Research Paper

Determining the Number of Measurements and Bootstrap Samples Required to Estimate of Long-Term Noise Indicators

Bartłomiej STĘPIEŃ

AGH University of Science and Technology
Department of Mechanics and Vibroacoustics
Kraków, Poland; e-mail: Bartlomiej.Stepien@agh.edu.pl

The minimum size of the bootstrap algorithm input parameters have been determined for estimation of long-term indicators of road traffic noise. Two independent simulation experiments have been performed for that purpose. The first experiment served to determine the impact of original random sample size, and the second to determine the impact of number of the bootstrap replications on the accuracy and uncertainty of estimation of long-term noise indicators. The inference has been carried out based on results of non-parametric statistical test at significance level $\alpha = 0.05$. The simulation experiments have shown that estimation of long-term noise indicators with uncertainty below ± 1 dB(A) requires all-day noise measurements during three randomly selected days during the year in a dense urban development. The maximum size of original random sample should not exceed $n = 50$ elements. The minimum number of bootstrap replications necessary for estimation should be $B = 5000$. The data used to the simulation experiments and carry out the analysis were results of continuous monitoring of road traffic noise recorded in 2009 in one of the main arteries of Krakow in Poland.

Keywords: bootstrap; bootstrap replications; long-term noise indicators; number of measurements; uncertainty; accuracy.

1. Introduction

The basis for creating noise maps for sites under protection are the values of the two basic long-term noise indicators: L_{DEN} and L_{N} (European Parliament, 2002). Any plans to prevent and reduce the harmful effects of noise in the environment are then associated with their values. These indicators characterise the acoustic climate over a long period. Most often it is assumed that this is one full calendar year, so values of the indicators depend on many factors (i.e. traffic intensity, structure of the vehicle stream, average vehicle velocity, type and technical condition of the road surface, distance of the nearest buildings from the road edge, technical condition of the vehicles). Estimation of long-term noise hazard indicators requires access to results of an all-year-long sound level monitoring program. In practice, it is almost impossible to meet such a requirement. Therefore estimations of indicators are usually done on the basis of a highly limited and correlated random sample. Sample size n is very small and ranges from few to a dozen or so elements. They are obtained as results of environmental sampling inspections (SCHOMER, DEVOR, 1981; GAJA *et al.*, 2003; ROMEU *et al.*, 2006).

The necessity of validation of the obtained results, which requires the analysis of uncertainty budget of estimation, is connected with the process of calculating the average long-term noise indicators determined by values L_{DEN} and L_{N}. Overrating or underestimating values of this indicators can have notable social and financial consequences.

Many authors have previously raised the question of number of control days during which the measurements should be conducted in order to determine the long-term noise indicators with sufficient accuracy. The authors from Italy (BRAMBILLA *et al.*, 2007) showed that despite the time variability of the sources and the sound propagation conditions it was found that, for road traffic noise and railway noise, sampling the dataset on 5–7 non-sequential days allowed for an estimate of the L_{DEN} annual indicator to within ± 1 dB(A). However, based on studies carried out during the whole one year, in Lithuania (JAGNIATINSKIS,

FIKS, 2014), the authors showed that the lowest uncertainty values of L_{DEN} occurred when the total measuring time was 7 consecutive days, provided that the measurement was done under normal weather and source emission conditions.

An essential component of such budget is the type A standard uncertainty defined as the standard deviation of the mean from the inspections results. The rules given in the ISO/IEC (2008) are based on the point estimation methods and commonly used in the calculations. They are based on the classic variance estimators on the condition of assigning the normal distribution and lack of correlation between elements of the sample as well as adequate sample size and observation equivalence to random results of the sampling inspections. Results of acoustical measurements usually do not meet these assumptions (DON, REES, 1985; GIMÉNEZ, GONZÁLEZ, 2009; PRZYSUCHA *et al.*, 2020).

For reasons presented above, there is a need to search for methods to more accurately determine the expected value of the noise indicators (MAKAREWICZ, 2011a) and also other acoustic parameters and their uncertainty (HEISS, KRAPF, 2007; MAKAREWICZ, 2011b; PILCH, 2018). That is why nonparametric methods of statistical inference which deviate from the assumptions of the classical statistical model, i.e. kernel density estimator (STĘPIEŃ, 2016; HUANG, ELHILALI, 2017) bootstrap resampling method (RUGGIERO *et al.*, 2016), Bayesian inference (SCHUMACHER *et al.*, 2012; STĘPIEŃ, 2018), and interval arithmetic (BATKO, PAWLIK, 2012; 2013) gain a wider use in acoustics. These methods are used mainly due to the accompanying interpretation assumptions, the most important of which include absence of limitations in terms of form and properties of the studied statistic and the sample size.

The analysis of papers published in recent years indicates a growing recognition among the researchers for the bootstrap resampling method. It is used with success in point (BATKO, STĘPIEŃ, 2010) and interval estimation (STĘPIEŃ, 2017) of the noise indicators expected value and uncertainty (FARRELLY, BRAMBILLA, 2003; MATEUS *et al.*, 2015), as well as in planning the measurement strategies (LIGUORI *et al.*, 2017a; 2017b). It is often used in statistical analysis of sound measurement results (BRAMBILLA *et al.*, 2015; VOS, 2017).

The papers have used various sizes of input parameters of the bootstrap algorithm which include the original sample size n, and the number of bootstrap replications B. The issue of original sample size can be equated with the number of control days or measurement events used to determine the noise indicators. MATEUS *et al.* (2015) have analysed the impact of the number of bootstrap samples on the estimation accuracy of A-weighted equivalent sound level for day, evening and night periods. The authors stated that the

number of repetitions $B = 10^5$ for the bootstrap algorithm satisfies the criteria (standard deviation equal to 0.01) for all three reference periods while still being practical in terms of the computation time required.

The reasons presented above indicate the need to specify the minimum size of the bootstrap algorithm input parameters, i.e. original sample size n, and the number of bootstrap replications B, in order to determine the expected values of long-term noise indicators L_{DEN} and L_{N} of road traffic noise with required accuracy which has been presented in the paper.

Discussion of the algorithms, together with an example illustrating their functioning, will be presented further in this paper. The reference base comprises the results of the constant noise monitoring recorded in 2009 in one of the main arteries of Krakow, Poland.

2. Bootstrap method background

Consider an observed random sample $\mathbf{x} = (x_1, x_2, ..., x_n)$ from an unknown probability distribution F with an intent to estimate a parameter of interest $\theta = t(F)$ on the basis of $\mathbf{x}$. For this purpose, let an estimate $\widehat{\theta} = s(\mathbf{x})$ from $\mathbf{x}$ be calculated.

The bootstrap method was introduced in 1979 by EFRON (1979) as a computer-based method for estimating the standard error of $\widehat{\theta}$. The bootstrap estimate of standard error requires no theoretical calculations and is available no matter how mathematically complicated the estimator $\widehat{\theta} = s(\mathbf{x})$ may be.

