Acoustic and Biomedical Engineering

Application of the Bootstrap Estimator for Uncertainty Analysis of the Long-Term Noise Indicators

W. Batko* and B. Stępień

Department of Mechanics and Vibroacoustics, AGH University of Science and Technology

Mickiewicza 30 Av., 30-059 Kraków, Poland

In the process of environmental noise hazards assessment besides estimation of uncertainty measurement of control results there should be made errors analysis related to estimation processes to long-term noise indicators. the condition of correct quantification of uncertainty budget components in the control process is using proper analysis method. The problem rest on determining density function of probability distribution of long-term noise indicators. in order to several conditioning characteristic for the problem it can not be to solved by classic estimation analysis applied in statistical researches, without different reservations. There was formulated the estimation idea of seeking density function of long-term noise indicators distribution by bootstrap method, which does not generate limitations for form and properties of analyzed statistics. There was presented theoretical basis of the proposed method, and the example of calculation process which make possible determining searching estimators of expected value and variance of long-term noise indicators L_{DEN} and L_{N}. The illustration for indicated solutions and usefulness analysis was continuous monitoring results of a traffic noise recorded on one of the main arteries of Kraków.

PACS numbers: 43.50.Rq, 43.50.Yw

1. Introduction

The European Union Directives [1] given in papers [2, 3], as well as the domestic regulations [4] concerning the long-term policy of environment protection against noises, formulated the need of estimating the long-term noise indicators. They are determined by annoyance values of the long-term average sound level A in the day-evening-night periods (L_{DEN}) and night periods (L_{N}) in dB. Their values constitute the basis for preparing the noise maps, in places where the acoustic climate is being assessed, and the selection of activities aimed at the prevention and limitation of harmful noise effects in the environment.

The long-term average sound levels A (L_{DEN}) and (L_{N}) in dB are determined on the basis of noise annoyance indicators $L_{\mathrm{DEN},i}$ for $i = 1, 2, \ldots, 365$ of all days in the calendar year at the day-evening-night periods:

$$L_{\mathrm{DEN},i} = 10\log(\frac{1}{24}(12 \times 10^{0,1 L_{\mathrm{D},i}} + 4 \times 10^{0,1(L_{\mathrm{E},i}+5)} +8 \times 10^{0,1(L_{\mathrm{N},i}+10)})/24), \quad (1)$$

where: $L_{\mathrm{D},i}$ — day sound A level, determined from the day–time noise exposure i.e. from 6:00 a.m. to 6:00 p.m., [dB],

$L_{\mathrm{E},i}$ — evening sound A level, determined from the noise exposures from 6:00 p.m. to 10:00 p.m., [dB],

$L_{\mathrm{N},i}$ — night sound A level, determined for the night periods i.e. from 10:00 p.m. to 6:00 a.m., [dB],

and night periods $L_{\mathrm{N},i}$ for $i = 1, 2, \ldots, 365$ determined by relation [5]:

$$L_{\mathrm{N},i} = 10\log\left[\frac{1}{K}\sum_{i=1}^{\mathrm{K}} 10^{0,1(L_{\mathrm{Aeq,T}})_i}\right], \quad (2)$$

where: K — sample size, $(L_{\mathrm{Aeq,T}})_i$ — equivalent sound level for the ith sample, [dB].

Estimation of the long-term indicators of the acoustic hazard for the environment L_{DWN} and L_{N} being the average value calculated from all calendar days:

$$L_{\mathrm{DEN}} = 10\log\left[\frac{1}{365}\sum_{i=1}^{365} 10^{0,1 L_{\mathrm{DEN},i}}\right], \quad (3)$$

$$L_{\mathrm{N}} = 10\log\left[\frac{1}{365}\sum_{i=1}^{365} 10^{0,1 L_{\mathrm{N},i}}\right], \quad (4)$$

forming a set of two indicators $L_{\mathrm{Aeq.LT}} = \{L_{\mathrm{DEN}}, L_{\mathrm{N}}\}$, requires an access to the results of the whole year sound level monitoring.

