ARCHIVES OF ACOUSTICS
34, 3, 295–303 (2009)

Non-Parametric Methods of Estimation of Type A Uncertainty of the Environmental Noise Hazard Indices

Wojciech BATKO, Bartłomiej STĘPIEŃ

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

(*received March 15, 2008; accepted August 19, 2009*)

A control of environmental noise hazards requires estimation of uncertainty of noise indices L_{DEN}, L_{N}. Assessment of the type A standard uncertainty in measurement results – expressed as the standard deviation of the mean, calculated the most often at the assumption of a normal distribution – is significant for the process. Such assumption – in relation to the noise measurement results – is of a relatively low likelihood. Thus, there is a need of looking for non-standard procedures of the standard deviation estimation of the mean of results, without any information of belonging to a certain class of distribution.

The aim of the hereby paper is an indication of the possibility of using non-parametric estimators of a density function in the calculation process of the type A standard uncertainty of environmental noise hazard indices. An attention was directed towards kernel estimators. The origin of their application, advantages and the method of constructing was described on the basis of a continuous monitoring of a traffic noise recorded on one of the main arteries of Kraków in 2004 and 2005. Usefulness of three forms of estimators, it means: kernel, unbiased and of maximum likelihood, was analysed.

Keywords: acoustic monitoring of environment, analysis of the results, type A standard uncertainty in measurements, kernel estimator

1. Introduction

Analysis of acoustic conditions of using advantageously the environment (according to the binding regulations in Poland [11], as well as in the European Union [4]) requires estimation of noise indices L_{DEN} and L_{N}. The problem of directive guidelines was described in many papers [9, 13]. Estimation of uncertainty of the measurements – when taking into account its significant components related to the applied measurement procedures, calibrating the measuring equipment, as well as input function conditions essential in the control process – is necessary in the calculation of indices. They are distinguished and quantified at the application of the proper analysis method.

An essential component of uncertainty is the *type A uncertainty* calculated by the statistical analysis of series of singular observations, at the assumption that the results are of a random character being submitted to the normal distribution. The value, which is an estimation of a standard deviation of an experimental mean, determined by classic variance estimators, is assigned to it.

The International Standard Organisation (ISO) issued the "Guide to the Expression of Uncertainty in Measurements" in 1995 [7]. The Polish version of the Guide was issued by the Central Office of Measures in 1999 [3]. This allowed a unified approach to several problems related to the subject. The Guide comprehensively explains the principles of determining uncertainty in measurements of repeatable values.

However, the application of its recommendations in respect of the estimation of the *type A standard uncertainty* of noise hazard indices in environment is rather **dubious**. As it results from papers: Wszołek, Kłaczyński [15] as well as Batko, Stępień [1], the assumption of a normal distribution of measurement results is difficult to be accepted. This is confirmed by the analysis of the measurement results of a traffic noise, which in significant majority required the rejection of the hypothesis concerning the possibility of using the normal distribution for the description. Extra-statistical information in relation to the occurrence of certain noise expositions in environment, especially in night hours (more than one maximum) also discredit this assumption.

An application of classic solutions of identification of probability distribution for the tested noise indices, allowing to estimate their type A uncertainty, is rendered difficult due to the fact that there are no likelihood indications on its form in the scientific literature. Therefore the authors propose to apply non-parametric statistical methods [5], allowing to determine the distribution of a random variable without any information of its belonging to the defined distribution class. Non-parametric estimation, especially kernel estimation [10], initiated by works of Rosenblatt [14] and Parzen [12] is an approach, which is based on a direct estimation of an unknown density function of random variable on the basis of data contained in a sample without postulating *a priori* the defined probability distribution function. Thus, it meets the requirements of the problem considered.

Properties of the proposed approach to the problem of estimation of the type A uncertainty of noise indices as well as the results of testing the usefulness of kernel estimators in the calculating process constitute the subject of the hereby paper. Considerations are illustrated by assessments of uncertainty in the continuous monitoring of traffic noise recorded on one of the main arteries of Kraków in 2004 and 2005.