Bootstrap methods depend on the concept of a bootstrap sample. Let $\widehat{F}$ be the empirical distribution, assigning probability $1/n$ to each of the observed values $x_i, i = 1, 2, ..., n$. A bootstrap sample is defined as a random sample of size n drawn from $\widehat{F}$, say $\mathbf{x}^{\text{b}} = (x_1^{\text{b}}, x_2^{\text{b}}, ..., x_n^{\text{b}})$ (EFRON, TIBSHIRANI, 1993)

$$\widehat{F} \to (x_1^{\text{b}}, x_2^{\text{b}}, ..., x_n^{\text{b}}). \tag{1}$$

The symbol "b" indicates that $\mathbf{x}^{\text{b}}$ is not the actual data set $\mathbf{x}$, but rather a resampled version of $\mathbf{x}$.

Symbolic expression (1) can be also verbalised as follows: the bootstrap data points $x_1^{\text{b}}, x_2^{\text{b}}, ..., x_n^{\text{b}}$ are a random sample of size n drawn with replacement from the population of n objects $(x_1, x_2, ..., x_n)$. The bootstrap data set $(x_1^{\text{b}}, x_2^{\text{b}}, ..., x_n^{\text{b}})$ consists of elements of the original data set $(x_1, x_2, ..., x_n)$.

Corresponding to a bootstrap data set $\mathbf{x}^{\text{b}}$ is a bootstrap replication of $\widehat{\theta}$

$$\widehat{\theta}_b = s(\mathbf{x}^{\text{b}}). \tag{2}$$

The quantity $s(\mathbf{x}^{\text{b}})$ is the result of applying to $\mathbf{x}^{\text{b}}$ the same function $s(\cdot)$ as this applied to $\mathbf{x}$.

2.1. Point estimation of distribution parameters

Point estimation of an unknown distribution parameter θ of the examined variable is based on as-

suming that the estimator value of this parameter at the given sample is its estimation. By applying the Monte Carlo method to the bootstrap, a bootstrap sample B is generated. The bootstrap samples are generated from the original data set (analysed sample). Each bootstrap sample has n elements generated by sampling with replacement n times from the analysed sample. Bootstrap replications $\widehat{\theta}_1, ..., \widehat{\theta}_b, ..., \widehat{\theta}_B$ are obtained by calculating the value of the statistics $s(\mathbf{x})$ on each bootstrap sample. The mean of these values can be assumed to be an assessment of parameter θ. Thus, the assessment of parameter θ can be expressed as (EFRON, TIBSHIRANI, 1993)

$$\overline{\theta}_B = \frac{1}{B} \sum_{i=1}^{B} \widehat{\theta}_b. \tag{3}$$

The bootstrap estimate of the standard error is the standard deviation of the bootstrap replications (EFRON, TIBSHIRANI, 1993):

$$\widehat{s}_B = \sqrt{\frac{\sum_{b=1}^{B} (\overline{\theta}_B - \widehat{\theta}_b)^2}{B-1}}. \tag{4}$$

Further, the bootstrap estimate of bias $\widehat{b}_B$ based on the B replications is defined by

$$\widehat{b}_B = \overline{\theta}_B - \widehat{\theta}, \tag{5}$$

where $\overline{\theta}_B$ is bootstrap estimate of parameter θ, and $\widehat{\theta}$ is estimate of parameter θ. The value of $\widehat{\theta}$ may be calculated from the original sample $\mathbf{x}$ or may differ from $\widehat{\theta} = s(\mathbf{x})$, e.g. it is determined from the population (EFRON, TIBSHIRANI, 1993). Note that both $\widehat{s}_B$ and $\widehat{b}_B$ can be calculated from the same set of bootstrap replications.

3. Research material

With the intention to solve increasing environmental noise problems and broaden the knowledge about acoustic phenomena observed in the Krakow urban area, a system of continuous noise monitoring has been put into operation as early as in the year 1996.

This measuring system is composed by the real time noise analyser Nor 110 (Norsonic, Norway) equipped with preamplifier and $1/2$ inch microphone with outdoor kit. This system was also equipped with an automatic calibration function. The system of continuous noise monitoring allows for simultaneous recording and analysis of the acoustic signal. The value of A-weighted sound pressure level (L_{pA}) is a basic parameter recorded by the system in an 1 second resolution.

The solution was implemented by the Małopolskie Voivodeship Environment Protection Inspectorate in co-operation with academics from AGH-UST's Department of Mechanics and Vibroacoustics and

Department of Robotics and Mechatronics. Location for the measuring station was selected bearing in mind the necessity to diagnose the acoustic climate in the vicinity of the most crowded traffic arteries in Krakow. The selected street crosses a dense urban development area. The measuring probe is situated in the middle of the green median belt separating two carriageways of the road with three lanes in each direction, with the 24-hour average traffic density about 3000 vehicles per hour. The average traffic density in the day period (from 06:00 a.m. to 06:00 p.m.) was about 3800 vehicles per hour, in the evening period (from 6:00 p.m. to 10:00 p.m.) was about 4500 vehicles per hour and in the night period (from 10:00 p.m. to 06:00 a.m.) was about 800 vehicles per hour. In the traffic stream approximately 6.5% of heavy vehicles was detected. The average speed of light vehicles was about 55 km/h, whereas heavy vehicles – about 30 km/h. Traffic intensity measurements on this section of road were made in 2009 by the Management of Municipal Infrastructure and Transport in Krakow.

The determination of useful (minimum) size of the bootstrap algorithm input parameters was carried out with the use of actual measurement results. To this end, A-weighted sound levels recorded by the above-mentioned noise monitoring system throughout the year 2009 were used. The analysis covered a total of 334 days of the year for which complete 24-hour-long records of the A-weighted equivalent sound levels were available. For the remaining days, the daily records were incomplete or did not exist at all.

On the grounds of the recorded A-weighted sound levels, values of 24-hour day-evening-night sound levels $L_{\mathrm{DEN},i}$ and night-time sound levels $L_{\mathrm{N},i}$ were calculated which constituted the examined populations with size $t = 334$. The values were used to determine long-term (annual) noise indicators. The obtained value of the long-term average day-evening-night sound level is $L_{\mathrm{DEN}} = 78.5$ dB(A) with the standard deviation $\sigma(L_{\mathrm{DEN},i}) = 0.8$ dB(A). The long-term night-time sound level was also determined as equalling $L_{\mathrm{N}} = 70.8$ dB(A) together with its standard deviation $\sigma(L_{\mathrm{N},i}) = 0.9$ dB(A).

Time plots of these quantities throughout the year are presented in Fig. 1. The calculated skewness values, which are 1.05 for $L_{\mathrm{DEN},i}$ and 0.37 for $L_{\mathrm{N},i}$, confirms the fact that examined populations have positively-skewed distributions. Kurtosis (excess kurtosis) for these populations is 7.33 (4.33) and 3.60 (0.60) for $L_{\mathrm{DEN},i}$ and $L_{\mathrm{N},i}$, respectively. The obtained values show that distributions of the examined populations are not normal.

A number of normality test have been carried out with the objective to confirm that the analysed population did not come from any normal distribution. The analysis included performing the Shapiro-Wilk test, the Jarque-Bera test, the Lilliefors test, and the
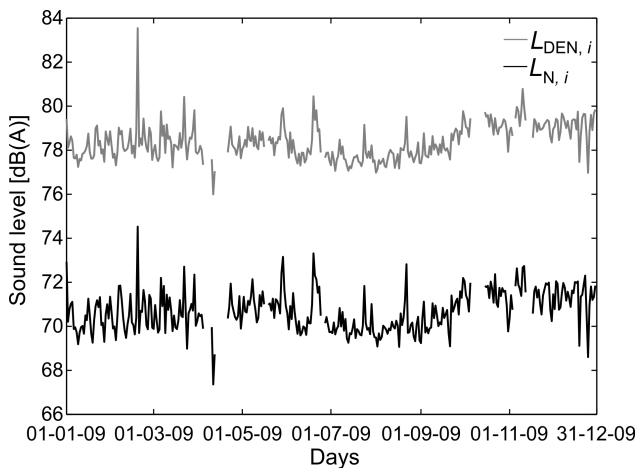
Fig. 1. Time history of noise annoyance indicators during the year 2009.

