ARCHIVES OF ACOUSTICS
32, 4 (Supplement), 235–245 (2007)

ANALYSIS OF STOCHASTIC ACOUSTICAL HAZARDS IN ENVIRONMENT

Wojciech BATKO, Renata BAL–PYRCZ

AGH University of Science and Technology
Department of Mechanics and Vibroacoustics
Al. Mickiewicza 30, 30-059 Kraków, Poland
e-mail: batko@agh.edu.pl

State Higher Vocational School in Krosno
Politechnical Institute
Rynek 1, 38-400 Krosno, Poland

(*received July 15, 2007; accepted October 5, 2007*)

Acoustical assessment of environment, realized on the basis of changes observation held as part of national environment monitoring for the need of long term noise protection policy in environment, demands long term average noise level L_{DEN} and L_N estimation. The estimation is based on noise measurement results from all the days in the year, taking into account day, evening and night times. The basis for correct statistical estimation when the data is incomplete (time limited monitoring of acoustic environment) is a choice of time sample so that results gained reflect the mechanism of noise hazard changes that works throughout the year. It should provide identification of particular event occurring frequency and time relations between events, e.g. sound extreme value appearance or recognition of periodical changes of important parameters in uncertainty assessment. Such process demands assumption of particular statistical techniques related to the assumed model of noise level changes where the random factor is important and always present. That problem is part of the article. The paper presents basic analysis of monitored phenomenon by usage of models framing its stochastic volatility. Important elements of modeling are described, illustrated by examples of yearly noise data analysis gathered at one of main streets in Kraków. Attention is paid to the analysis problem with the process periodicity, measurement uncertainty related to the measurement conditions volatility or time uniformity of the monitored noise processes in view of its internal causality relations.

They are loaded with random factor whether we assign randomness to noise source volatility, whole year measurement conduct inability or distortions related to proper sampling of given field of noise hazards.

Keywords: noise analysis, noise control, noise condition, hipothesis testing in time series analysis.

1. Introduction

It is necessary for the process of environmental noise control programs development to properly estimate long-term average noise levels $L_{\mathrm{DEN}}$and L_N. The estimation is based on noise measurement results from all the days in the year, taking into account day, evening and night times. The basis for correct statistical estimation when the data is incomplete (time limited monitoring of acoustic environment) is a choice of time sample so that results gained reflect the mechanism of noise hazard changes that works throughout the year. They should be representative in terms of their equality of the value calculated from the sample to proper noise level ratio characteristic for the year. From their nature the acoustical influence processes are stochastic processes, where the random factor is always present. It might be connected to incomplete knowledge about the factors exciting the volatility of the noise hazards levels, inability of conducting well-conditioned exact measurement, or the proper sampling of the process. For this reason in long term assessment of the noise level ratios the statistical analysis methods of time series of noise level volatility are playing a significant role. That applies to such problems where the frequency of appearance of particular events (like extreme values appearance) is calculated, uncertainty assessment, search for relations between tested quantities by model verification etc. For the needs of stochastic analysis of noise hazards in environment a research schematic was developed, presented in Fig. 1, showing testing actions leading to the predictive recognition of noise hazards in environment.

