	2.58%

As we can see from Table 1, the variables proposed for an analysis are characterised by a sufficiently low coefficient of variation.

THE CRYTERION OF CHOOSING THE MODEL OF THE MULTIFACTORAL ANALYSIS OF VARIANCE

The multifactoral analysis of variance (ANOVA) is a statistical method that helps us discover the influence of the considered factors on a chosen random variable. Since the factors determined by exploitation conditions can interact with each other (i.e. the effect of interaction can take place), the best thing would be to have many measurements of the variable at the determined levels of the factors. As it was mentioned above, in case of a vibration spectrum's examination we usually have only one spectrum's measurement. This allows for just 1 observation in the ANOVA subclass, and does not give any opportunity to evaluate error or perform any statistical analysis.

Let *a* be the number of factors considered in the research. In such a situation we suggest performing the analysis in two stages:

1. From the range of all possible analyses of variance with a-1 factors (there are a of them) choose the one which gives the least mean square root for error [10]. The factor eliminated in that way has the least influence on the considered random variable and is used for an evaluation of the statistical error.

2. For the model chosen in item 1. perform an analysis of variance with a-1 factors. If one of the remaining factors, with all the including it interactions, turns out to be insignificant (at the selected level of significance), remove this factor and limit the analysis to the remaining a-2 (or fewer) factors. The factor removed in such a way will increase the number of error tolerance stages.

During an interpretation of the results of the analysis of variance the highest level of interaction should be considered, the one in which the significant (on the selected significance level) interactions are involved. For example, if the interactions of the factors 1 and 3, as well as those of 2 and 3 turn out to be significant, the interaction 123 ought to be interpreted, even if it is insignificant in itself (1, 2, 3 are the measures of factors). If all the assumptions of the analysis of variance (see the next chapter) are fulfilled, the test LSD (Least Significance Difference), [7, 10] can be used for a more detailed evaluation of the influence of the factors on the examined value.

THE ASSUMPTIONS OF THE ANALYSIS OF VARIANCE AND THE CONSEQUENCES OF THEIR TRANSGRESSION

ANOVA is built on the following assumptions:

1. The homogeneity of variance: it is assumed that variances within subclasses are the same. This hypothesis can be verified with, for example, Bartlett's, Cochran's or Hartley's test. According to many authors (Lindman, 1974, after Statistic, 1997) the F statistic, on which the analysis of variance is based, is fairly robust to violations of such an assumption. In one specific case, however, the F statistic can be misleading, i.e. the case when means and variances within subclasses are correlated (Statistica PL for Windows, 1997, Statsoft Inc). If, for example, in the subclass with the highest mean the highest variance occurs, this mean can influence the variance analysis which is based on the estimation of a common variance in the subclasses. The same mean could be essentially similar to others if we took into account different variances in subclasses. If we discover such kind of transgression from ANOVA assumptions, we should try to transform data using functions known as the Box-Cox transformations. One of such monotone, order-preserving, variance stabilising transformation is logarithm or square root.

2. *The normal distribution of response*. This assumption can be checked by plotting sorted residuals against corresponding quantities of the normal distribution [10]. Despite the fact that ANOVA is robust to departures from normality, the use of an appropriate transformation (often the same which stabilises variance) improves the normality of the residuals.

THE ILLUSTRATION OF THE PRESENTED METHODOLOGY

The object of the examination was the dependence of the vibrations of a S 4002/3 engine system on the following factors:

- TPC condition (new, worn),

- presence of a compression chamber (without compression, with compression),

- 3 vibration planes (transversal, longitudinal and vertical),

-2 points of measurement (N – opposite to and N' – concordant with the hand of rotation of the engine's crank-shaft),

- oil contamination (pure oil, oil with silicon dust 8 mg/m³, oil with silicon dust 16 mg/m^3).

For each of the 72 combinations $(2 \times 2 \times 3 \times 2 \times 3)$ of the listed factors the values of vibrations with 16 frequencies from the range of 315-10000 Hz were recorded. The statistical analysis of the maximum vibration level was performed according to the methodology presented in this paper. There are 5 factors here which influence the v_{Max} value and only 1 v_{Max} value in a subclass. The initially performed 4-factor analyses of variance gave errors' evaluation (mean square roots for the Mse error) presented in Table 2.

Table 2. Mean square roots for error in 4-factoral ANOVA for the variable ν_{Max}

Omitted factor	TPC	Chamber	Plane	Point	Oil
MS _e	252.93	248.58	67.70	37.16	34.09

Thus, 4-factoral analysis TPC \times Chamber \times Plane \times Point was chosen. The Bartlett's homogeneity test gives p – value 0,008745. This means that there is heterogeneity of variance here. Fig. 3 shows the normality of the residuals graph, and Fig. 2 points out that, unfortunately, a transgression of essential ANOVA assumptions concerning lack of correlation between means and standard aberrations in groups takes place here.



Fig. 2. Standard aberrations and means in subclasses for the v_{Max} variable



Fig. 3. Normality of the residuals graph for the v_{Max} variable

Considering the correlation of means and standard aberrations in groups, we decided to use root transformation of data. For the $\sqrt{v_{Max}}$ variable, tab. 3 gives errors' evaluations in the 4-factoral ANOVA.

