www.czasopisma.pan.pl PAN www.journals.pan.pl

Metrol. Meas. Syst., Vol. 24 (2017), No. 1, pp. 69–78.

PAN POLSKA AKADEMIA NAUK

METROLOGY AND MEASUREMENT SYSTEMS

Index 330930, ISSN 0860-8229

www.metrology.pg.gda.pl

M&MS

ESTIMATION OF CONDITIONAL EXPECTED VALUE FOR EXPONENTIALLY AUTOCORRELATED DATA

Adam Kowalczyk, Anna Szlachta, Robert Hanus, Rafał Chorzępa

Rzeszow University of Technology, Department of Metrology and Diagnostic Systems, Powstańców Warszawy 12, 35-959 Rzeszów, Poland (kowadam@prz.edu.pl, ✉ *annasz@prz.edu.pl,* +48 17 743 2462, *rohan@prz.edu.pl, rchorz@prz.edu.pl)*

Abstract

Autocorrelation of signals and measurement data makes it difficult to estimate their statistical characteristics. However, the scope of usefulness of autocorrelation functions for statistical description of signal relation is narrowed down to linear processing models. The use of the conditional expected value opens new possibilities in the description of interdependence of stochastic signals for linear and non-linear models. It is described with relatively simple mathematical models with corresponding simple algorithms of their practical implementation.

The paper presents a practical model of exponential autocorrelation of measurement data and a theoretical analysis of its impact on the process of conditional averaging of data. Optimization conditions of the process were determined to decrease the variance of a characteristic of the conditional expected value. The obtained theoretical relations were compared with some examples of the experimental results.

Keywords: conditional averaging, conditional expected value, auto-correlated data, random signals.

1. Introduction

Intense worldwide technological development poses new challenging tasks for metrology. In measurement models, relations between stochastic variables (signals) should be taken into account. Stochastic relations make it harder to estimate the statistical characteristics of signals and measurement data. The majority of existing documents and measurement recommendations intended for use do not take into account the impact of stochastic relations between data [1–4].

Familiarity with the characteristics which describe stochastic relations is the primary question in solving many issues in science and technology. The most commonly used characteristics of stochastic relations are correlation functions. The need to take data correlation into account in evaluation of the measurement result uncertainty is indicated by many authors in specialised publications. An analysis of time series in the form of auto-correlated numerical sequences is presented in the study [5]. Determination of the impact of autocorrelation through an alternative measure, the so-called effective number of independent observations, was undertaken in a number of publications [6–8]. The derivation, analysis and examples of the use of formulae on the unbiased single measurement variance estimators and the arithmetic mean for correlated data, and a discussion on metrological usability of the proposed characteristics were presented in the works [9–11]. Simulation research into the impact of the instability of estimates of a normalised autocorrelation function on the uncertainty of the arithmetic mean value was carried out using the Monte Carlo method in the paper [12]. The application of Allan's variance in the analysis of correlated data is shown in the paper [13].

The theoretical correlation characteristics, which give effective results with linear relations, lose their benefits in the analysis of signals and systems with nonlinear characteristics of relations. In practice, in measurement of data with stochastic relations, correlation characteristics tend to create computational difficulties. There is a need to determine

Article history: received on May 11, 2016; accepted on Sep. 19, 2016; available online on Mar. 15, 2017; DOI: 10.1515/mms-2017-0005.

the behaviour of the autocorrelation function and the sign of correlation. Moreover, the literature does not provide accurate results in the estimation of the variance of a characteristic for any probability distributions describing signals.

The above-listed limitations pose a barrier to normative applications of correlative relations in the final measurement assessments with stochastic relations. Therefore, in studies and publications, other probabilistic characteristics describing the relationships of a stochastic nature are introduced and applied [14]. Methods of measurement and analysis of such characteristics are being developed intensively.

In performing a metrological identification of signals and systems, developing models and making research into stochastic signals, the authors use the theory and techniques of conditional signal averaging. In the paper there is examined a real approximate model of exponential correlation and its impact on the process of conditional data averaging. Continuity of the derivative of autocorrelation function was evidenced for the examined model of signal correlation. This makes possible the theoretical research into the correlation of conditionally averaged implementations of the signal. The conditions of averaging aimed at reducing the variance of the conditional average value were determined. The theoretical model of exponential correlation was compared with the results of experimental research.