Kolmogorov-Smirnov test at the significance level $\alpha = 0.05$. Probability values (p-values) for these tests are presented in Table 1.

Table 1. Probability values for the performed normality tests.

Normality test	p-values	
	L_{DEN}	L_{N}
Shapiro-Wilk	$2.5120 \cdot 10^{-9}$	$2.0221 \cdot 10^{-4}$
Jarque-Bera	$1.0000 \cdot 10^{-3}$	0.0072
Lilliefors	$1.0000 \cdot 10^{-3}$	0.0079
Kolmogorov-Smirnov	$8.7688 \cdot 10^{-294}$	$8.7688 \cdot 10^{-294}$

The obtained values are much less than the assumed significance level. This is an evidence of significant "distance" between the distribution of probability of the variable describing the 24-hour day-evening-night sound levels and the normal distribution.

4. Simulation experiments, results and discussion

Two simulation experiments have been conducted in order to specify the minimum size of the bootstrap algorithm input parameters, i.e. original sample size n, and the number of bootstrap replications B in order to determine the expected values of long-term noise indicators L_{DEN} and L_{N} with required accuracy.

4.1. Experiment #1

The first experiment served to determine the impact of original random sample size n on the estimation accuracy of long-term noise indicators is presented in Fig. 2.

For that reason, 1000 random samples with sizes $n = 2, 3, ..., 334$ were drawn from the examined population. The original random sample size n simulates the number of control days based on which the L_{DEN} and L_{N} indicators are estimated. In order to eliminate the impact of the number of bootstrap replications B on the estimation result of the expected value of noise indicators, the reconstruction of probability distributions was performed based on the same number of replications B for each sample with size n. The distributions were determined based on $B = 10\,000$ replications, thus receiving 1000 bootstrap probability distributions with $10\,000$ elements for each original sample size n. Each distribution was used to determine the bootstrap estimate of the expected value of noise indicators from the following equations:

$$L_{\text{DEN},m} = 10 \log \left(\frac{1}{B} \sum_{b=1}^{B} 10^{0.1 L_{\text{DEN},b}} \right),$$

$$L_{\text{N},m} = 10 \log \left(\frac{1}{B} \sum_{b=1}^{B} 10^{0.1 L_{\text{N},b}} \right). \tag{6}$$
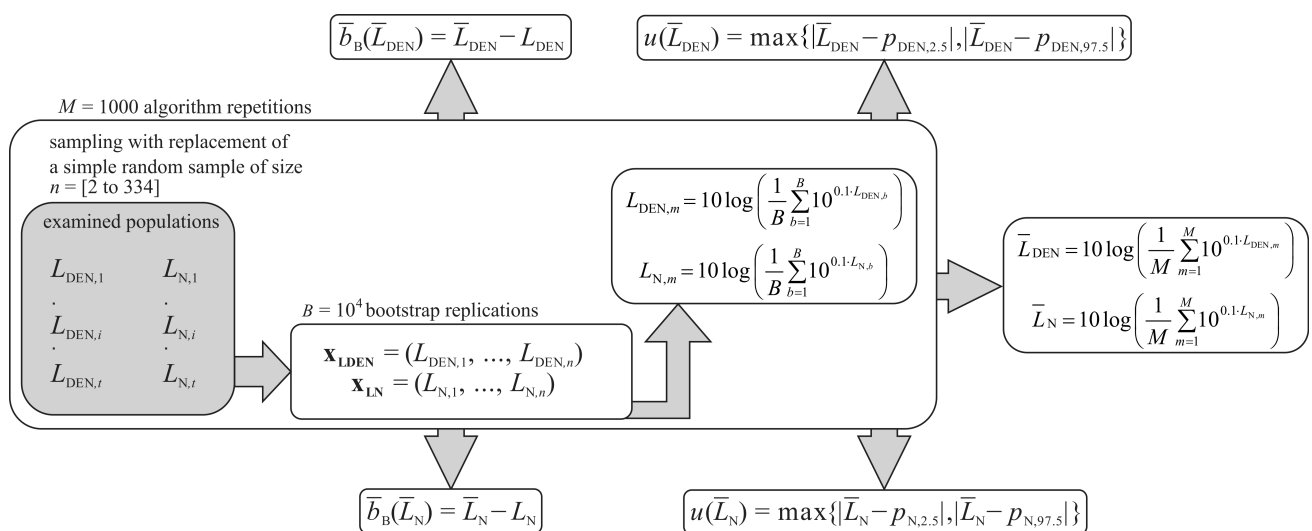

Fig. 2. Schematic diagram of the Experiment #1.

The result was 1000-element probability distributions of L_{DEN} and L_{N} indicators which were subjected to further statistical analysis.

First, the Kruskal-Wallis non-parametric test has been performed at the significance level $\alpha = 0.01$ in order to check if there are statistically significant differences in estimated long-term noise indicators for various original sample sizes. The test gave the probability values of $p = 7.57 \cdot 10^{-108}$ for L_{DEN} and $p = 4.93 \cdot 10^{-14}$ for L_{N}. These values are much less than the assumed level of significance which proves the existence of statistically significant differences in values of estimated indicators. The Tukey-Kramer multiple comparison test at the level of significance $\alpha = 0.05$ was conducted in order to find out between which groups there are differences. The probability values from this test are presented as a matrix in Fig. 3. White colour in the figure indicates the probability values greater than the assumed level of significance, and the black colour the values which are less than assumed level of significance. Such presen-

a)

b)

Fig. 3. The probability values (p-values) of the Tukey-Kramer test: a) for L_{DEN}, b) for L_{N}. The values higher than the assumed significance level $\alpha = 0.05$ are marked in white colour, while the lower values are marked in black colour.

tation facilitates the analysis of results and inference. The white points indicate absence of statistically significant differences between the analysed groups, and the black points indicate presence of such differences.

The results of the Tukey-Kramer test for L_{DEN} presented in Fig. 3a indicate three intervals of the original random sample size n based on which the estimated L_{DEN} expected values are not statistically different at the assumed level of significance. The first interval includes the original sample sizes from 2 to 51, the second from 19 to 98, and the third from 51 to 334. The determination of the third interval omitted the bootstrap distribution obtained from the sample size 65 whose expected value of L_{DEN} is statistically different only from 21 values obtained from the bootstrap distributions reconstructions of which were based on more numerous original samples. In case of the long-term average night-time sound level, there are two such intervals (Fig. 3b). The first includes the original sample sizes from 2 to 31, and the second from 25 to 334, omitting only five pairs of bootstrap distributions based on which the determined L_{N} are statistically different at the assumed level of significance.