In practice, it is not possible to meet such requirement. Therefore estimations of indicators are usually done on the basis of reference model methods, assigned to various types of noises: road traffic, railway, air or industrial. However, as it results from numerous researches, such estimation method is carrying a high uncertainty [6]. This is caused by: difficulties in a precise determination of input data for calculation models, errors of the calculation

* corresponding author; e-mail: batko@uci.agh.edu.pl

method generated by the structure of the applied model or conditions of their validation [7]. Model calibrations realised on the basis of the limited sample of results of a constant acoustic monitoring, resulting from the adopted time schedule of random sampling of the environment are necessary in such situations.

The problem of the necessity of performing the estimation — on the basis of limited random sample — of the expected value and the variance of the controlled long-term noise indicators is related to this task. The binding legal regulations impose the obligation of performing not only estimations of noise hazard indicators for the environment but also analyses of the uncertainty budget of those indicators. The component related to the probability density function of the long-term indicators of noise hazards in the environment i.e. the estimation of type A uncertainty — expressed by standard deviation of the mean [8] is the substantial component of the combined uncertainty. In order to estimate this component one must know the probability distribution of the long-term indicators of the noise hazards in the environment.

Finding the probability density function is difficult — as presented in papers [9, 10] — due to: a lack of knowledge of the distribution class to which probability distributions of this indicators belong, a limited sample size and inadequacy of the classis estimation solutions. The authors proposed, in their prior works, kernel estimators [11], allowing to obtain estimators for the relatively large samples size control. However, there is a need for methods which — in the case of small samples size control — could predict how reliable is assuming of an asymptotic character if the sample is small.

The possibility provided by resampling techniques i.e. by methods based on the multiple sampling, will be presented in the hereby paper. An application of this idea, based on the bootstrap method was introduced in the year 1979 by Bradley Efron [12], constitutes a premise for increasing the estimation accuracy of the expected value and variances of the long-term noise indicators $L_{\mathrm{Aeq.LT}}$ on the bases of the results obtained by sampling inspections. Discussion of the method, together with the example illustrating its functioning, will be contained in the present paper. The reference base constitute the results of the constant noise monitoring recorded in one of the main arteries of Kraków, Poland.

2. Assumptions and ideas of the bootstrap

Applying the bootstrap method in the investigation of the population, in consideration of one-dimensional random variable X, we assume that [13, 14]:

— we have the "primary", finite simple random sample $x = (X_1, X_2, \ldots, X_n)$ from the investigated population, while values $(x_1, x_2, \ldots, x_n)$ constitute the realisation of random variable $X = \{X_i, i = 1, 2, \ldots, n\}$,

— we do not know the probability distribution F of the investigated random variable in the parent population,

— $R(x, F)$ means certain statistics determined on the sample space.

Standard steps at the basic bootstrap procedure are as follows [13, 14]:

1. We design the probability distribution by means of the following function
$$P\left(X_{\mathrm{B}} = x_k\right) = \frac{1}{n} \quad \text{for} \quad k = 1, 2, \ldots, n \tag{5}$$
called the bootstrap distribution from sample and denoted by $\hat{F}$, where n is the sample size.

2. We sampling, independently, according to distribution of values $(x_1^*, x_2^*, \ldots, x_n^*)$, which are treated as the realisation of variable $x^* = (X_1^*, X_2^*, \ldots, X_n^*)$ and it is called the bootstrap sample.

3. Distribution of statistics R is approximated by means of the bootstrap distribution:

$$R^* = R\left(x^*; \hat{F}\right). \tag{6}$$

Distribution of variable R^* is approximated by means the Monte Carlo method. The histogram of statistics R^* is determined on the bases of the bootstrap samples repeated N-times.