2. Models of autocorrelation function

Linear and exponential models of the *autocorrelation function* (ACF) are most frequently applied in descriptions of data autocorrelation. Linear and exponential autocorrelation functions $R_x(\tau)$ of an argument $\tau = 0$ have a common feature – a lack of continuity of their derivatives, which in many situations makes analysis and calculations difficult when processing the signals. Such functions are called non-differentiable and they are characterized by the infinite value of the derivative variance.

A binary synchronous signal of parameters A, T and with an even distribution of moments of changes in the signal value (Fig. 1a) is characterized by a linear autocorrelation function (Fig. 1b). These kinds of signals occur in digital processing systems, *e.g.* after sampling and quantization; the analogue signals are transformed into bivalent signals.

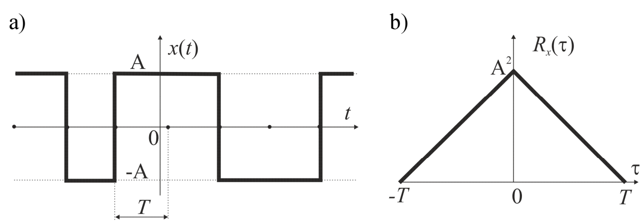

Fig. 1. A binary synchronous signal: a waveform (a); an autocorrelation function (b).

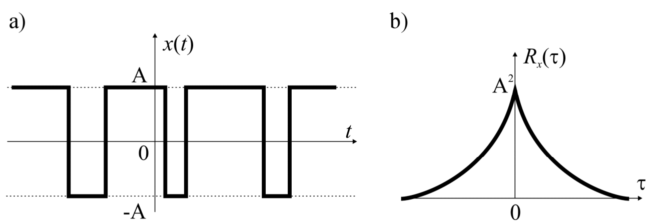

Fig. 2. A binary asynchronous signal: a waveform (a); an autocorrelation function (b).

Metrol. Meas. Syst., Vol. 24 (2017), No. 1, pp. 69–78.

Binary asynchronous signals with a random distribution of moments of changes in the signal value (Fig. 2a) are characterized by the ACF with an exponential shape, as presented in Fig. 2b. These kinds of signals occur in radioactive radiation trajectories.

An autocorrelation function with an exponential shape is also obtained by the output signal of a low-pass RC filter, when applying a white noise filter with a constant power spectral density $G_x(f) = G_0$ at the input:

$$R_x(\tau) = \frac{G_0}{RC} e^{-\frac{|\tau|}{RC}}. \tag{1}$$

The exponential shape of the ACF is a relatively frequent model when processing and describing analogue stochastic signals. In practice, an approximate model of exponential correlation is obtained for signals with a limited bandwidth with the features of white noise, passing physical inertial systems. Distributions of physical signals are usually normal or quasi normal, due to the central limit theorem and the inertia of typical processing systems.

3. Conditional averaging of auto-correlated data

The autocorrelation function is the main characteristic in the time domain describing the relation of stochastic signals. As a mixed second-order moment, it creates certain calculation difficulties, especially in assessment of the characteristic variance of auto-correlated data. The scope of usefulness of the ACF is narrowed down to linear models in probabilistic relations.

Restrictions for the measurement applications of correlation characteristics cause seeking other forms of description of stochastic relations for signals in the time domain. New possibilities in this scope for linear and non-linear models in metrological applications appear thanks to the use of functional and numerical conditional characteristics, in particular those of the conditional expected value and the conditional variance [14, 15]. The conditional expected value ensures the best estimate of interdependencies of stochastic signals in the mean square sense.

The conditional expected value in the time domain as the first-order central moment is described by relatively simple mathematical models with equally simple algorithms of their practical implementation corresponding to them [15]. In metrological applications of conditional averaging a right selection of the averaging condition enables to reduce the variance of estimates of experimental characteristics, which is one of the main objectives in measurement.

