

UNCERTAINTY OF THE CONVERSION FUNCTION CAUSED BY SYSTEMATIC EFFECTS IN MEASUREMENTS OF INPUT AND OUTPUT QUANTITIES

Mykhaylo Dorozhovets^{1,2)}

1) Rzeszow University of Technology, Faculty of Electrical and Computer Engineering, Department of Metrology and Diagnostic Systems, Wincentego Pola, 2A, 35-959 Rzeszow, Poland (✉ michdor@prz.edu.pl)

2) Lviv Polytechnic National University, Institute of Computer Technologies, Automation and Metrology, Department of Information Measuring Technology, Bandera str., 12, 79013 Lviv, Ukraine

Abstract

The paper presents an evaluation with the Type A and B methods for standard uncertainties of coefficients of a polynomial function of order k determined by n points obtained by measurement of input and output quantities. A method for deriving a posteriori distributions of function coefficients based on the transformation of estimator distributions without assuming any a priori distributions is presented. It was emphasized that since the correct values of the standard uncertainty of type A depend on the $\sqrt{n - k - 3}$ and not on the $\sqrt{n - k - 1}$, therefore, with a small number of measurement points, the use of the classical approach leads to a significant underestimation of uncertainty. The relationships for direct evaluation with the type B method of uncertainties caused by uncorrected systematic additive (offset error) and multiplicative (gain error) effects in the measurements of both input and output quantities are derived. These standard uncertainties are determined on the basis of the manufacturers' declared values of the maximum permissible errors of the measuring instruments used. A Monte Carlo experiment was carried out to verify the uncertainties of the coefficients and quadratic function, the results of which fully confirmed the results obtained analytically.

Keywords: uncertainty, systematic, effects, polynomial function, measurement system.

© 2023 Polish Academy of Sciences. All rights reserved

1. Introduction

In measurement practice polynomial function of order k :

$$Y = F(X) = \beta_0 + \beta_1 \cdot X + \dots + \beta_k X^k = \sum_{m=0}^k \beta_m \cdot X^m \quad (1)$$

(where $\beta = (\beta_0, \beta_1, \dots, \beta_k)^T$ is a vector of $k+1$ coefficients, X is input and Y is output quantities) is used in various tasks, such as to determine (i) the conversion functions of the measurement systems and sensors, (ii) the calibration function of measurement instruments and systems, (iii) the parameters of time series functions which represent certain technological process, *etc.* [1–8]. Each sensor can be used in a measurement system only when its conversion function as dependence between the output and input quantities is known. In many cases such conversion functions have the polynomial form (1). For example, traditionally, conversion functions of resistive copper, platinum

and nickel temperature sensors and others are described with polynomial functions of order from 1 to 3 [4]. Function (1) also can be used to model the dynamic properties of sensors, namely their complex frequency response [5]. One of the most important tasks in calibration of measuring instruments and systems is to determine their calibration curve [6–8]. The calibration curve, which represents the relation between quantity values provided by the measurement standards and the corresponding indications of a measuring system or instrument, is usually presented in the polynomial form (1) of a relatively small order. Polynomials of type (1) have also been often used to describe time processes, especially relatively with smooth process parameters [9–11].

In the classical approach, such a function is called regression. Various issues relating to polynomial regression have been deeply researched and presented in extensive literature, only some of them listed in [11–13]. To determine regression coefficients, the n ($n > k$) measurement pairs $(x_i; y_i, i = 1, 2, \dots, n)$ of values x_i (vector $\mathbf{x} = (x_1, x_2, \dots, x_n)^T$) of the input quantity X that are matched by values y_i (vector $\mathbf{y} = (y_1, y_2, \dots, y_n)^T$) of the output quantity Y are used. In a simple case, it is assumed that values y_i of the output quantity Y are distorted by normally distributed uncorrelated random noise ε_i with zero expected value ($\mu = 0$) and standard deviation σ . The estimators $\mathbf{b} = (b_0, b_1, \dots, b_k)^T$ of regression coefficients $\boldsymbol{\beta} = (\beta_0, \beta_1, \dots, \beta_k)^T$ are usually determined with the *maximum likelihood estimation* (MLE) or the *least squares method* (LSM) [11–13]. When using the LSM, the search values $\mathbf{b} = (b_0, b_1, \dots, b_k)^T$ are the solution of the matrix equation:

$$\mathbf{b} = (\boldsymbol{\Phi}^T \boldsymbol{\Phi})^{-1} \boldsymbol{\Phi}^T \cdot \mathbf{y} = \mathbf{M}^{-1} \cdot \mathbf{Y}, \quad (2)$$

where

$$\begin{aligned} \Phi_{i,m} &= x_i^m, \quad i = 1, \dots, n; \quad m = 0, \dots, k, \\ \mathbf{M} &= (\boldsymbol{\Phi}^T \boldsymbol{\Phi}), \quad \mathbf{M}^{-1} = \mathbf{D} = (\boldsymbol{\Phi}^T \boldsymbol{\Phi})^{-1}, \quad \mathbf{Y} = \boldsymbol{\Phi}^T \cdot \mathbf{y}. \end{aligned} \quad (3)$$

In order to show a more transparent relationship, we will introduce a simple normalization: matrices \mathbf{M} and \mathbf{Y} in (2) and (3) are divided by the number n of observations:

$$\mathbf{M}n = \frac{\mathbf{M}}{n}, \quad \mathbf{Y}n = \frac{\mathbf{Y}}{n}, \quad \mathbf{D}n = n \cdot \mathbf{D}, \quad \mathbf{b} = \mathbf{D}n \cdot \mathbf{Y}n. \quad (4)$$

Using this normalization, solution (2) will not change. On the other hand, the variances and standard deviations of the coefficients and functions will receive a normalizing factor $1/\sqrt{n}$, which is analogous to the factor in the evaluation uncertainty of measurement while processing multiple observations *i.e.*, where the standard deviation of the mean value is σ/\sqrt{n} . After estimation of the coefficients $\mathbf{b} = (b_0, b_1, \dots, b_k)^T$, the estimated dependence $y(x)$ between input and output quantities can be presented by function:

$$y(x) = b_0 + b_1 \cdot x + \dots + b_k x^k = \sum_{m=0}^k b_m \cdot x^m. \quad (5)$$

The variance $\sigma^2(b_m)$ and standard deviation $\sigma(b_m)$ of the estimated coefficient and standard deviation of function $\sigma[y(x)]$ are given by formulas:

$$\begin{aligned} \sigma^2(b_m) &= \frac{\sigma^2}{n} \cdot \mathbf{D}n_{m,m}; \quad \sigma(b_m) = \frac{\sigma}{\sqrt{n}} \cdot \sqrt{\mathbf{D}n_{m,m}}, \\ \sigma[y(x)] &= \frac{\sigma}{\sqrt{n}} \sqrt{\sum_{m=0}^k \sum_{l=0}^k \mathbf{D}n_{m,l} x^{m+l}}, \end{aligned} \quad (6)$$

where $\mathbf{D}n_{m,m}$ is the corresponding diagonal element of inverse matrices $\mathbf{D}n$. As variation σ^2 of population is usually unknown in praxis, so-called unbiased estimation is used [14]:

$$S^2 = \frac{1}{n-k-1} \sum_{i=1}^n \left(\sum_{m=0}^k b_m \cdot x_i^m - y_i \right)^2. \quad (7)$$

Therefore, the approximated standard uncertainties of coefficients and function traditionally are presented as [14]:

$$u_A(\beta_m) \approx \frac{S}{\sqrt{n}} \cdot \sqrt{\mathbf{D}n_{m,m}}, \quad u_A[Y(x)] \approx \frac{S}{\sqrt{n}} \cdot \sqrt{\sum_{m=0}^k \sum_{l=0}^k \mathbf{D}n_{m,l} x^{m+l}}. \quad (8)$$

The model used in the above analysis of the influence of uncorrelated random noise only on the output quantity, is very simplified. In practice, noise may be correlated or influenced by both output and input quantities. It is also possible to deviate the probability density function (PDF) of random noise from Gaussian. In addition to these, a very important factor, which is not included in the classical regression model, are systematic effects in the measurement results of both quantities. Therefore, extensive research is being carried out taking into account the influence of random effects on input and output variables, the influence of correlated random effects and the influence of systematic effects on the measurements of output quantities in the uncertainty assessment. Namely, in [15] and [16] there were presented results of the estimation of linear regression confidence bands in the case of correlated noise in the output quantity. In [17] and [18] results of the evaluation of regression straight line uncertainty due to correlated random effects in both quantities are presented. In this article, these problems are not being discussed.