3. Point estimation of distribution parameters with the application of the bootstrap method

Point estimation of an unknown distribution parameter β of the investigated variable is based on assuming that the estimator value of this parameter at the given sample is its estimation. If the statistics R is this estimator, then the statistics R^* (6) value can be adopted as the assessment of unknown parameter β. Applying the Monte Carlo method for the bootstrap distribution of statistics R, the mean of values $r_1^*, r_2^*, \ldots, r_{\mathrm{N}}^*$ — being values of variable R^*, obtained by the bootstrap sample repeated N-times of values $(x_1^*, x_2^*, \ldots, x_n^*)$ can be assumed as assessment of parameter β [13]. Thus, the assessment of parameter β takes the value [14, 15]:

$$\bar{r}^* = \frac{1}{N} \sum_{k=1}^{N} r_k^*. \tag{7}$$

On the bases of values $(r_1^*, r_2^*, \ldots, r_{\mathrm{N}}^*)$ obtained from successive replicates of the bootstrap sample we can approximate the standard error (standard deviation) of parameter β [14]:

$$se_{\mathrm{B}} = \sqrt{\frac{\sum_{k=1}^{N} \left(r_k^* - \bar{r}^*\right)^2}{N-1}}. \tag{8}$$

4. Estimation of characteristics of the long-term noise indicators

Point estimation of characteristics of the long-term noise indicators $L_{\mathrm{Aeq.LT}}$ (an expected value and type A uncertainty) was done on the bases of sampling inspections — when utilising the classical and the bootstrap method — in the simulation experiment.

The value determined on the basis of the results recorded during the whole calendar year by the continuous noise monitoring system installed in one of the main arteries in Kraków was taken as the measured value of indicators $L_{\mathrm{Aeq.LT}}$. Due to this there will be a possibility to determine accurately the assessment error of the expected value — by means of the selected methods.

Type A uncertainty of the long-term noise hazard indicators was also determined with the application of two types of standard deviation estimators:

— classical, which assumes the normality of results of the long-term noise indicators,

— bootstrap, which does not generate limitations regarding forms and properties of the investigated statistics.

In our case the investigated population constitute the results of 24-hours average sound A level and the night sound A level determined on the basis of constant monitoring. Simple random samples of size $n = 5$, 10, 30, 50 (simulating the number of controlled days on the basis of which the levels: L_{DEN} and L_{N} will be estimated) were sampling from the above mentioned population. Utilising the application developed in the Matlab package the expected values and type A uncertainties of the long-term noise indicators were assessed by the classical and bootstrap method for individual cases.

The expected values of indicators — in the classical approach — were determined by Eq. (9):

$$\bar{L}_{\mathrm{Aeq.LT}} = 10\log\left(\frac{1}{n}\sum_{i=1}^{n} 10^{0.1 L_{\mathrm{Aeq.LT,i}}}\right), \tag{9}$$

where: n — sample size, $L_{\mathrm{Aeq.LT,i}}$ — index level for the ith sample, [dB].

The type A uncertainty of the long-term noise hazard indicators were determined by Eq. (10):

$$s\left(\bar{L}_{\mathrm{Aeq.LT}}\right) = \sqrt{\frac{\sum_{i=1}^{n}\left(L_{\mathrm{Aeq.LT,i}} - \bar{L}_{\mathrm{Aeq.LT}}\right)^2}{n\,(n-1)}}\,. \tag{10}$$

Algorithm of the characteristics estimation of the long-term noise indicators performed by the classical method is presented in Fig. 1.