In basic applications of conditional averaging of Gaussian random signals, the characteristics of linear regression are used. For a single stationary signal with a distribution $N(0, \sigma_x)$ and a normalized ACF $\rho_x(\tau)$, the conditional expected value and the conditional variance are described by the following relations:

$$E(x_2|x_1) = \rho_x(\tau)x_1, \tag{2}$$

$$Var(x_2|x_1) = \sigma_x^2\left(1 - \rho_x^2(\tau)\right), \tag{3}$$

where: x_1 and x_2 are values of signal $x(t)$ at moments t_1 and t_2, respectively; $\tau = t_2 - t_1$.

In a simplified model of averaging non-correlated M fragments of $x(t)$, after exceeding the $x(t) = x_p$ level [14], the assessment of the relative standard uncertainty of the conditional value of arithmetic mean is:

$$\varepsilon = \frac{\sigma_x}{\sqrt{M} x_p} \frac{\sqrt{1 - \rho_x^2(\tau)}}{\rho_x(\tau)}. \tag{4}$$

In algorithms of conditional averaging using the maximum number of conditions $x(t)=x_p$ initiating the averaging, correlation of subsequently averaged fragments of the signal becomes problematic.

The following part of the paper presents the results of studies into correlation depending on a level x_p, which initiates conditional averaging. The studies were carried out assuming a normal distribution and an exponential correlation of a signal $x(t)$.

4. Assessment of auto-correlated data

Transition of white noise with a flat power spectral density in the B band, equal to $G(\omega)=\sigma^2/B$, through an RC inertial system is described by the relation of one-sided spectral density at the system output:

$$G_x(\omega)=G(\omega)\left(\frac{1}{\sqrt{1+(\omega RC)^2}}\right)^2=\frac{\sigma^2}{B\left[1+(\omega RC)^2\right]}. \tag{5}$$

The ACF at the inertial system output is described by:

$$R_x(\tau)=\frac{1}{2\pi}\int_0^{2\pi B}G_x(\omega)\cos\omega\tau d\omega=\frac{\sigma^2}{2\pi B}\int_0^{2\pi B}\frac{\cos\omega\tau}{1+(\omega RC)^2}d\omega . \tag{6}$$

After transformations, for $\tau=0$ we arrive at:

$$R_x(0)=\sigma_x^2=\frac{\sigma^2}{2\pi BRC}arctag2\pi BRC . \tag{7}$$

Example 1:

$B=25\cdot10^3\ Hz,$

$RC=100\cdot10^{-6}\ s,$

$2\pi BRC=2\pi\cdot25\cdot10^3\cdot10^2\cdot10^{-6}=5\pi,$

$R_x(0)=\sigma_x^2=\dfrac{\sigma^{\ 2}}{2\pi BRC}\dfrac{\pi}{2}=\dfrac{\sigma^2}{4BRC}=\dfrac{\sigma^2}{10}.$

The autocorrelation of subsequent instances exceeding a given level x_p by a signal $x(t)$ can be determined provided that the ACF $\rho_x(\tau)$ and statistical assessments of time intervals between appropriate instances exceeding the level x_p are known.

In order to assess the autocorrelation of subsequent signal fragments, after exceeding the given level x_p, the ratio of the maximum interval correlation τ_{km} and the average interval $\bar{\tau}_p$ is determined for the signal $x(t)$ exceeding the level x_p.

The average time of passing the level x_p by a signal $x(t)$ is described by [16]:

$$\bar{\tau}_p=\frac{1}{\overline{M}(x_p)}=\frac{2\pi}{\omega_{1x}}e^{\frac{x_p^2}{2\sigma}}, \tag{8}$$

where: $\overline{M}(x_p)$ τ an average number of signals $x(t)$ passing the level x_p with a derivative of one sign in a given time unit; ω_{1x} – an average frequency of the spectrum of a random signal $x(t)$ determined by:

$$\omega_{1x}^2=\sqrt{-\rho_x''(0)}=\frac{1}{2\pi\sigma_x^2}\int_0^{2\pi B}\omega^2G_x(\omega)d\omega, \tag{9}$$

Metrol. Meas. Syst., Vol. 24 (2017), No. 1, pp. 69–78.