On the other hand, it should be noted that there are limited research results for the impact of uncorrected systematic effects on the measurement results of both quantities. Measurements are realized with corresponding instruments, whose readings are never perfect, even after appropriate correction. Each measuring instrument is characterized by the values of the maximum permissible errors (MPE), which takes into account possible systematic effects. These systematic effects have many components, but in practice, the main components are additive effects (independent of the measurand), traditionally called offset errors Δ_0 , and multiplicative (proportional to the measurand), traditionally called gain errors δ_g . Not including the instrumental components of uncertainty caused by systematic effects in measurement results in some cases can lead to unjustifiably optimistic values of uncertainty of coefficients and function.

In [19], the influence of additive systematic component on the output variable is taken into account. In [20], the uncertainty of regression line is analyzed taking into account the Type A and B uncertainties of the dependent (output) variable. It was noted in [9] that in some cases it would be necessary to consider also the effects of the measurement error in the regression variables on the model and also the correlation between the regression variable and the measurement error. An example was given of the effect of additive influence on the value of a function in the linear regression model. However, a method for evaluating the uncertainty caused by such an influence is not provided. Multiplicative influences are generally not paid much attention.

It should be noted that since matrices (3) and (4) depend only on the input vector $\mathbf{x} = (x_1, x_2, \dots, x_n)^T$, therefore systematic effects on the output quantity do not change these matrices. Therefore, such systematic effects can be introduced into the model of output quantity (in matrix \mathbf{Y} in (3)), and then an appropriate uncertainty assessment can be performed. On the other hand, additive and multiplicative systematic effects in the results of the measurement of the

input quantity (vector $\mathbf{x} = (x_1, x_2, \dots, x_n)^T$) change the structure of matrices (3) and (4). Since the values of these systematic effects are not known, taking them into account in these matrices directly causes significant complications related to solving equation (2).

In addition to the problems related to random and systematic effects in measurements, there is a very important methodological problem related to the incompatibility of the classical way of evaluating uncertainty with the definition of uncertainty in the Guide [14]. Formally, the method of evaluating of Type A uncertainty presented in the Guide [14], strictly speaking, directly relates to statistical properties of the result determined after processing registered observations, but not to the properties of the measurand. This is due to the fact that standard uncertainty (8) applies to coefficient estimators \mathbf{b} (which are the results) and not to the coefficients $\boldsymbol{\beta}$ (which are the measurands in (1)). It should be clearly noted that according to the definition [14], uncertainty concerns the measurands, which are the coefficients $\boldsymbol{\beta}$, and not the measurement results, which are the estimators \mathbf{b} . It is quite reasonably stated in [21] that the classical (so called “frequentist”) statistical theory approach to the estimation of Type A uncertainty [14] is inconsistent with the definition of the “uncertainty” of the measurement.

Hence, according to the definition, in order to correctly calculate the uncertainty, one should use an appropriate PDF of the measurand referring to the measurement result and the parameter characterizing the dispersion of its possible values. One possible way to obtain such a PDF of the measurand is to use the Bayesian approach. Various aspects and problems related to the choice of the Bayesian approach to uncertainty assessment with the Type A method are presented in numerous literature [21–24]. Bayesian uncertainty analysis provides an a posteriori PDF of the measurand under the assumption of an appropriate a priori PDF for it. A priori PDFs represent the state of knowledge about population parameters before the measurements are performed. It is obvious that the use of any other a priori PDF entails a change in the a posteriori PDF. In general, choosing the proper a priori PDF is the fundamental problem of this method that gives rise to many discussions on this subject. Usually, in the absence of information on the a priori PDF, Jeffrey’s rule is used [25]. Namely, for a population described by an expected value μ and a variance σ^2 , these a priori PDFs are [22]:

$$p_a(\mu) \propto 1, \quad -\infty \leq \mu \leq \infty, \quad p_a(\sigma^2) \propto \frac{1}{\sigma^2}, \quad 0 \leq \sigma^2 \leq \infty. \quad (9)$$