Values of the estimated noise indicators were determined — for the needs of the performed simulation experiment — on the basis of $N = 500$, 1000, 5000, 200000 bootstrap replications. The sequence: $L^*_{\mathrm{Aeq.LT,1}}$, $L^*_{\mathrm{Aeq.LT,2}}$, ..., $L^*_{\mathrm{Aeq.LT,N}}$ was obtained as a result of values: $L_{\mathrm{Aeq.LT,1}}$, $L_{\mathrm{Aeq.LT,2}}$, ..., $L_{\mathrm{Aeq.LT,n}}$ generated N-times according to Eq. (5), it means sampling independently from samples of size n and calculating each time

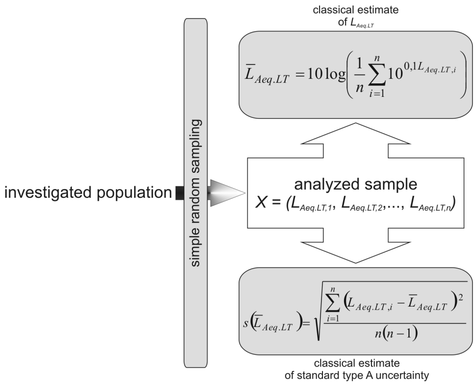

Fig. 1. Algorithm of the estimation of the expected value and the type A uncertainty of the long-term noise indicators — performed by the classical method.

the statistics value $L^*_{\mathrm{Aeq.LT,k}}$ from Eq. (9). This sequence was used for determining histograms, which illustrate the bootstrap distribution of L_{DEN} and L_{N}. The bootstrap estimation of the long-term noise indicators are:

$$\bar{L}^*_{\mathrm{Aeq.LT}} = \frac{1}{N}\sum_{k=1}^{N} L^*_{\mathrm{Aeq.LT,k}}\,, \tag{11}$$

where: N — number of bootstrap replications, $L^*_{\mathrm{Aeq.LT,k}}$ — level of the k^{th} bootstrap estimate of index $L_{\mathrm{Aeq.LT}}$.

The bootstrap estimation of the type A uncertainty can be determined as follows:

$$s_{\mathrm{B}}\left(\bar{L}^*_{\mathrm{Aeq.LT}}\right) = \sqrt{\frac{\sum_{k=1}^{N}\left(L^*_{\mathrm{Aeq.LT,k}} - \bar{L}^*_{\mathrm{Aeq.LT}}\right)^2}{N-1}}\,. \tag{12}$$

Algorithm of the estimation of the characteristics of the long-term noise indicators performed by the bootstrap method is presented in Fig. 2.

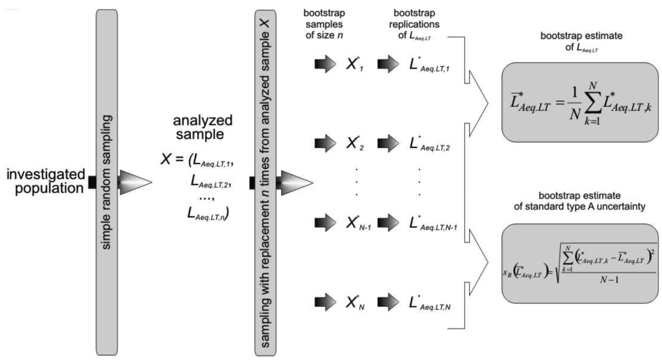

Fig. 2. Algorithm of the estimation of the expected value and the type A uncertainty of the long-term noise indicators — performed by the bootstrap method.

5. Results of the experiment

The long-term environmental noise hazard indicators i.e. the day-evening-night level (L_{DEN}) and the night level

($L_{\rm N}$) were determined on the basis of the results recorded in the year 2005 by the constant acoustic monitoring station installed at the Krasiński Avenue in Kraków.

The expected value and the type A uncertainty were determined by means of two methods, the classical one and the bootstrap included to methods of repeated drawings. The proceedings of the experiment are presented in Figs. 1 and 2, respectively.

There are certain discrepancies between the known values of $L_{\rm DEN}$ and $L_{\rm N}$, and the estimated values (classical and bootstrap), however only four bootstrap estimate had a significant error (marked bold in Table I). It might be assumed that in this case the number of replications was too small for the case $L_{\rm DEN}$ for $n = 50$, however case $n = 10$ is one of the special cases described below. A higher number of replications substantially improved the estimation results. Generally, it can be noticed that estimates performed for larger samples are burdened with smaller errors.