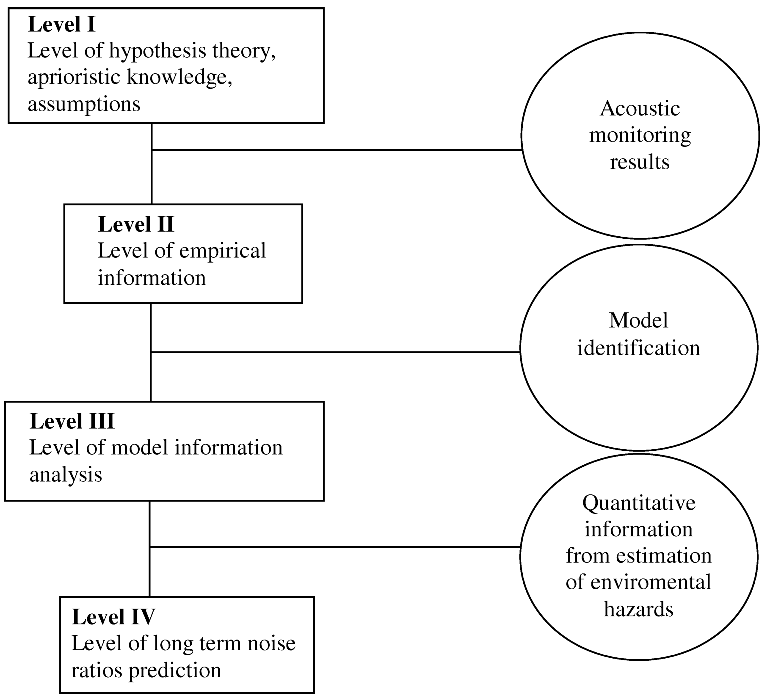

Fig. 1. Testing actions important for diagnostic information development on noise hazards state in environment.

The first level of testing is related to the accepted base of assumptions related to the modeled reality. It might be related to recommended by directive 2002/49/EC [4] models of calculation in noise environment: road, railroad, industrial or aviation. The model formalism related to that is defined by source emission calculation module and its propagation conditions. Here the expected properties of the tested processes might be assumed, including the information about its potential statistical properties.

The second level is related to the ways of acquiring the empirical information on the analyzed process of the noise hazards. It is defined by available database from the acoustical monitoring of the environment, being a sampled stochastic process. Their processing with proper analysis methods of time series allows identification of determinants of volatility of noise hazards in environment including periodicity of the process, its stationary or volatility defined as the measure of uncertainty of future changes of the tested noise ratios. Presenting methods of this stage in context of their application in acoustical monitoring procedures important from theoretical and practical point of view is a subject of this paper.

Levels three and four are the levels of application of the accepted model (see BATKO [3]). It is related to usefulness of the accepted model in explaining volatility and value prediction of the long-term noise ratios. Those levels differentiate values of the related qualitative and quantitative results. It is caused by the fact that information on level 3 used for parameters estimation of the accepted model and its verification are more precise than on level 4. The prediction of the model values is usually based on information with much larger uncertainty, including additionally uncertainty related to keeping in the assumed time span in terms of the model or input data for the analyzed phenomenon.

2. Structure analysis of a process of stochastic volatility of noise hazards in environment

2.1. Randomness study

Volatility of noise level in environment $L_{eq\,A}$, given as time series $L_{eq\,i}$, $i = 1, 2, \ldots, n$ need their randomness nature hypothesis solving. Chaos theory brings tools that allow distinguishing random series from non-random series. One of such tools is Hurst exponent (see ANIS, LLOYD [1]. Its application for randomness assessment was first described in master's thesis MORZYK [6]. Its value allows distinguishing random series from non-random series provided granted access to any time series long enough. The measure created by Hurst is based on *rescaled range* R/S analysis related to analysis of a quotient of the analyzed parameter to the mean value. It identifies random change series of the analyzed parameter to the range of variation proportion to the square root of time $(t^{1/2})$. By means of rescaled range R/S Hurst proposed the following form of the model:

$$\left(\frac{R}{S}\right)_n = cn^H, \tag{1}$$

where S is increment standard deviation in time n, c some positive constant, n – number of observations and H is Hurst's exponent.

When calculating H, first average R/S values for different n need to be calculated, then by means of linear regression the following equation should be solved:

$$\ln E(R/S)_n = \ln c + H \ln(n).$$

So it is enough to plot $E(R/S)$ against n in double logarithmic scale. The slope of the curve will then estimate H. Hurst expanded Einstein's model from $t^{0.5}$ to t^1. Determination how much Hurst's exponent is above 0.5 allows to determine the trends. Depending on the resulting value of the Hurst exponent the tested time series has the following properties:

- $0 \leq H \leq 0.5$ ergodicity which means phenomenon of returning to mean value.
- $H = 0.5$ the time series is random, that is its volatility is random. Those are the time series where present is independent of past and future is independent of present. Such series are unpredictable.
- $0.5 \leq H \leq 1$ there is a long term memory in the series that means the values of the series are correlated. In those series past influences present and present influences future Morzyk M., [6].