where: $\rho_x''(0)$ – the second derivative of normalized ACF of a signal $x(t)$ for $\tau = 0$. Taking (5) and (9) into account, after necessary calculations, we obtain:

$$\omega_{1x} = \sqrt{-\rho_x''(0)} = \sqrt{\frac{1}{(RC)^2}\left(\frac{2\pi BRC}{arctg 2\pi BRC} - 1\right)}. \tag{10}$$

Based on the relations (8) and (10), the average time between signals $x(t)$ passing the level x_p with a derivative of one sign is:

$$\bar{\tau}_p = \frac{2\pi}{\sqrt{\frac{1}{(RC)^2}\left(\frac{2\pi BRC}{arctg 2\pi BRC} - 1\right)}} e^{\frac{x_p^2}{2\sigma_x^2}}. \tag{11}$$

For the maximum interval of autocorrelation τ_{k_m} of a signal $x(t)$ with an exponential ACF, equal to $\tau_{k_m} = 3\tau_k = 3RC$, the ratio $\bar{\tau}_p / \tau_{k_m}$ is described by:

$$\frac{\bar{\tau}_p}{\tau_{k_m}} = \frac{2\pi}{\sqrt{\frac{1}{(RC)^2}\left(\frac{2\pi BRC}{arctg 2\pi BRC} - 1\right)} \cdot 3RC} e^{\frac{x_p^2}{2\sigma_x^2}} = \frac{\frac{2}{3}\pi}{\sqrt{\frac{2\pi BRC}{arctg 2\pi BRC} - 1}} e^{\frac{x_p^2}{2\sigma_x^2}}. \tag{12}$$

Example 2:
For given relative values of the level $\nu_p = x_p / \sigma_x$ initiating conditional averaging and for the data included in Example 1, the ratio $\bar{\tau}_p / \tau_{k_m}$ is:

$$\frac{\bar{\tau}_p}{\tau_{k_m}} = \frac{\frac{2}{3}\pi}{\sqrt{\frac{5\pi}{\frac{\pi}{2}} - 1}} e^{\frac{x_p^2}{2\sigma_x^2}} = \frac{2}{9}\pi e^{\frac{\nu_p^2}{2}}. \tag{13}$$

Calculated and rounded values of the ratio $\bar{\tau}_p / \tau_{k_m}$ for several values ν_p are presented in Table 1.

Table 1. The values of ratio $\bar{\tau}_p / \tau_{k_m}$.

ν_p	0	1	$\sqrt{2}$	2
$\frac{\bar{\tau}_p}{\tau_{k_m}}$	0.70	1.15	1.90	5.16

For values $x_p \geq \sigma_x$, the averaged implementations of a signal $x(t)$ exceeding the level x_p with a derivative of one sign can be practically considered to be non-correlated.

Average implementations of $x(t)$, initiated by subsequent instances exceeding the level x_p with a derivative of any sign, can be described by the average time of a signal $x(t)$ being above the level $\nu_p = x_p / \sigma_x$:

$$\bar{\tau}_{p\pm} = \frac{\pi}{\omega_{1x}} e^{\frac{x_p^2}{2\sigma_x^2}} \left[1 - 2\Phi(\nu_p)\right], \tag{14}$$

where: $\Phi(\nu)=\frac{1}{\sqrt{2\pi}}\int_0^{\nu_p} e^{\frac{-z^2}{2}}\,dz$ – the Laplace's integral.

A chart of relation (14) is presented in Fig. 3.

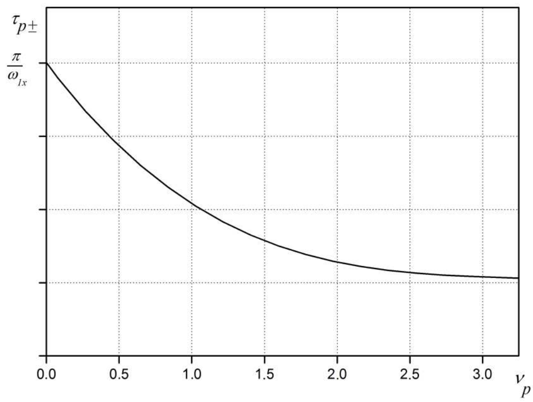

Fig. 3. An average time during which a signal $x(t)$ remains above the level ν_p.