The results of estimates of the type A uncertainty of the long-term noise indicators shown in Table II were determined on the basis of the same random samples as the expected values of indicators shown in Table I. It can be seen that type A uncertainties estimated by the bootstrap method are smaller than the ones obtained by the classical estimation method. This indicates that values of the long-term noise indicators determined by the bootstrap method are more accurate.

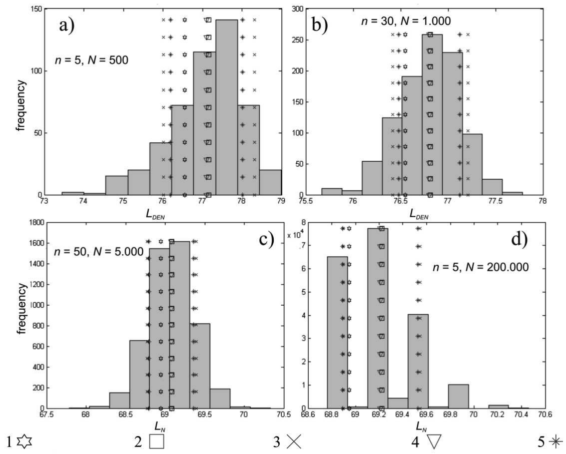

Fig. 3. The bootstrap estimates of the probability density function of the long-term noise indicators Legend: 1 — measured value of index $L_{\rm Aeq.LT}$, 2 — classical estimate of index $\bar{L}_{\rm Aeq.LT}$, 3 — classical interval $\bar{L}_{\rm Aeq.LT} \pm s(\bar{L}_{\rm Aeq.LT})$, 4 — bootstrap estimate of index $\bar{L}^*_{\rm Aeq.LT}$, 5 — bootstrap interval $\bar{L}^*_{\rm Aeq.LT} \pm s_{\rm B}(\bar{L}^*_{\rm Aeq.LT})$.

Examples of histograms that were obtained during the experiment of various n and various bootstrap samples replications N are presented in Fig. 3. Shadowed fields in Tables I and II were taken into account when drawing Fig. 3.

Examples given in Table III, where all bootstrap estimates are characterised by larger errors than the classical ones, are the special cases. The measured value of indicators is neither contained within a range: $\bar{L}_{\rm Aeq.LT} \pm s(\bar{L}_{\rm Aeq.LT})$ estimated by the classical method nor within a range: $\bar{L}^*_{\rm Aeq.LT} \pm s_{\rm B}(\bar{L}^*_{\rm Aeq.LT})$ estimated by the bootstrap method. This is shown in Fig. 4.

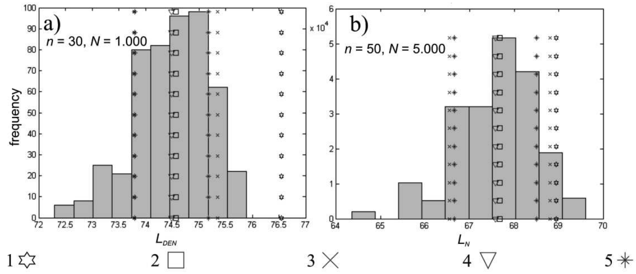

Fig. 4. Bootstrap estimates of the probability density function of the long-term noise indicators for the special cases Legend: 1 — measured value of index $\bar{L}_{\rm Aeq.LT}$, 2 — classical estimate of index $\bar{L}_{\rm Aeq.LT}$, 3 — classical interval $\bar{L}_{\rm Aeq.LT} \pm s(\bar{L}_{\rm Aeq.LT})$, 4 — bootstrap estimate of index $\bar{L}^*_{\rm Aeq.LT}$, 5 — bootstrap interval $\bar{L}^*_{\rm Aeq.LT} \pm s_{\rm B}(\bar{L}^*_{\rm Aeq.LT})$.