The results of the Tukey-Kramer multiple comparison test confirm the claim that it is possible to estimate the long-term noise indicators L_{DEN} and L_{N} using the bootstrap algorithm based on a small sample. Figure 3 shows clearly that in the sample size from 2 to 50 elements there are values of samples size n for which the values of estimated indicators are not statistically different from each other regardless of the number of elements in the original sample. There are 26 and 35 such values for L_{DEN} and L_{N}, respectively.

The dispersion of obtained results was also analysed by determining the 95% confidence interval width using the percentiles of the bootstrap distribution for each probability distribution obtained using the original random sample of size n. This method was chosen because some analysed distributions were not Gaussian and some had a large skewness. The values of this parameter were from -2.99 to 1.39 for L_{DEN}, and from -1.34 to 0.35 for L_{N}. This method can be successfully used to determine to measurement uncertainty which was analysed in detail in previous work of the author (STĘPIEŃ, 2017).

The uncertainty of long-term noise indicators was determined as the function of size of original random sample n, which is shown in Fig. 4. The uncertainty was defined as

$$u\left(\overline{L}_{\text{DEN}}\right) = \max\left\{\left|\overline{L}_{\text{DEN}} - p_{\text{DEN},2.5}\right|, \left|\overline{L}_{\text{DEN}} - p_{\text{DEN},97.5}\right|\right\},$$
$$u\left(\overline{L}_{\text{N}}\right) = \max\left\{\left|\overline{L}_{\text{N}} - p_{\text{N},2.5}\right|, \left|\overline{L}_{\text{N}} - p_{\text{N},97.5}\right|\right\},$$

$$(7)$$

where $\overline{L}_{\text{DEN}}$ and $\overline{L}_{\text{N}}$ are the bootstrap estimates of the long-term noise indicators, $p_{\text{DEN},2.5}$ and $p_{\text{DEN},97.5}$ are the 2.5th and 97.5th empirical percentiles of the bootstrap distributions of L_{DEN}, while $p_{\text{N},2.5}$ and $p_{\text{N},97.5}$
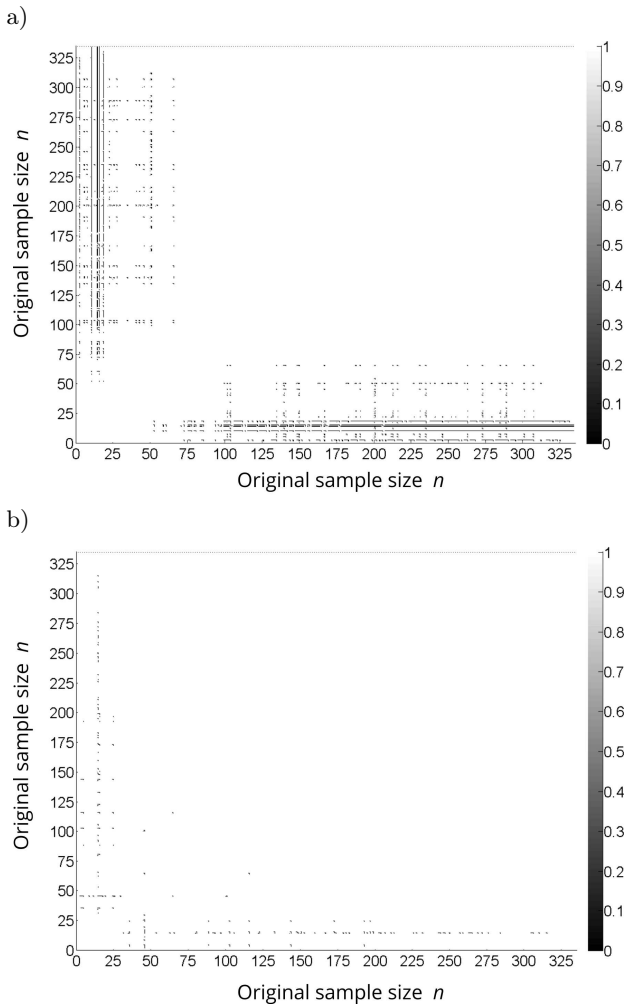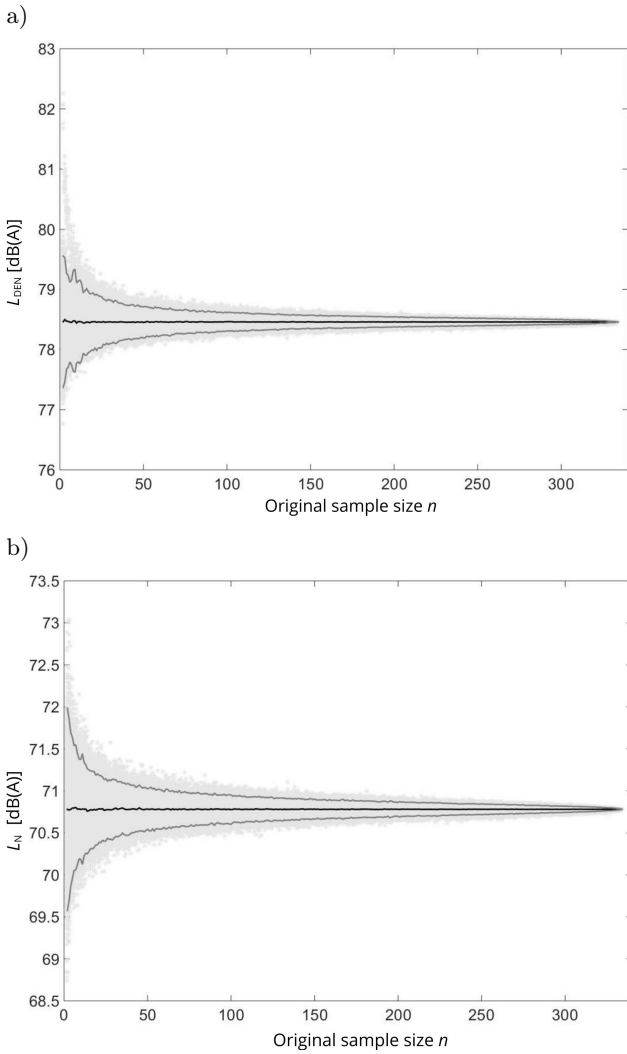
a)

b)

Fig. 4. The 1000-element bootstrap distributions of long-term noise indicators (light grey stars) obtained on the base of original samples of size n: a) for L_{DEN}, b) for L_{N}. The black line in each panel shows the bootstrap estimate of long-term noise indicator. The values $\overline{L}_{\text{DEN}} \pm u\left(\overline{L}_{\text{DEN}}\right)$ and $\overline{L}_{\text{N}} \pm u\left(\overline{L}_{\text{N}}\right)$ in the panels a) and b), respectively, are marked by the dark grey line.

are the 2.5th and 97.5th empirical percentiles of the bootstrap distributions of L_{N}. The results clearly show that the uncertainty decreases when the size of original random sample increases. From the results, it has been concluded that the uncertainty of long-term noise indicators below ± 1 dB(A) requires all-day noise measurement on three randomly selected days a year.