Calculated and rounded values of the ratio $\bar{\tau}_{p\pm}/\tau_{k_m}$ for several values ν_p are presented in Table 2.

Table 2. The values of ratio $\bar{\tau}_{p\pm}/\tau_{k_m}$.

ν_p	0	1	$\sqrt{2}$	2
$\frac{\bar{\tau}_{p\pm}}{\tau_{k_m}}$	0.35	0.18	0.15	0.12

It can be concluded from the provided comparison that the time intervals between subsequent instances exceeding the level $x_p \geq 0$ by a signal $x(t)$ are on average significantly lower than the maximum interval of correlation τ_{k_m}. The ratio of the arithmetic mean of two time intervals between subsequent instances exceeding the level $\nu_p=\sqrt{2}$ with derivatives of various signs and the maximum correlation interval is:

$$\frac{\frac{\bar{\tau}_{p\pm}}{2}}{\tau_{k_m}}=\frac{\pi}{9}e^{1}\approx 0.95. \tag{15}$$

In the publications [17, 18] it was indicated that the optimal value of the level initiating conditional averaging of a signal with a normal distribution and an exponential ACF is included in the following interval:

$$\sqrt{2}\sigma_x \leq x_p \leq 2\sigma_x. \tag{16}$$

Using the obtained results to optimize the process of estimating the conditional expected value, we can perform the following sequence of calculations, basing on a random signal digitally registered in time:

1. Assume an averaging level x_p, $\sqrt{2}\sigma_x \leq x_p \leq 2\sigma_x$.

Metrol. Meas. Syst., Vol. 24 (2017), No. 1, pp. 69–78.

2. In the time interval $0-T_r$ $\left(T_r \geq \tau_{k_m}\right)$ average subsequent $M/2$ of implementations exceed the level x_p with a positive derivative.
3. Start and perform averaging with a delay by the first implementation in time T_r from the previous point also in the time interval $0-T_r$; subsequent $M/2$ of implementations exceed the level x_p with a negative derivative.
4. Perform synchronic averaging of values of partial characteristics from points 2 and 3.

Assessment of the relative standard uncertainty of determining the conditional average value for $\left(0 \leq \tau \leq T_r\right)$ can be calculated from the relation (4).

5. Experimental studies

The low-pass white noise $x(t)$ with a distribution N(0 V, 0.3 V) and a frequency band $B = 25$ kHz was applied on the inputs of first-order inertial systems with three different time constants T_c of: 10 µs, 30 µs and 100 µs. Figs. 4b–4d provide the obtained functions of the conditional average value (CAV) $\bar{x}(\tau)\big|x_p$, which are proportional to appropriate autocorrelation functions $\rho_x(\tau)$. For comparison, in Figure 4a the behaviour $\bar{x}(\tau)\big|x_p = f(\tau)$ of the original signal $x(t)$ was presented.