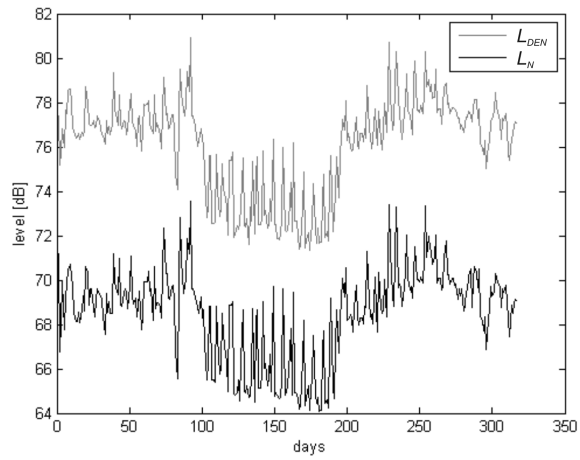

Fig. 5. Time history of the long-term noise indicators during the year.

Those cases are the effects of drawing simple samples of elements characterised by significantly lower values than the measured value. Time-history of indicators' changes during the calendar year is presented in Fig. 5. A similar situation can also occur during the actual measurements since a person performing sampling inspections is not able to determine whether the assumed time schedule will be representative for the whole year. Special cases of the numerical experiment presented in Fig. 4 are shadowed in Tables III and IV.

TABLE I

Estimates of the expected values of the long-term noise indicators.

index	measured value [dB]	sample size n	classical estimate [dB]	bootstrap estimate $N = 500$ [dB]	bootstrap estimate $N = 1000$ [dB]	bootstrap estimate $N = 5000$ [dB]	bootstrap estimate $N = 200000$ [dB]
L_{DEN}	76.5461	5	77.1623	77.1020	77.0696	77.0953	77.0747
		10	76.5619	**76.5237**	**76.5247**	76.5318	**76.5291**
		30	76.8123	76.7789	76.8022	76.7977	76.8009
		50	76.6152	**76.6176**	76.6045	76.6085	76.6064
L_{N}	68.9467	5	69.2217	69.1945	69.1932	69.2102	69.2097
		10	69.1873	69.1038	69.1608	69.1166	69.1315
		30	69.1123	69.0661	69.0985	69.1005	69.0992
		50	69.0822	69.0712	69.0686	69.0764	69.0738

TABLE II

Estimates of the type A uncertainty of the long-term noise indicators.

index	sample size n	classical estimate [dB]	bootstrap estimate $N = 500$ [dB]	bootstrap estimate $N = 1000$ [dB]	bootstrap estimate $N = 5000$ [dB]	bootstrap estimate $N = 200000$ [dB]
L_{DEN}	5	1.1519	0.9099	0.9110	0.8879	0.8978
	10	0.7126	0.5258	0.5233	0.5167	0.5266
	30	0.3970	0.3015	0.3220	0.3281	0.3250
	50	0.3249	0.2663	0.2793	0.2770	0.2754
L_{N}	5	0.3249	0.2950	0.3116	0.3136	0.3166
	10	0.7849	0.7682	0.6881	0.7206	0.7149
	30	0.3837	0.3411	0.3404	0.3437	0.3468
	50	0.3110	0.2803	0.2796	0.2834	0.2841

TABLE III

Special cases of estimation the expected value of the long-term noise indicators.

index	measured value [dB]	sample size n	classical estimate [dB]	bootstrap estimate $N = 500$ [dB]	bootstrap estimate $N = 1000$ [dB]	bootstrap estimate $N = 5000$ [dB]	bootstrap estimate $N = 200000$ [dB]
L_{DEN}	76.5461	5	74.5715	**74.4916**	**74.5270**	**74.5291**	**74.5202**
L_{N}	68.9467	5	67.6773	**67.5832**	**67.6060**	**67.5726**	**67.5815**