2. Indices of environmental noise hazards

Directive 2002/49/EC of the European Parliament and of the Council relating to the assessment and management of environmental noise, was established on 25th June 2002.

One of the most important regulations of the Directive and following it the domestic legal acts – is an obligatory introduction of the realisation of acoustic maps.

In order to guarantee a uniform form and contents of such maps as well as the comparability of results the maps must be based on the common indices determined in the regulations:

- L_{DEN} – day-evening-night sound A level, dB, defined via sound level values determined in the characteristic day-times given by the equation [6]:

$$L_{\mathrm{DEN}} = 10 \log \left[\frac{1}{24} \left(12 \times 10^{0.1 L_{\mathrm{D}}} + 4 \times 10^{0.1(L_{\mathrm{E}}+5)} + 8 \times 10^{0.1(L_{\mathrm{N}}+10)} \right) \right], \qquad (1)$$

where L_{D} – day sound A level, determined within the whole day period, understood as the time interval from 6:00 a.m to 6.00 p.m., dB, L_{E} – evening sound A level, determined within the evening period, understood as the time interval from 6:00 p.m. to 10 p.m., dB, L_{N} – night sound A level, determined within the whole night, understood as the time interval from 10:00 p.m to 6.00 a.m., dB, L_{D}, L_{E}, L_{N} values needed for the determination of L_{DEN}, can be calculated from the following dependency [6]:

$$L_{\mathrm{D,E,N}} = 10 \log \left[\frac{1}{N} \sum_{i=1}^{N} 10^{0.1 \left(L_{\mathrm{Aeq},T} \right)_i} \right], \qquad (2)$$

where N – number of samples, $(L_{\mathrm{Aeq},T})_i$ – time equivalent of sound A for the i-th sample, dB.

- L_{N} – night sound A level, dB.

It should be mentioned, that the above given noise index L_{N}, being one of the parameters for calculating L_{DEN} level, is simultaneously the second index used for the preparation of acoustic maps and it can be calculated from Eq. (2).

3. Kernel density estimation for mean standard deviation estimation – as the type A uncertainty measure

When one is delivering the measurement result of a physical quantity he should also provide quantitative information on the result quality to enable the user of the result to estimate its likelihood. Without such information the measurement results cannot be compared neither with each other nor with the reference ones given in the specification or in the standard. Thus, the procedure of calculating and expressing uncertainty, convenient in application, easily understood and generally accepted, is really needed, [3].

Components of uncertainty of measurement results can be grouped into two categories according to the way of calculating their numerical values:

- type A method – calculated by statistical methods,
- type B method – calculated by other methods.

Category A components are characterised both by the variance estimate and by the standard deviation of the mean value. Trials of determination of the estimate of standard deviation of the expected value, which can be found from Eq. (3) [3] were undertaken in the hereby paper:

$$s\left(\overline{x}\right) = \sqrt{\frac{s^2\left(x_k\right)}{n}}, \tag{3}$$

where $s^2(x_k)$ – experimental variance of a random sample, n – sample size.

The experimental variance was – for the needs of this paper – determined by three estimation methods in order to check an influence of the selected method on the value of the estimated parameter.

The method of the kernel density estimation – belonging to the group of nonparametric methods – was used as the first one. The kernel density estimator was determined according to [10]:

$$\hat{f}(x) = \frac{1}{nh}\sum_{k=1}^{n} K\left(\frac{x - x_k}{h}\right), \tag{4}$$

where K – kernel function. A normal kernel function determined by the following equation was used in the paper:

$$K = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{x^2}{2}\right), \tag{5}$$

h – smoothing parameter, x_k – elements of a random sample.

On the basis of the kernel density estimation the experimental variance of the random sample was determined from the dependency [8]:

$$s^2(x_k) = \int\limits_{-\infty}^{+\infty}\left(x - \hat{E}(x)\right)^2 \hat{f}(x)\,\mathrm{d}x, \tag{6}$$

where $\hat{E}(x)$ – estimate of the expected value, determined from the equation:

$$\hat{E}(x) = \int\limits_{-\infty}^{+\infty} x\,\hat{f}(x)\,\mathrm{d}x, \tag{7}$$

$\hat{f}(x)$ – kernel density estimator determined on the basis of Eq. (4).