The next analysed parameter was bias of bootstrap estimator calculated according to the Eq. (5), where $\overline{\theta}_{\text{B}}$ is a bootstrap estimate of a long-term noise indicator, and $\widehat{\theta}$ is a value of long-term noise indicator from measurements. The parameters values range from -0.02 dB(A) to 0.04 dB(A) for L_{DEN}, and from -0.03 dB(A) to 0.02 dB(A) for L_{N}. Figure 5 presents the mean values of cumulative sums of bias $M_{\text{CSB},j}$ calculated according to the equation

$$M_{\text{CSB},j} = \frac{1}{j}\sum_{i=1}^{j}\widehat{b}_{\text{B},j}, \tag{8}$$

for $j = 1, 2, \ldots, N$, where N is the sample size ($N = 333$), and $\widehat{b}_{\text{B},j}$ is a bias of bootstrap estimator (5). The values of this statistics show convergence of bootstrap estimator towards the expected value of bias. The $M_{\text{CSB},j}$ value stabilizes around the value determined for $N = 333$ at the original sample size $n = 50$ for L_{DEN} and $n = 34$ for L_{N}. The graph of this statistics for the expected value of long-term noise indicators is the same as in Fig. 5, only it is shifted towards higher values by the value of this indicators. Based on the graph of statistics $M_{\text{CSB},j}$, it was concluded that the maximum size of the original sample n used to estimate the long-term noise indicators should not exceed $n = 50$, because further increase of the sample size does not improve estimation accuracy and only increases measurement costs.
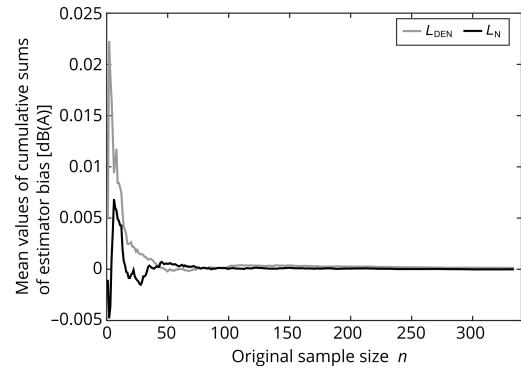

Fig. 5. The convergence processes of the bootstrap algorithm to the expected value of estimator bias of long-term noise indicators. The mean values of cumulative sums of estimator bias for L_{DEN} and L_{N} are indicated by the grey line and the black line, respectively.

4.2. Experiment #2

In the first experiment, it was proved that for the original sample size $n > 50$ the estimation accuracy is significantly not improved. For that reason, the experiment determined the impact of the number of bootstrap replications B on the estimation accuracy of long-term noise indicators L_{DEN} and L_{N} for sets of various sizes of the basic sample n in the 5 to 50 range with an increment of 5.

This experiment (Fig. 6) was similar to the first experiment (Fig. 2). One thousand original samples each were randomly drawn from the examined populations for each analysed size from 5 to 50 elements. Then, based on these original samples generated were B bootstrap samples from the interval from 100 to 10 000 with an increment of 100. Thus, we obtained 1000-element L_{DEN} and L_{N} probability distributions for each analysed number of bootstrap replications in each set which were then further statistically processed.
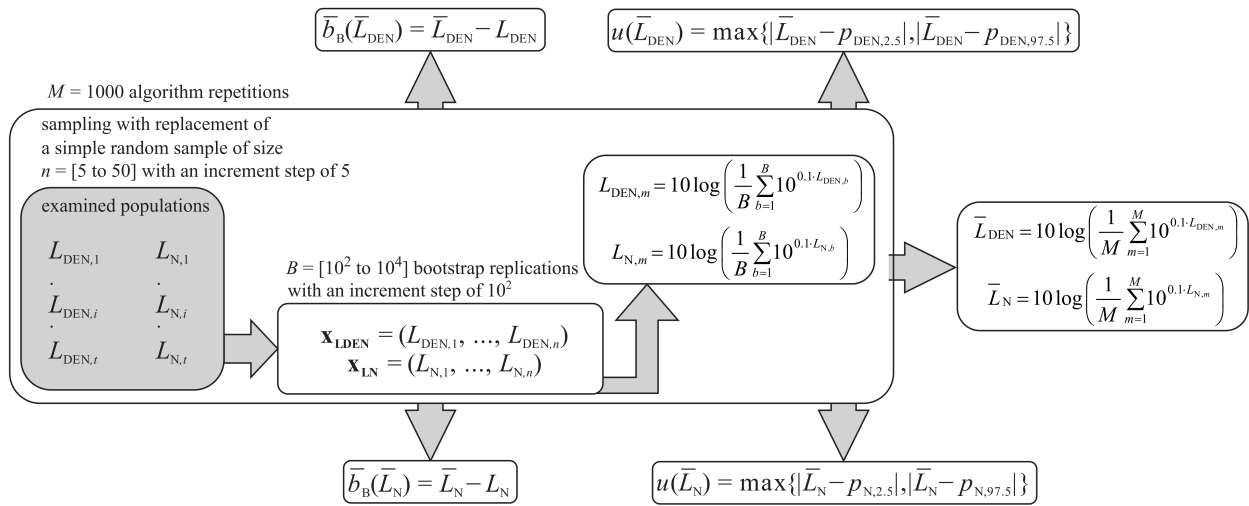
Fig. 6. Schematic diagram of the Experiment #2.

Similarly, firstly Kruskal-Wallis test was performed for each dataset at the level of significance $\alpha = 0.01$ in order to check if there were statistically significant differences in estimated long-term noise indicators depending on the number of bootstrap replications B. The probability values in all analysed sets were higher than the assumed level of significance α: for L_{DEN} they were in the 0.2275 to 0.8389 range, and for L_N in the 0.1215 to 0.9938 range. The results clearly show that in all analysed sets the estimates of expected value of long-term noise indicators are not statistically different regardless of the number of bootstrap replications which was used to determine them.

The convergence of the bootstrap algorithm towards the expected value of long-term noise indicators has been also analysed in the function of number of bootstrap replications based on the mean value of cumulative sums of bootstrap estimates $M_{CSE,j}$ described by the equation