If the result of the calculation of the Hurst exponent is greater than 0.5 that denotes the fact of existing trend in the series, existing repeatable sequences of the results. If the values of the Hurst exponent is smaller than 0.5, the tested process converges to the mean value.

For exemplary analysis database was used from the continuous monitoring station of road noise in Kraków, from the year 2004. Experimental data was changing values of $L_{A\,eq\,T}$ calculated in 15 minutes intervals. The calculations of Hurst exponent are presented in the Tables 1 and 2.

Table 1. Hurst exponent values for monitored values of L_{eq}.

Year 2004	Hurst exponent	Average	Deviation
January	0.439	70.4	3.7
February	0.376	71.7	3.2
March	0.454	71.9	2.8
April	0.449	72.3	2.9
May	0.384	71.6	2.7
June	0.384	71.7	2.9

Year 2004	Hurst exponent	Average	Deviation
July	0.445	72.1	2.9
August	0.452	71.5	2.9
September	0.413	72.1	2.9
October	0.419	72.5	2.6
November	0.401	72.1	3.3
December	0.479	71.7	3.1

Table 2. Hurst exponent values for monitored values of L_{eq} for day, evening and night times.

Year 2004	Hurst exponent	Average	Deviation
Day	0.678	73.4	1.3
Evening	0.743	72.8	0.9
Night	0.826	68.6	1.4

As one can see by choosing different times one might receive different values of Hurst exponent. They point to existing different in time mechanisms shaping volatility of the monitored noise hazards. The road noise level L_{eq} volatility in monthly periods has a characteristic of convergence to the mean values, which is presented in the Table 1. The process defining the volatility is ergodic. The analyzed volatility in cycles of day evening and night does not have that property. There exists long term memory, which means that the values in the series are correlated and there is a trend in them. As it is seen from the Table 2 the strongest correlation shows up in the night time. Their identification might be done by moving average.

2.2. Periodicity study

The volatility process of noise level might be described as stochastic. In characteristic of such process it is important to estimate the spectral density function. Based on the spectral density curve it is possible to identify the existing trend or the presence of periodic changes (see TALAGA, ZIELIŃSKI, [8]. The value of spectral density function might be interpreted as percentage of harmonic variation of given frequency in total variance of the process.

Each real stochastic process might be presented as:

$$X_t = \int_0^\pi \left[\cos\omega t\,\mathrm{d}U(\omega) + \sin\omega t\,\mathrm{d}V(\omega)\right], \qquad t = 0,\ \pm 1,\ .... \tag{2}$$

where $\mathrm{d}U(\omega)$ and $\mathrm{d}V(\omega)$ are random functions fulfilling the following conditions:

$$E\left[\mathrm{d}U(\omega)\ \mathrm{d}U(\lambda)\right] \ = \ E\left[\mathrm{d}V(\omega)\ \mathrm{d}V(\lambda)\right] = \begin{cases} 0, & \text{for } \omega \neq \lambda, \\ 2\mathrm{d}F(\omega), & \text{for } \omega = \lambda \end{cases}; \tag{3}$$

$$E\left[\mathrm{d}U(\omega)\ \mathrm{d}U(\lambda)\right] \ = \ 0, \qquad 0 < \omega, \quad \lambda < \pi.$$

Function $F(\omega)$ is a spectral function and the function $F(\omega)$ is differentiable

$$\mathrm{d}F(\omega) = f(\omega)\,\mathrm{d}\omega, \tag{4}$$

where $f(\omega)$ – spectral density function.

As an example the figures below represent the spectrum of noise level L_{eq} volatility for the data from yearly noise monitoring in day (Fig. 2), evening (Fig. 3) and night (Fig. 4) times. It was calculated with use of GRETL (see KUFEL [5]).