An unbiased estimator, properly describing a normal distribution variance, was used for determining the estimate of the experimental variance [8]:

$$s^2(x_k) = \frac{1}{n-1}\sum_{k=1}^{n}(x_k - \overline{x})^2, \tag{8}$$

where $\overline{x}$ – mean value in a sample.

In order to compare estimates determined by means of an unbiased estimator (8) with biased estimators using maximum likelihood estimator – belonging to the group of biased estimators – was decided. This estimator is in a form [8]:

$$s^2(x_k) = \frac{1}{n}\sum_{k=1}^{n}(x_k - \overline{x})^2. \tag{9}$$

4. The results of analysis

Noise hazard indices in the environment, which means day, evening and night levels of sound A, as well as day-evening-night level, were determined on the bases of the equivalent sound A levels recorded by the continuous monitoring acoustic station operating in Krasiński Avenue in Kraków in the year 2004 and 2005.

On the basis of the obtained results the kernel estimators of a probability density function (4) with using a normal kernel function (5) for the mentioned above indices, which are presented in Figs. 1 and 2 together with histograms for the year 2004 and 2005, were determined.

Estimates of the expected value $\hat{E}(x)$ on the basis of Eq. (7) were also calculated and compared with the mean value determined on the basis of the random sample of the twenty-four hours indices of environmental noise hazards. The results are presented in the table below.

As can be seen from Table 1, the estimate of the expected value is equal the mean value – made even to four places after the dot. This confirms the statement, that the arithmetic mean from a sample, regardless of the distribution of the variable under testing, is the most effective estimator of the mean value from the population in the class of unbiased estimators of this parameter [5].

Then the type A standard uncertainty was determined on the basis of Eq. (3), while the experimental variance estimate value of the random sample (appearing in the numerator (3)) was determined by means of three different estimators:

- estimator 1 – kernel estimator of the form (6),
- estimator 2 – unbiased estimator of the form (8),
- estimator 3 – maximum likelihood estimator of the form (9).

The results are listed in Table 2.

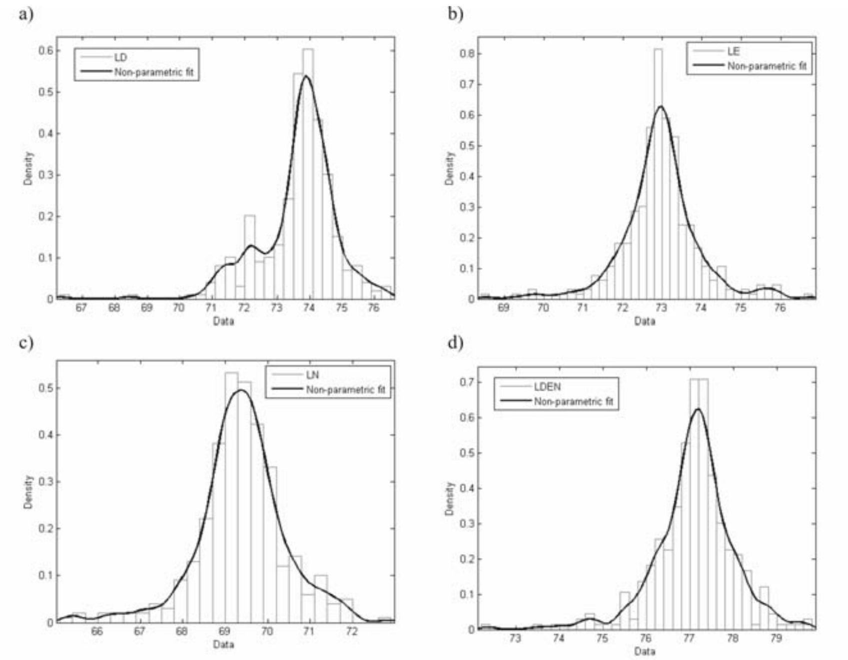

Fig. 1. Kernel density estimators together with histograms (2004) for: a) day level, b) evening level, c) night level, d) day-evening-night level.