Table 3. Mean root squares for error in the 4-factoral ANOVAs for the $\sqrt{v_{Max}}$ variable



0,4 0,2 0,0 -0,2

2.5

3.5

Fig. 4. Standard aberrations and means in subclasses for the $\sqrt{v_{Max}}$ variable

5.5

6,5

7,5

8.5

4.5



Fig. 5. Normality of the residuals graph for the $\sqrt{v_{Max}}$ variable

The Bartlett's homogeneity of variance test for the $\sqrt{v_{Max}}$ variable gives p-value 0.923. This fact, together with figures 4 and 5, indicate that the used root transformation fulfilled its task, i.e. it has equalised variance in groups, diminished the correlation between the means and the standard aberrations (correlation coefficient is now insignificant, on the level 0.05) and normalised the data.

In table 4 the essential effects of 4-factoral ANOVA TPC \times Chamber \times Plane \times Oil l on the level 0.05 were presented.

1-TPC, 2-Chamber, 3-Plane, 4-Oil									
effect	df effect	MS effect	df error	MS error	F	p-value			
1	1	76.85421	36	0.315573	243.5387	0.000000			
2	1	65.95370	36	0.315573	208.9968	0.000000			
3	2	2.37324	36	0.315573	7.5204	0.001866			
12	1	1.98111	36	0.315573	6.2778	0.016887			
13	2	1.80941	36	0.315573	5.7337	0.006892			
23	2	3.19891	36	0.315573	10.1368	0.000322			
123	2	1.15735	36	0.315573	3.6675	0.035511			
124	2	1.31904	36	0.315573	4.1798	0.023316			

Table 4. The outcomes of 4-factoral ANOVA for the $\sqrt{v_{Max}}$ variable

Although the 1234 interaction is not essential, we have to interpret it in view of the mutual entanglement of interactions 124 and 123.



Figures 6 and 7 show the graph of a 4-direction interaction (because of the number of factors this interaction cannot be presented in one figure). In the graphs the vertical bar shows the size of LSD – Least Significant Difference on significance level 0.05. Only differences between means which are greater than LSD are significant.



Fig. 7. Interaction TPC × Chamber × Plane × Oil, without compression chamber

CONCLUSIONS

The used methodology of statistical calculations by the method of interaction of multifactoral variance has found its practical application in the simulated examination of vibration of the kinematic pair t-p-c for different exploitation conditions. The essential requirements for using this method were taken into account here.

The description of the calculation methodology was illustrated with the concrete examples of its use, which gives detailed information for the interested.

From the analysis of LSD and Fig. 6 and 7, used as an example for calculations, follows that:

- there is no essential influence of the measurement point on the maximum vibration level,

- a new TPC has a lower maximum vibration level than a worn out one, the difference is insignificant only in the presence of compression chamber, with pure oil in transversal plane, and without compression chamber, with oil + dust 16 mg/m^3 in longitudinal plane,

- the presence of compression chamber usually causes reduction of the maximum vibration level, the differences are insignificant only for new TPC, transversal plane, pure oil; new TPC, longitudinal plane, pure oil and oil + dust 8 mg/m^3 ; worn out TPC, longitudinal plane, oil + dust 16 mg/m^3 ,

- the influence of oil contamination with silicon dust is slight, shows only for a worn out TPC with the compression chamber in a transversal plane. Pure oil gives a significantly lower maximum vibration level than dust contaminated oil.

REFERENCES

- Burski Z. i in.: Analiz tribologiczeskich simptomow kolebanij kinematiczeskoj pary traktornogo dwigatelja w modelirowanych isledowanijach. Zbiornik naukowych prac Nacionalnogo Agrarnogo Uniwersitetu Ukraini. T. X. Kijew, 2001, s. 189-198.
- Burski Z. i in.: The simulation researches on the vibrations of a cylinder sleeve of tractors engine connection with its strain caused by the pressure of gases in the condition of limitary consumption. Proceedings of National Agricultural University of Ukraine. V. X, Kijiv: NAUU, 2001, p.160-171.
- Burski Z., Tarasińska J.: Analiza dynamiki rozruchu silnika ciągnikowego metodą interakcji wieloczynnikowej wariancji. PAN o/ Lublin, Polit. Szczecińska, nr 6. Szczecin – Trelleborg. 2002, s. 21-27.
- 4. Cempel Cz.: Podstawy wibroakustycznej dynamiki maszyn. WN-T. Warszawa. 1982.
- Cempel Cz.: Multi dimensional condition monitoring of mechanical systems in operation. XXIX Sympozjum "Diagnostyka Maszyn". Polit. Sląska. Węgierska Górka 2002, s. 101-118.
- Konprobst H., Woschni G., Zeilinger K.: Simulation des Verhaltes von Kolberingen im Motobetrieb. MTZ 11, 12, 1989.
- Oktaba W.: Elementy statystyki matematycznej i metodyka doświadczalnictwa. PWN. Wyd. 4. Warszawa 1976.
- 8. Statistica PL. Dla Windows, Statsoft. Inc., 1997.
- Tarasińska J., Burski Z., Sadkiewicz R.: Application of Analysis of Variance in the Simulation Researches on the Dynamics of the Vibrations of a Kinematic Pair of an Engine Under the Conditions of Air Dustiness. Polish Academy of Sciences Branch in Lublin. Commission of Motorization and Energetics in Agriculture. Vol. II. Lublin 2002, p. 173-181.
- 10. Yandell B. S.: Practical data analysis for designed experiments. Chapman & Hall, London. 1997.

SUMMARY

The paper presents methodological aspects of using analysis of variance in simulated examinations of vibration dynamics of the kinematic pair piston – piston rings – cylinder sleeve of an agricultural machine set. The choice of variables for a statistical analysis by the method of multifactoral analysis of variance, repetitiveness of the evaluation results and cryteria of choice were presented. The accepted methodology was illustrated with the examples of different exploitation conditions of a tractor's engine.