Conclusions: The shape of the spectral density curve proves the importance of periodic variations (presence of peaks). There are very visible peaks present, very steep which means the periodicity is close to single frequency. There are also flat peaks which means the periodicity is volatile and irregular. The presence of peaks for seasonal frequencies proves the pertinence of seasonal changes. For each time period the shape of the curve is similar which proves identical harmonic structure of the tested processes. Seasonal variations with the period of 6 months are important. Variations with the period of 3 and 2 months are irregular and have small amplitude. The calculated curves prove

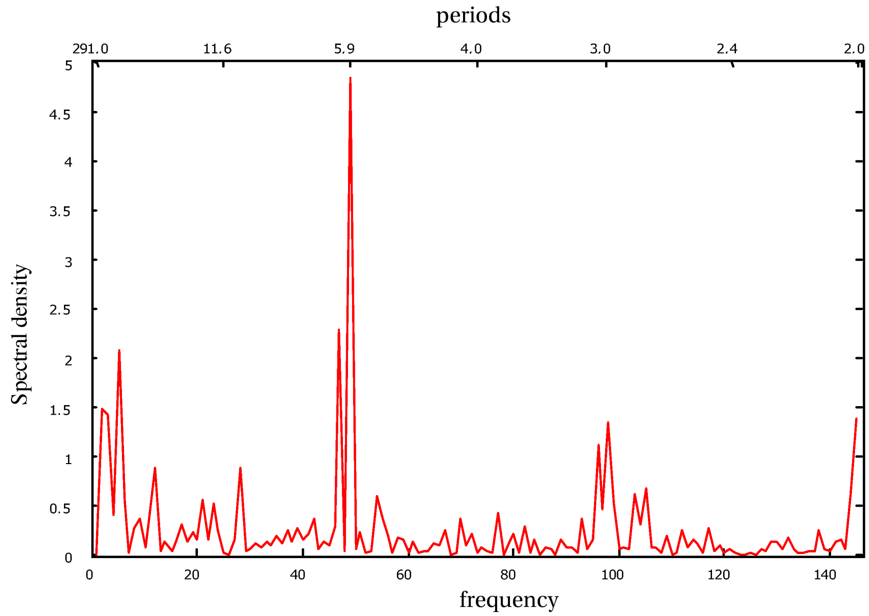

Fig. 2. The spectrum of road noise level changes $L_{eq\,D}$ for the day time.

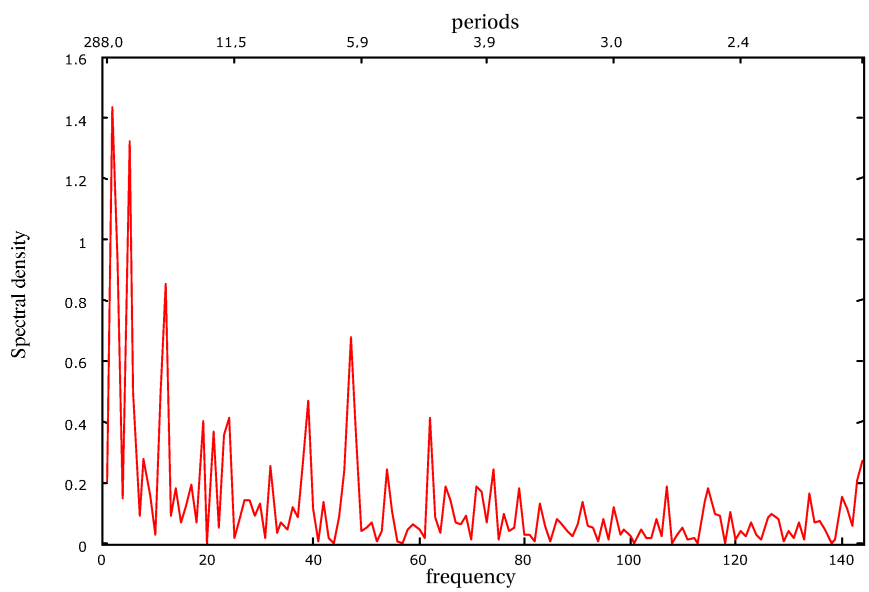

Fig. 3. The spectrum of road noise level changes $L_{eq\,E}$ for the evening time.

the existence of a trend, which is interfered by random fluctuations which have significant role in total process variation. Rather small values of density function for low frequencies prove a weaker trend. Strong trend exist for high frequencies. Long term fluctuations are significant. Short term fluctuations (period of 2.4, 3) are insignificant.