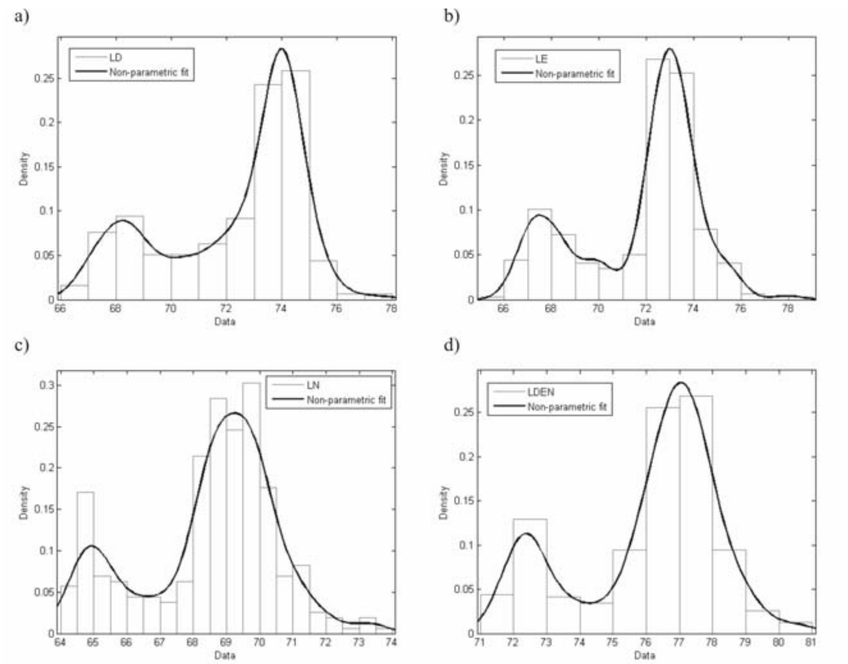

Fig. 2. Kernel density estimators together with histograms (2005) for: a) day level, b) evening level, c) night level, d) day-evening-night level.

Table 1. Comparison of the estimate of the expected value with the mean value of indices.

Year 2004		
Index	Expected value $\hat{E}(x)$ [dB]	Mean value $\bar{x}$ [dB]
L_{D}	73.6382	73.6382
L_{E}	72.9380	72.9380
L_{N}	69.3742	69.3742
L_{DEN}	77.0914	77.0914
Year 2005		
Index	Expected value $\hat{E}(x)$ [dB]	Mean value $\bar{x}$ [dB]
L_{D}	72.2828	72.2828
L_{E}	71.7527	71.7527
L_{N}	68.5019	68.5019
L_{DEN}	76.0702	76.0702

Table 2. Type A, standard uncertainty.

Year 2004			
Index	Estimator 1 [dB]	Estimator 2 [dB]	Estimator 3 [dB]
L_{D}	0.0662	0.0648	0.0647
L_{E}	0.0556	0.0544	0.0544
L_{N}	0.0589	0.0573	0.0573
L_{DEN}	0.0511	0.0500	0.0499
Year 2005			
Index	Estimator 1 [dB]	Estimator 2 [dB]	Estimator 3 [dB]
L_{D}	0.1472	0.1443	0.1440
L_{E}	0.1470	0.1445	0.1442
L_{N}	0.1183	0.1152	0.1150
L_{DEN}	0.1242	0.1215	0.1213

The results given in Table 2 allow to state, that the standard deviation of the expected value determined by means of the kernel estimator is of a higher value than the ones determined by means of other estimates. Type A standard uncertainty determined by means of the unbiased and maximum likelihood estimator was in the year 2004 of a very similar value. However – in the year 2005 – the uncertainty values determined by means of the unbiased estimator were higher than the ones determined by means of the maximum likelihood estimator (belonging to the group of biased estimators). This is the most probably caused by the bimodal distribution of noise hazard indices in the year 2005, contrary to the distribution of indices in 2004, what can be noticed when comparing Fig. 1 and Fig. 2.