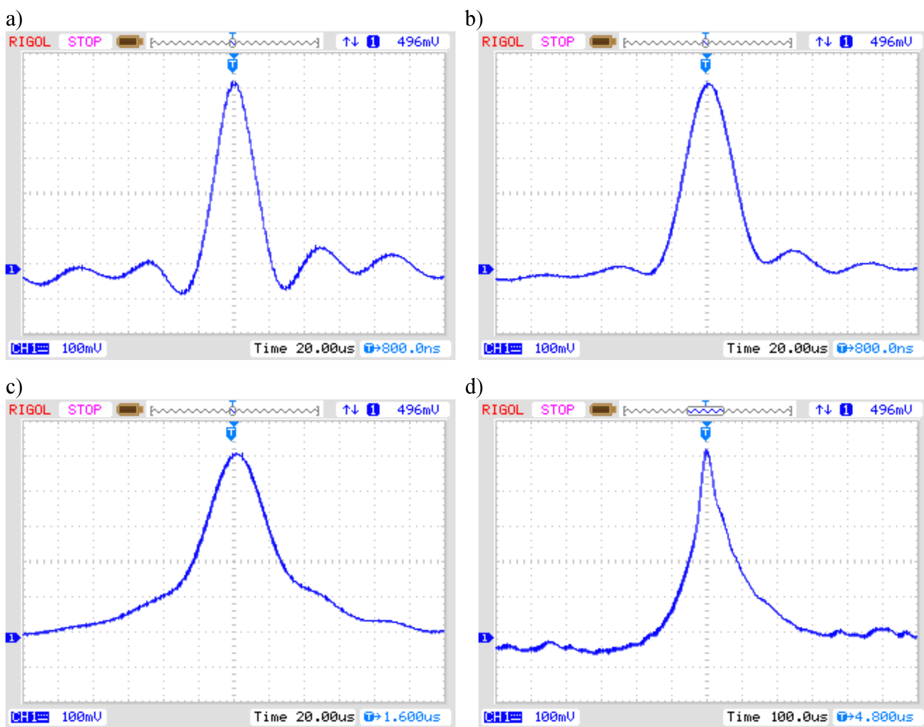

Fig. 4. The behaviour of functions of CAV: the low-pass white noise: N(0 V, 0,3 V), $B = 25$ kHz and after the noise passed the first-order inertial system with various time constants T_c (a): $T_c = 10$ µs (b); $T_c = 30$ µs (c); $T_c = 100$ µs (d).

The characteristics were designated using a RIGOL digital oscilloscope, with the level initiating averaging of $x_p = 0.5$ V and the number of averages $M = 256$.

The compliance of the experimentally obtained values $\rho_x(k\Delta t)$, determined based on the conditional function of the average value, was confirmed by calculating the normalized value of autocorrelation from the relation $\rho(k) = e^{-k}$. The obtained analytical results for T_c = 100 µs, Δt = 100 µs and $k = \Delta t / T_c$ are presented in Table 3.

Table 3. The values of function $\rho(k) = e^{-k}$ obtained from calculations.

k	0	1	2	3	4	5
$\rho(k)$	1	0.36	0.14	0.05	0.018	0.007

Figure 5 illustrates the behaviour of experimentally determined normalized ACF of the low-pass white noise: N(0 V, 0.3 V), B = 25 kHz when it passed the first-order inertial system with a time constant T_c = 100 µs. The values $\rho(k)$ presented in Table 3 were marked on the chart with crosses. A significant compliance of the results obtained experimentally and analytically is visible.

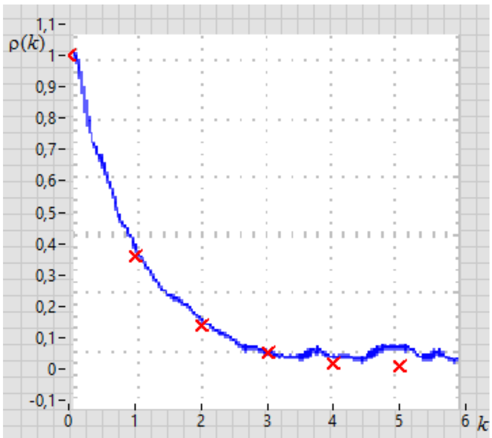

Fig. 5. The experimental behaviour of normalized autocorrelation function $\rho_x(k\Delta t)$ of the low-pass white noise N(0 V, 0.3 V), B = 25 kHz after it passed the first-order inertial system with a time constant T_c = 100 µs (continuous line) and the calculated autocorrelation values (points x).

6. Summary

1. In practice, the exponential autocorrelation model is obtained using white noise signals with a limited B band, passing physical inertial systems with a time constant T_c. In the implemented experiment for a product $BT_c \geq 2.5$, the function $\rho_x(k\Delta t)$ for $\tau = 0$ has a derivative, and for $\tau > 0$ the function shape is exponential. The presented realistic signal model is useful in practical applications.
2. Due to simple theoretical and practical models, when developing the auto-correlated measurement data with normal and quasi-normal distributions, conditional averaging algorithms can be used. It is especially beneficial in the case of strong data autocorrelation. The conditional average value is proportional to the ACF, therefore it can be used to assess interdependencies of measurement data, *e.g.* exponentially auto-correlated data.