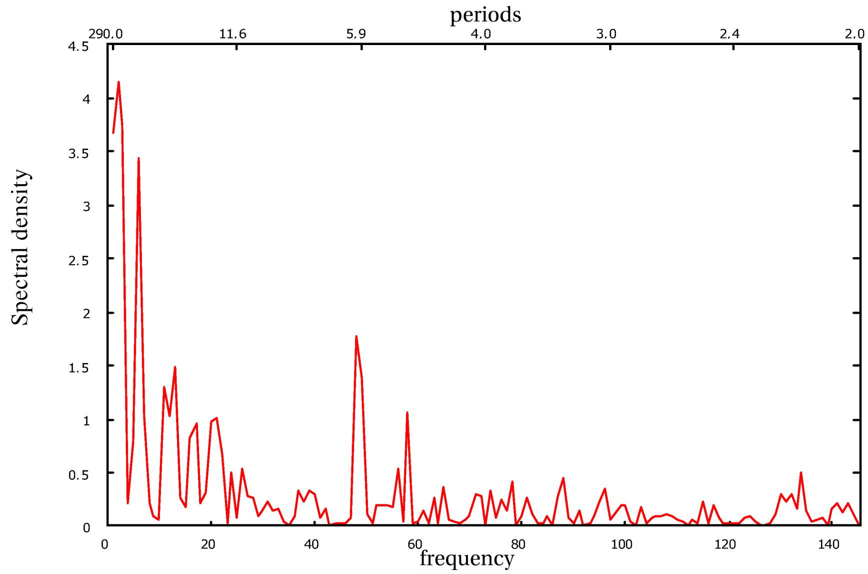

Fig. 4. The spectrum of road noise level changes $L_{eq\,N}$ for the night time.

2.3. *Stationarity study*

By accepting of the value volatility randomness model of the analyzed time series (see ZIELIŃSKI [10]) it is important to answer the question of stationarity. The commonly observed noise level L_{eq} volatility randomness takes its place in measurements which result is determined by unrepeatability of the measurement conditions. Therefore the stationarity analysis is needed in particular time ranges. The stationarity of the tested process is related to its properties. The stochastic process is stationary if certain conditions are met:

• the average value is constant

$$m(t) = E\left[X(t)\right] = \text{const}; \tag{5}$$

• covariance function depends only on the moment difference

$$K(t_1, t_2) = K(t_2 - t_1) = K(\tau), \qquad \text{where} \quad \tau = t_2 - t_1\,. \tag{6}$$

One of the more important characteristic of random variable distribution creating stationary process is an autocorrelation factor. Correlogram of the process is a set of autocorrelation factors of the stochastic process. The Fig. 5 presents such correlogram for the process of noise volatility. There are three functions of correlogram for day, evening and night time.

Based on the chart one might say that the control process is stochastic and stationary. Knowledge of properties of different types of stochastic processes and ability to distinguish their types based on the observation of the process allows finding possibly most effective estimators of noise level volatility models parameters.