TABLE IV

Estimates of the type A uncertainties of the long-term noise indicators for the special cases.

index	sample size n	classical estimate [dB]	bootstrap estimate $N = 500$ [dB]	bootstrap estimate $N = 1000$ [dB]	bootstrap estimate $N = 5000$ [dB]	bootstrap estimate $N = 200000$ [dB]
L_{DEN}	5	0.7844	0.6881	0.6741	0.6613	0.6689
L_{N}	5	1.1341	0.9142	0.9049	0.9208	0.9253

6. Conclusions

Methodological backgrounds of the utilisation of the resampling technique (the bootstrap method) for the estimation of the probability density function of the long-term indicators of the environment acoustic hazards are presented in the hereby paper. They provide the new estimation perspective for the uncertainty of results as well as the more strict integration of the applied assumptions with the practical conditions of their applications.

The simulation experiment of the determination of the expected value and the type A uncertainty was carried out by means of two methods, classical and bootstrap, and its results presented. The results that were obtained allow to state, that:

— the bootstrap estimates were characterised by smaller errors (the difference between the measured value and the estimated one was smaller),

— the type A uncertainty determined by the bootstrap method is of a smaller value, what indicates that the estimation of the expected value of the long-term noise indicators is more accurate.

The possibility of using the bootstrap method for the determination of the type A uncertainty as well as the expected value of the long-term noise indicators seems to be a promising calculation tool. This widens the existing calculation algorithms. Interpretation assumptions accompanying this method (among others: a lack of limitations concerning shape and character of the investigated statistics as well as the primary sample size), made it more likelihood than the classical estimation analysis applied up to the present.

Acknowledgments

This work has been supported by the Polish Ministry of Science and Higher Education from means on science in 2009–2011 (research project no. NN504345636).

References

[1] *Directive 2002/49/WE of the European Parliament and of the Council of 25 June 2002*, relating to the assessment and management of environmental noise, Official Journal of the European Communities 18.07.2002.

[2] R. Kucharski, *Archives of Acoustics* **32 2**, 293 (2007).

[3] D.I. Popescu, *Archives of Acoustics* **32 2**, 329 (2007).

[4] *Environmental protection law*, Law of 27.04.2001, (in Polish), Official Diary 2001/62/627 with amendments.

[5] *Guidelines for developing of acoustic maps* (in Polish), Ed. R.J. Kucharski, Institute of Environmental Protection, Warszawa 2006.

[6] *Imagine Project IMA32TR-040510-SP08*, available from `http://www.imagine-project.org` .

[7] *Good Practice Guide for Strategic Noise Mapping and the Production of Associated Data on Noise Exposure*, version 2, European Commission Working Group Assessment of Exposure to Noise (WG-AEN) 2006.

[8] *Guide to the Expression of Uncertainty in Measurement*, ISO 1995.

[9] W. Batko, B. Stępień, in: *35th Winter School on Vibroacoustical Hazards Suppressions*, Ed. M. Roczniak, Upper Silesian Division of the Polish Acoustical Society, Gliwice 2007, p. 5.

[10] T. Wszołek, M. Kłaczyński, *Archives of Acoustics* **31**, 4 (supplement), 311 (2006).

[11] W. Batko, B. Stępień, *Archives of Acoustics* **34 3**, 295 (2009).

[12] B. Efron, *Annals of Statistics* **7**, 1 (1979).

[13] Cz. Domański, K. Pruska, *Non-classical statistical methods* (in Polish), Polish Economics Publishers, Warszawa 2000.

[14] B. Efron, R.J. Tibshirani, *An Introduction to the Bootstrap*, Chapman & Hall/CRC, New York 1993.

[15] P. Koronacki, J. Mielniczuk, *Statistics for students of technical and natural discipline* (in Polish), WNT, Warszawa 2004.