Metrol. Meas. Syst., Vol. 24 (2017), No. 1, pp. 69–78.

3. In exponential autocorrelation models for a threshold condition value $x_p \geq \sigma_x$, initiating subsequent conditional averaging, averaged implementations of a signal $x(t)$ exceeding the level x_p with a derivative of one sign can be practically considered to be non-correlated.
4. In the exponential autocorrelation model, the subsequent averaged implementations of a signal $x(t)$ exceeding the level x_p with a derivative of any sign are significantly correlated.
5. The process of conditional averaging a signal $x(t)$ can be optimized: through selecting a value of the level x_p, and average conditional components determined with a lack of data correlation for instances exceeding the level $|x_p|$ with derivatives of any sign.
6. When estimating the conditional expected value for exponential oscillatory data autocorrelation models, relatively large values of maximum intervals of correlation τ_{km} should be taken into account when averaging with the use of non-correlated samples with a time $T_p \geq \tau_{km}$. Assessing values and signs of autocorrelations when averaging using correlated samples also needs to be considered.

References

[1] Guide to the expression of uncertainly in measurement. (1995). International Organisation for Standardisation.

[2] International vocabulary of basic and general terms in metrology (VIM). (2004). International Organization for Standardization. (Revision of the 1993 edition).

[3] Evaluation of measurement data. Supplement 1 to the "Guide to the expression of uncertainty in measurement" – Propagation of distributions using a Monte Carlo method (2008). JCGM 101:2008.

[4] Evaluation of measurement data. Supplement 2 to the "Guide to the expression of uncertainty in measurement" – Extension to any number of output quantities. (2011). JCGM 102:2011.

[5] Box, G.E.P., Jenkins, G.M., Reinsel, G.C. (1994). *Time series analysis: forecasting and control*. Prentice Hall, Englewood Cliffs.

[6] Bayley, G.V., Hammersley, G.M. (1946). The "effective" number of independent observations in an autocorrelated time-series. *J. Roy. Stat. Soc. Suppl.* 8, 184–197.

[7] Zhang, N.F. (2006). Calculation of the uncertainty of the mean of autocorrelated measurements. *Metrologia* 43, 276–281.

[8] Dorozhovets, M., Warsza, Z. (2007). Evaluation of the uncertainty type A of autocorrelated measurement observations. *Measurement Automation and Monitoring*, 53(2), 20–24.

[9] Witt, T.J. (2007). Using the autocorrelation function to characterize time series of voltage measurements. *Metrologia*, 44, 201–209.

[10] Zięba, A. (2010). Effective number of observations and unbiased estimators of variance for autocorrelated data – an overview. *Metrol. Meas. Syst.*, 17(1), 3–16.

[11] Zięba, A., Ramza P. (2011). Standard deviation of the mean of autocorrelated observations estimated with the use of the autocorrelation function estimated from the data. *Metrol. Meas. Syst.*, 18(4), 529–542.

[12] Dorozhovets, M. (2009). Influence of lack of a priori knowledge about autocorrelation functions of observations on estimation of their average value standard uncertainty. *Measurement Automation and Monitoring*, 55(12), 989–992.

[13] Zhang, N.F. (2008). Allan variance of time series models for measurement data. *Metrologia*, 45, 549–561.

[14] Kowalczyk, A., Szlachta, A., Hanus, R. (2012). Standard uncertainty determination of the mean for correlated data using conditional averaging. *Metrol. Meas. Syst.*, 19(4), 787–796.

[15] Kowalczyk, A. (2015). *Measurement applications of conditional signal averaging*. Oficyna Wydawnicza Politechniki Rzeszowskiej, Rzeszów.

[16] Bendat, J.S., Piersol, A.G. (2010). *Random data. Analysis and measurement procedures.* John Wiley & Sons, New York.

[17] Kowalczyk, A. (2008). Classical method for determination of linear system dynamic properties using signal conditional averaging. *Measurement Automation and Monitoring*, 54(12), 820–823.

[18] Szlachta, A., Kowalczyk, A., Wilk, G. (2009). Accuracy investigations of impulse response estimation obtained by conditional averaging. *Measurement Automation and Monitoring*, 55(12), 981–984.