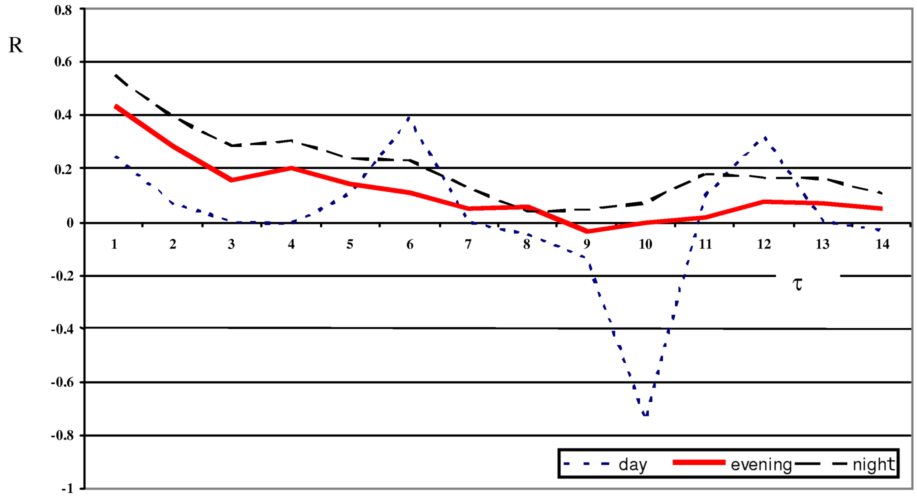

Fig. 5. Correlogram of road noise level volatility L_{eq} for the day, evening and night time.

2.4. Measurement data uniformity study

In the assessment of the results of conducted control noise measurement it is important to solve the questions if the distribution of the random vector L_{DEN} or its components L_D, L_E and L_N in each month, day, week etc are uniform. To answer such question it is important to put up the proper hypothesis and the choice of the proper method of verification. To separate the set of uniform results it is important to test if the average values in the range of given levels for example weekday differ in radical way and the variance should be calculated (see PAWŁOWSKI [7], STECZKOWSKI, ZELIAŚ [9]). The conclusions from those analyses should provide unequivocal answer to the questions put. For the assessment of differences between the average values in chosen groups we use variance analysis. That allows to state if the classification used is proper in terms of the division criteria choice, that is if we have uniform groups inside but being different from each other at the same time.

The analysis is illustrated by the results of yearly monitoring of noise from year 2004 measured by continuous monitoring station in Kraków. It was analyzed if the variances in each month differ from each other in day, evening and night times. For this hypothesis verification a Bartlett test was used.

Based on the results I check the hypothesis H$_0$: $\sigma_1^2 = \sigma_2^2 = ... = \sigma_k^2$ $(k = 1, ..., 12)$ confronting alternative hypothesis that at least two of the variances form freely chosen month differ from each other relevantly: H$_1$: $\sigma_g^2 \neq \sigma_d^2$ $(g \neq d)$. In this case the statistic was used:

$$U^2 = \frac{2.3026}{1+\dfrac{1}{3(k-1)}\displaystyle\sum_{i=1}^{k}\left(\frac{1}{n_i-1}-\frac{1}{n-k}\right)}\left[(n-k)\log\sigma^2-\sum_{i=1}^{k}(n_i-1)\log\sigma_i^2\right], \tag{7}$$

where

$$\sigma_i^2 = \frac{1}{n_i - 1} \sum_{j=1}^{n_i} (y_{ij} - \overline{y}_i)^2, \qquad i = 1, \ldots, k, \tag{8}$$

$$\sigma^2 = \frac{1}{n-k} \sum_{i=1}^{k} (n_i - 1)\sigma_i^2. \tag{9}$$

For given freedom degrees the critical value χ_α^2 needs to be read. If $U^2 > \chi_\alpha^2$ than the zero hypothesis should be rejected.

The results:

Table 3. The analysis results or variations in particular month differ. Separate results for day, evening and night are presented.

The source of volatility month	Statistics U^2 calculated	Table statistics 11 degrees of freedom, $\alpha = 0.05$
Day	72.22	19.675
Evening	114.56	19.675
Night	180.85	19.675

For each time the critical value is smaller than the calculated value. Hence the zero hypothesis H_0, stating that the variances of L_{eq} in different months differ relevantly, should be rejected.

After such analysis there are more questions coming: do averages for months differ relevantly from each other? To verify that hypothesis we use proper test depending on the count of the groups. Based on the above example we check if averages for different months differ form each other relevantly. Formally for verification we use a zero hypothesis stating that the averages L_{eq} in particular months are identical H_0: $m_1 = \ldots = m_{12}$ confronting the alternative that at least two of the averages differ from each other relevantly H_1: $m_d \neq m_g$.