$$M_{CSE,j} = \frac{1}{j}\sum_{i=1}^{j}\overline{\theta}_{B,j}, \qquad (9)$$

for $j = 1, 2, ..., N$, where N is the sample size ($N = 100$), and $\overline{\theta}_{B,j}$ is a bootstrap estimate of the expected value of long-term noise indicator. The graph of the statistics determined based on the Eq. (9), showing the convergence of the bootstrap algorithm for extreme cases, is presented in Fig. 7. The analysis of graphs shown in Fig. 7 clearly indicates that the estimates of expected values of long-term noise indicators obtained before the algorithm stabilization have a higher bias than the estimates obtained after the stabilization. The number of bootstrap replications at which the algorithm has stabilized, for both expected value and uncertainty, is presented in Table 2. The analysis of values included in Table 2 indicates that the number of bootstrap replications B at which the algorithm has been considered
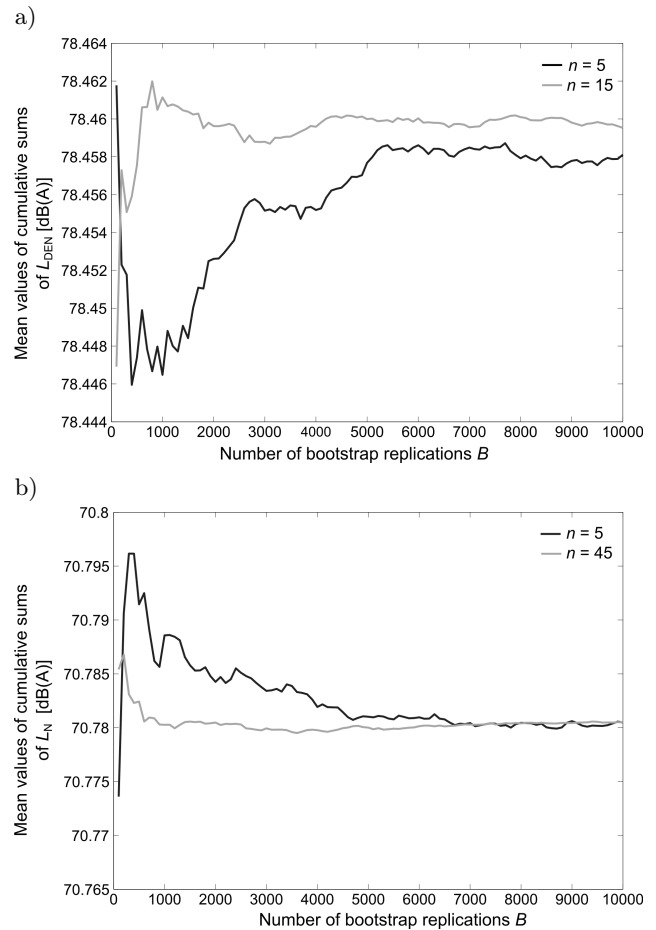

Fig. 7. The selected convergence processes of the bootstrap algorithm to the expected values of long-term noise indicators as a function of the number of bootstrap replications: a) for L_{DEN}, b) for L_N. The grey line shows the fastest process of convergence to the expected value, while the slowest one is marked by the black line.

stable was different depending on the size of the original random sample on which the estimation was based.

Table 2. Minimum number of bootstrap replications that guarantees the stability of bootstrap algorithm for different
sizes n of the original random samples for estimation expected value and uncertainty of L_{DEN} and L_{N}.

Indicator	Number of bootstrap replications B									
	Original sample size n									
	5	10	15	20	25	30	35	40	45	50
Expected value										
L_{DEN}	5000	3000	1000	5000	3000	1500	3500	2000	2000	2000
L_{N}	4500	1000	2000	3000	3500	2500	3500	2000	1000	4000
Uncertainty										
L_{DEN}	1500	2000	2000	3000	2000	1200	3000	2000	1500	2000
L_{N}	4000	2000	1500	1000	2000	3000	2000	3000	3000	2000

The results do not show any trend which could indi-
cate any relationship between the required number of
bootstrap replications B depending on the size of orig-
inal random sample in order to stabilize the algorithm.
These are random values which depend on the struc-
ture of the examined population. The number of boot-
strap replications B at which the algorithm was stable
is from 1000 to 5000 for L_{DEN} and from 1000 to 4500
for L_{N}.

   The next parameter analysed for each dataset was
uncertainty of L_{DEN} and L_{N} in the function of the

number of bootstrap replications. The uncertainty
was defined identically as in Experiment #1 on the
basis of expression (7). The uncertainty values of
long-term noise indicators are presented in Fig. 8. The
results do not show any trend which could indicate
any relationship between the uncertainty of long-term
noise indicators and the number of bootstrap replica-
tions. The presented graphs oscillate around some set
values, that is expected values of uncertainty of long-
term noise indicators which are included in Table 3.
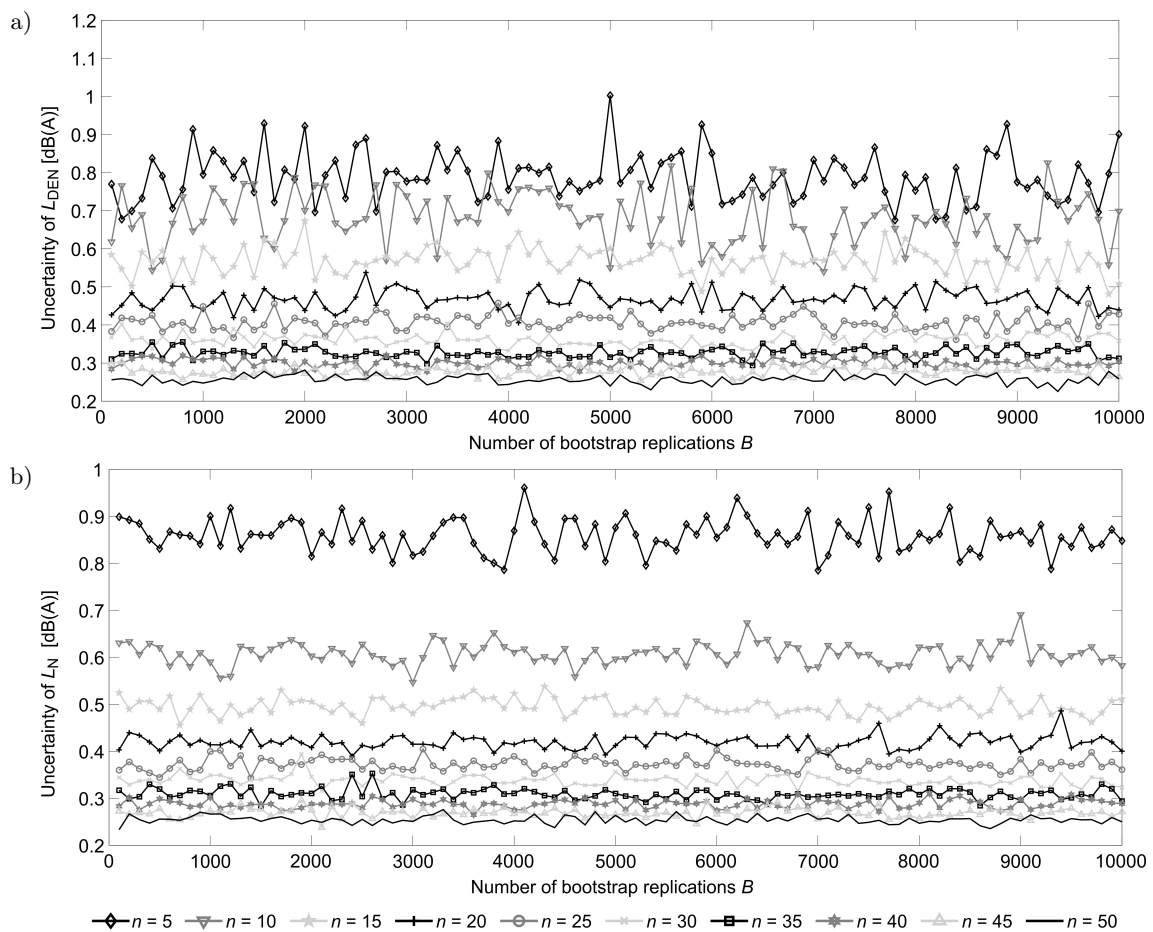The values are from 0.2565 to 0.7909 for L_{DEN}, and

Fig. 8. The uncertainty of long-term noise indicators for different sizes n of the original random samples as a function
of the number of bootstrap replications for: a) L_{DEN}, b) L_{N}.

Table 3. Expected value of uncertainty of L_{DEN} and L_{N} for different sizes n of the original random samples.