5. Conclusions

As can be seen from the performed investigations the possibility of applying kernel estimators for the determination of the type A uncertainty – presented in the hereby paper – seems to be a promising tool for calculating errors in the measuring process of assessing noise hazards in the environment. This enriches the existing – in this scope – calculating algorithms, making them more likelihood due to the realization assumptions accompanying their application.

Conceptual approach to constructing kernel estimators is natural, interpretively clear, and its form is suitable for the mathematical analysis [2].

The statement, that the arithmetic mean of the sample is the most effective estimator of the population mean in the class of unbiased estimators, was confirmed by the performed calculations.

Values of the type A uncertainty, determined by means of the kernel estimator are higher than the values obtained by means of the unbiased and the maximum likelihood estimators. As the result the probability, that the measured real value will be contained within the assigned range, increases, which – in turn – decreases the risk of committing the 1st degree error.

References

[1] Batko W., Stępień B., *Analysis of traffic noise probability distribution* [in Polish:] *Analiza rozkładu prawdopodobieństwa hałasu drogowego*, 35th Winter School on Vibroacoustical Hazards Suppressions, pp. 5–16, Wisła, Poland, February 26 – March 02, 2007.

[2] Batko W., Stępień B., *Nonparametric estimate of probability density function for long-term indicators of environmental noise risk* [in Polish:] *Estymacja nieparametryczna funkcji gęstości prawdopodobieństwa długookresowych wskaźników zagrożeń hałasowych środowiska*, 36th Winter School on Vibroacoustical Hazards Suppressions, pp. 5–14, Wisła, Poland, February 25–29, 2008.

[3] Central Office of Measures, *Guide to the expression of uncertainty in measurement* [in Polish:] *Wyrażanie niepewności pomiarów. Przewodnik*, Warszawa 1999.

[4] *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, Official Journal of the European Communities 18.07.2002.

[5] Domański Cz., Pruska K., *Non-classical statistical methods* [in Polish:] *Nieklasyczne metody statystyczne*, Polish Economics Publishers, Warszawa 2000.

[6] Institute of Environmental Protection, *Guidelines for developing of acoustic maps* [in Polish:] *Wytyczne opracowania map akustycznych*, pp. 28–39, Vesign – Graf, Warszawa 2006.

[7] International Organization for Standardization, *Guide to the expression of uncertainty in measurement*, 1995.

[8] Koronacki P., Mielniczuk J., *Statistics for students of technical and natural discipline* [in Polish:] *Statystyka dla studentów kierunków technicznych i przyrodniczych*, WNT, Warszawa 2004.

[9] Kucharski R., *Complex noise indicator for noise map ping based on the EU working groups' and polish results of the annoyance investigations*, Archives of Acoustics, **32**, 2, 293–302 (2007).

[10] Kulczycki P., *Kernel estimators for analysis of systems* [in Polish:] *Estymatory jądrowe w analizie systemowej*, pp. 7–163, WNT, Warszawa 2005.

[11] Law of 27.04.2001, *Environmental protection law* [in Polish:] *Ustawa z dnia 27 kwietnia 2001 roku, Prawo ochrony środowiska,* Official Diary 2001/62/627 with amendments.

[12] Parzen E., *On estimation of a probability density function*, Annals of Math. Statist., **33**, 1065–1076 (1962).

[13] Popescu D.I., *Noise mapping in Romania within the framework of EU directive 2002/49/EC*, Archives of Acoustics, **32**, 2, 329–337 (2007).

[14] Rosenblatt M., *Remarks on some non-parametric estimates of a density function*, Annals of Math. Statist., **27**, 832–837 (1956).

[15] Wszołek T., Kłaczyński M., *Effect of traffic noise statistical distribution on $L_{Aeq,T}$ measurement uncertainty*, Archives of Acoustics, **31**, 4 (Supplement), 311–318 (2006).