The test used for this verification was F Snedecore test, for which the statistics is:

$$F = \frac{\dfrac{\sum_{i=1}^{k} n_i \left(\overline{y}_i - \overline{y}\right)^2}{k-1}}{\dfrac{\sum_{i=1}^{k} \sum_{j=1}^{n_i} \left(y_{ij} - \overline{y}_i\right)^2}{n-k}}, \tag{10}$$

where

$$\begin{aligned} &\frac{\sum\limits_{i=1}^{k} n_i \left(\overline{y}_i - \overline{y}\right)^2}{k-1} && \text{– average square value of cross group deviation;} \\ &\frac{\sum\limits_{i=1}^{k} \sum\limits_{j=1}^{n_i} \left(y_{ij} - \overline{y}_i\right)^2}{n-k} && \text{– average square value of inter group deviation} \\ & && \quad i = 1, \ldots, k \text{ with } n \text{ values of observation.} \end{aligned} \tag{11}$$

Table 4. The analysis results or variations in particular month differ. Separate results for day, evening and night are presented.

The source of volatility month	Statistics F calculated	Table statistics $\nu_1 = 11$, $\nu_2 = 283$, $\alpha = 0.05$
Day	3.18896	1.82
Evening	5.05960	1.82
Night	9.5633	1.82

It can be stated that the calculated values of the statistics are greater than the table statistics, which means that averages in particular month of the year are statistically different.

3. Final conclusions

Presented in this paper ideas represent possible methods of environmental acoustical monitoring data analysis. They define possible directions of research, which might be a useful tool for answering questions concerning the realization of control process of acoustical environment state. They are dedicated to estimation of long term noise level ratios and their uncertainty assessment. They generate directions useful for proper control method choice providing representative measurement estimates of long term noise ratios or the control timing choice for long term noise ratios L_{DEN} and L_N estimation.

The generate suggestions helpful for: choice of proper control method providing representativeness of measurement assessment of long term noise ratios or choice of control schedule for long term noise ratios L_{DEN} and L_N estimation.

References

[1] ANIS A. A., LLOYD E. H., *The expected value of the adjusted rescaled Hurst range of independent normal summands*, Biometrica, **63**, 283–298 (1976).

[2] BATKO W., BAL–PYRCZ R., *Analiza zmienności poziomu hałasu drogowego*, Materiały XXXV Zimowej Szkoły Zwalczania Zagrożeń Wibroakustycznych, Gliwice – Ustroń 26.02–2.03. 2007.

[3] BATKO W., *Modelowanie akustycznych zagrożeń środowiska dla potrzeb diagnostyki jego stanu*, Materiały XXXIII Sympozjum Diagnostyka Maszyn, CD, Węgierska Górka 2006.

[4] Directive 2002/49/EC of the European Parliament and the Council of 25 June 2002 relating to the assessment and management of environmental noise.

[5] KUFEL T., *Ekonometria. Rozwiązywanie problemów z wykorzystywaniem programu Gretl*, PWN, Warszawa 2004.

[6] MORZYK M., *Analiza struktury monitorowanego hałasu dla potrzeb estymacji długookresowych wskaźników hałasu,* MSc. Thesis, AGH 2006.

[7] PAWŁOWSKI Z., *Wstęp do statystyki matematycznej*, PWN, Warszawa 1965.

[8] TALAGA L., ZIELIŃSKI Z., *Analiza spektralna w modelowaniu ekonometrycznym*, PWN, Warszawa 1986.

[9] STECZKOWSKI J., ZELIAŚ A., *Analiza wariancyjna i kowariancyjna w badaniach ekonomicznych*, PWN, Warszawa 1982.

[10] ZIELIŃSKI Z., *Metody analizy dynamiki i rytmiczności zjawisk gospodarczych*, PWN, Warszawa 1979.