Indicator	Expected value of uncertainty [dB(A)]									
	Original sample size n									
	5	10	15	20	25	30	35	40	45	50
L_{DEN}	0.7881	0.6836	0.5683	0.4646	0.4059	0.3624	0.3269	0.3015	0.2772	0.2574
L_{N}	0.8639	0.6082	0.4977	0.4193	0.3717	0.3406	0.3112	0.2876	0.2693	0.2546

from 0.2549 to 0.8610 for L_{N}. The analysis of graphs in Fig. 8 and values in Table 3 indicates that the uncertainty is inversely proportional to the size of original random sample based on which the long-term noise indicators L_{DEN} and L_{N} were estimated.

The results of this experiment are in line with the results obtained in Experiment #1 at the constant number of bootstrap replications (B = 10 000). The similarity of results of two independent experiments proves a very good stability of the bootstrap algorithm, confirming that this approach can be used successfully in estimating the expected value and standard uncertainty of long-term noise indicators.

The convergence of the bootstrap algorithm towards the expected value of uncertainty long-term noise indicators has been analysed in the function of the number of bootstrap replications, which is presented in Table 2. Similarly to the algorithm convergence towards the expected value of long-term noise indicators, there is no trend which could indicate any

relationship between the required number of bootstrap replications B depending on the size of original random sample in order to stabilizes the algorithm. The number of bootstrap replications at which the algorithm has been stabilized ranges from 1200 to 5000 for L_{DEN} and from 1000 to 5000 for L_{N}. The graph of mean value of cumulative sums of uncertainty

$$M_{\text{CSU},j} = \frac{1}{j} \sum_{i=1}^{j} u_{\text{B},j}, \qquad (10)$$

for $j = 1, 2, ..., N$, where N is the sample size ($N = 100$), and $u_{\text{B},j}$ is an uncertainty of the expected value of long-term noise indicators determined based on the expression (7), showing the convergence of the bootstrap algorithm towards the expected value of uncertainty is presented in Fig. 9.

Based on the presented results of algorithm convergence, it was concluded that the minimum number of bootstrap replications for estimation of the expected
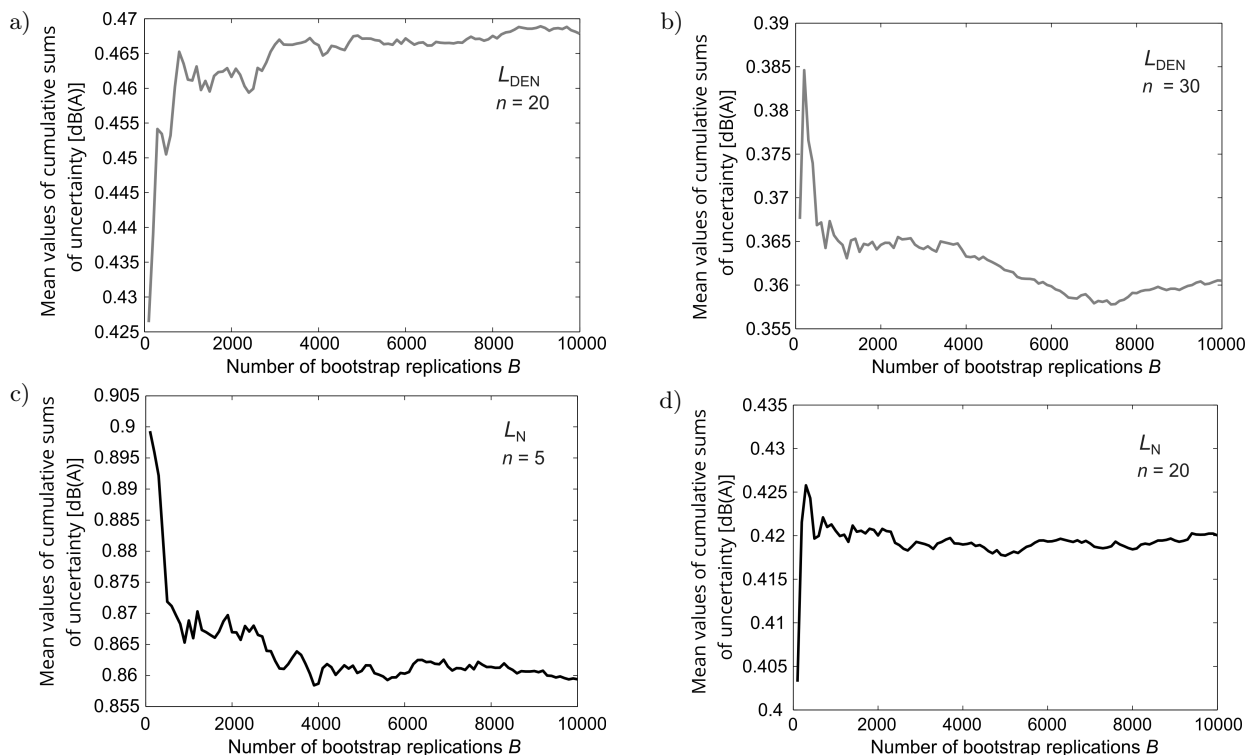

Fig. 9. The selected convergence processes of the bootstrap algorithm to the expected value of uncertainty of long-term noise indicators as a function of the number of bootstrap replications for: a) and b) L_{DEN}, c) and d) L_{N}. The b) and d) panels show the fastest process of convergence to the expected value of uncertainty, while the slowest one are show in panels a) and c).

value and uncertainty of long-term noise indicators should be $B = 5000$ in order to ensure an adequate algorithm convergence and consequently a satisfactory accuracy of estimated statistics.

5. Conclusions

The paper determines the minimum size of the bootstrap algorithm parameters (size of the original random sample and the number of bootstrap replications) necessary to estimate the expected value and the uncertainty of long-term indicators of road traffic noise with required accuracy. To this end, two independent simulation experiments were conducted. Experiment #1 served to determine the size of the original random sample, and Experiment #2 was used to determine impact of the number of bootstrap replications on the estimation accuracy of long-term noise indicators and their uncertainty.

The statistical analysis was carried out on the basis of Kruskal-Wallis test. Next, multiple comparison procedures were used for pairwise comparisons between the means using non-parametric Tukey-Kramer test at significance level $\alpha = 0.05$.

Two independent simulation experiments described above show that the estimation of long-term noise indicators with uncertainty below ±1 dB(A) requires all-day noise measurements during three randomly selected days during the year in a dense urban development.

The statistical analysis has shown that the maximum size of original random sample n used to estimate the values of long-term noise indicators should not exceed 50 elements.

The estimates of long-term noise indicators do not have a statistically significant difference regardless of the number of bootstrap replications B based on which they were determined.

The minimum number of bootstrap replications necessary to estimate the expected value and the uncertainty of long-term noise indicators should be $B = 5000$ in order to ensure an adequate algorithm convergence and consequently a satisfactory accuracy of estimated statistics.

The results of both experiments clearly indicate that the uncertainty decreases as the original random sample grows, thus proving a very good stability of the bootstrap algorithm and confirming that this approach can be successfully used to estimate not only long-term indicators of road traffic noise, but also other acoustic parameters.

The numerical experiment results presented in this paper refer only to one measuring cross-section located in a dense urban area. The minimum size of the bootstrap algorithm input parameters can be different in other measurement conditions. Therefore the minimum size of these parameters used for the cha-

racterization of each other place must be adapted according to the characteristics of the noise source and the local propagation conditions, because the proposed methodology may be applied to other situations where the temporal variation patterns of the environmental noise are known.

Acknowledgements

References

1. BATKO W., PAWLIK P. (2012), Uncertainty evaluation in modelling of acoustic phenomena with uncertain parameters using interval arithmetic, *Acta Physica Polonica A*, **121**(1–A): A-152–A-155, doi: 10.12693/APhysPolA.121.A-152.

2. BATKO W., PAWLIK P. (2013), New method of uncertainty evaluation of the sound insulation of partitions, *Acta Physica Polonica A*, **123**(6): 1012–1015, doi: 10.12693/APhysPolA.123.1012.

3. BATKO W., STĘPIEŃ B. (2010), Application of the bootstrap estimator for uncertainty analysis of the long-term noise indicators, *Acta Physica Polonica A*, **118**(1): 11–16, doi: 10.12693/APhysPolA.118.11.

4. BRAMBILLA G., CASTRO F.L., CERNIGLIA A., VERARDI P. (2007), Accuracy of temporal samplings of environmental noise to estimate the long-term L_{den} value, [in:] *36th International Congress and Exhibition on Noise Control Engineering INTER-NOISE 2007*, Turkish Acoustical Society, Istanbul, Turkey, pp. 4082–4089.

5. BRAMBILLA G., GALLO V., ZAMBON G. (2015), Prediction of accuracy of temporal sampling applied to non-urban road traffic noise, *Journal of Pollution Effects and Control*, **3**(3): 147, doi: 10.4172/2375-4397.1000147.

6. DON C.G., REES I.G. (1985), Road traffic sound level distributions, *Journal of Sound and Vibration*, **100**(1): 41–53, doi: 10.1016/0022-460X(85)90341-4.

7. EFRON B. (1979), Bootstrap methods: another look at the Jackknife, *Annals of Statistics*, **7**(1): 1–26.

8. EFRON B., TIBSHIRANI R.J. (1993), *An introduction to the bootstrap*, Chapman & Hall/CRC, New York.

9. European Parliament (2002), *Directive 2002/49/EC of the European Parliament and of the Council of 25 June 2002 relating to the assessment and management of environmental noise*.

10. FARRELLY F.A., BRAMBILLA G. (2003), Determination of uncertainty in environmental noise measurements by bootstrap method, *Journal of Sound and Vibration*, **268**(1): 167–175, doi: 10.1016/S0022-460X(03)00195-0.

11. Gaja E., Giménez A., Sancho S., Reig A. (2003), Sampling techniques for the estimation of the annual equivalent noise level under urban traffic conditions, *Applied Acoustics*, **64**(1): 43–53, doi: 10.1016/S0003-682X(02)00050-6.

12. Giménez A., González M. (2009), A stochastic model for the noise levels, *Journal of the Acoustical Society of America*, **125**(5): 3030–3037, doi: 10.1121/1.3109980.

13. Heiss A., Krapf K.G. (2007), Quantification of uncertainty by real time confidence limits in separation of sound immission levels, *Noise Control Engineering Journal*, **55**(1): 149–158, doi: 10.3397/1.2710947.

14. Huang N., Elhilali M. (2017), Auditory salience using natural soundscapes, *Journal of the Acoustical Society of America*, **141**(3): 2163–2176, doi: 10.1121/1.4979055.

15. ISO/IEC (2008), *ISO/IEC Guide 98-3:2008: Uncertainty of measurement. Part 3: Guide to the expression of uncertainty in measurement (GUM:1995)*.

16. Jagniatinskis A., Fiks B. (2014), Assessment of environmental noise from long-term window microphone measurements, *Applied Acoustics*, **76**: 377–385, doi: 10.1016/j.apacoust.2013.09.007.

17. Liguori C., Ruggiero A., Russo D., Sommella P. (2017a), Estimation of the minimum measurement time interval in acoustic noise, *Applied Acoustics*, **127**: 126–132, doi: 10.1016/j.apacoust.2017.05.032.

18. Liguori C., Ruggiero A., Russo D., Sommella P. (2017b), Innovative bootstrap approach for the estimation of minimum measurement time interval in road traffic noise evaluation, *Measurement*, **98**: 237–242, doi: 10.1016/j.measurement.2016.12.008.

19. Makarewicz R. (2011a), Estimation of L_{den} for road-, aircraft-, and train noise, *Acta Acustica united with Acustica*, **97**(3): 425–431, doi: 10.3813/AAA.918423.

20. Makarewicz R. (2011b), Uncertainty of approximation of the yearly average sound level of the road traffic noise, *Acta Acustica united with Acustica*, **97**(3): 416–424, doi: 10.3813/AAA.918422.

21. Mateus M., Carrilho J.A.D., da Silva M.C.G. (2015), Assessing the influence of the sampling strategy on the uncertainty of environmental noise measurements through the bootstrap method, *Applied Acoustics*, **89**: 159–165, doi: 10.1016/j.apacoust.2014.09.021.

22. Pilch A. (2018), Sources of measurement uncertainty in determination of the directional diffusion coefficient value, *Applied Acoustics*, **129**: 268–276, doi: 10.1016/j.apacoust.2017.07.016.

23. Przysucha B., Szeląg A., Pawlik P. (2020), Probability distributions of one-day noise indicators in the process of the type A uncertainty evaluation of long-term noise indicators, *Applied Acoustics*, **161**: 107158, doi: 10.1016/j.apacoust.2019.107158.

24. Romeu J., Jiménez S., Genescà M., Pàmies T., Capdevila R. (2006), Spatial sampling for night levels estimation in urban environments, *Journal of the Acoustical Society of America*, **120**(2): 791–800, doi: 10.1121/1.2215219.

25. Ruggiero A., Russo D., Sommella P. (2016), Determining environmental noise measurement uncertainty in the context of the Italian legislative framework, *Measurement*, **93**: 74–79, doi: 10.1016/j.measurement.2016.07.007.

26. Schomer P.D., DeVor R.E. (1981), Temporal sampling requirements for estimation of long-term average sound levels in the vicinity of airports, *Journal of the Acoustical Society of America*, **69**(3):, 713–719, doi: 10.1121/1.385569.

27. Schumacher T., Straub D., Higgins C. (2012), Toward a probabilistic acoustic emission source location algorithm: a Bayesian approach, *Journal of Sound and Vibration*, **331**(19): 4233–4245, doi: 10.1016/j.jsv.2012.04.028.

28. Stępień B. (2016), Confidence intervals for the long-term noise indicators using the kernel density estimator, *Archives of Acoustics*, **41**(3): 517–525, doi: 10.1515/aoa-2016-0050.

29. Stępień B. (2017), Comparison of selected methods of the confidence intervals for long-term noise indicators, *Acta Acustica united with Acustica*, **103**(2): 339–348, doi: 10.3813/AAA.919062.

30. Stępień B. (2018), A comparison of classical and Bayesian interval estimation for long-term indicators of road traffic noise, *Acta Acustica united with Acustica*, **104**(6): 1118–1129, doi: 10.3813/AAA.919276.

31. Vos J. (2017), Estimating parameter values in a model for rating shooting sounds from field survey data, *Journal of the Acoustical Society of America*, **141**(2): 864–877, doi: 10.1121/1.4976215.