For Gaussian uncorrelated random noise in the measurements of the output quantity, using the Bayesian approach with a priori PDFs such as (9), the a posteriori PDF $p_\beta(\beta_m | b_m, s)$ of regression coefficient β_m can be presented in the form of well-known *t*-Student’s distribution:

$$p_\beta(\beta_m | b_m, s) = \frac{1}{\sqrt{\pi} \sqrt{\mathbf{D}n_{m,m}} \cdot s} \cdot \frac{\Gamma\left(\frac{n-k}{2}\right)}{\Gamma\left(\frac{n-k-1}{2}\right)} \cdot \frac{1}{\left[1 + \left(\frac{\beta_m - b_m}{s}\right)^2 \cdot \frac{1}{\mathbf{D}n_{m,m}}\right]^{\frac{n-k}{2}}}, \quad (10)$$

$$m = 0, 1, \dots, k,$$

where

$$s^2 = \frac{1}{n} \sum_{i=1}^n \left(\sum_{m=0}^k b_m \cdot x_i^m - y_i \right)^2 = \frac{n-k-1}{n} S^2. \quad (11)$$

is the so-called biased estimate of variation.

Based on (10), the standard uncertainty of the coefficient β_m as standard deviation of Student's distribution for a number degrees of freedom $d = n - k - 1$ is:

$$u_A(\beta_m | s) = \frac{s}{\sqrt{n - k - 3}} \cdot \sqrt{\mathbf{D}n_{m,m}} = \frac{s}{\sqrt{d - 2}} \cdot \sqrt{\mathbf{D}n_{m,m}}, \quad d \geq 3. \quad (12)$$

It means that when standard uncertainty is determined with classical approaches, then the correct value of standard uncertainty must be changed by factor [21]:

$$k_{u_A}(d) = \sqrt{\frac{n - k - 1}{n - k - 3}} = \sqrt{\frac{d}{d - 2}}, \quad d \geq 3. \quad (13)$$

In the next part of the article, it will be shown that in order to obtain a posteriori PDF $p_\beta(\beta_m | b_m, s)$ of regression coefficients, in general, there is no need to use any a priori PDF. All you require is a joint PDF $p_{b,s}(b_m, s | \beta_m, \sigma)$ of the estimated coefficients b_m and estimated standard deviation s .

The aims of this paper are: (i) using the joint PDF of the estimators, deriving the a posteriori PDF of the coefficients of the function, which is then used to correctly evaluate Type A standard uncertainty; (ii) direct including the systematic (additive and multiplicative) components in measurement results of both input and output quantities in the uncertainty relationship; (iii) performing the Monte Carlo test for verification of the proposed approach to include systematic components in evaluating the Type B component of standard uncertainty as well as combined and expanded uncertainties.

2. Type A standard uncertainties

2.1. Derivation of the PDF of function coefficients

Similarly to the processing of multiple normally distributed uncorrelated observations [26], the PDF $p_\beta(\beta_m | b_m, s)$ of measurand β_m , can be derived from the joint PDF $p_{b,s}(b, s | \beta, \sigma)$ of the estimators $\mathbf{b} = (b_0, b_1, \dots, b_k)^T$ and s when coefficients $\boldsymbol{\beta} = (\beta_0, \beta_1, \dots, \beta_k)^T$ and standard deviation σ are known. As was mentioned above, in solving the estimation problem the coefficients β_m and standard deviation σ are constants and their estimated values b_m and s are random variables, which are described by the joint PDF $p_{b,s}(b_m, s | \beta_m, \sigma)$. It is well known [13] that when the random noises ε_i have normal distribution, then the PDF $p_b(b_m | \beta_m, \sigma)$ of estimate b_m of coefficient β_m ($m = 0, 1, \dots, k$) is normal too, with the expected value β_m and variance $\sigma^2 \cdot [\mathbf{M}^{-1}]_{m,m} = \frac{\sigma^2}{n} \mathbf{D}n_{m,m}$ (6):

$$p_b(b_m | \beta_m; \sigma) = \frac{\sqrt{n}}{\sigma \sqrt{2\pi} \sqrt{\mathbf{D}n_{m,m}}} e^{-\frac{n \cdot (b_m - \beta_m)^2}{2\sigma^2 \cdot \mathbf{D}n_{m,m}}}. \quad (14)$$

The covariance between estimates β_m and β_l is $\frac{\sigma^2}{n} \mathbf{D}n_{m,l}$, therefore the joint PDF $p_b(b | \beta, \sigma)$ of all estimates b_m has a $k+1$ -dimension normal distribution [13]. Besides, the estimate s^2 of variance σ^2 is distributed independently of b_m , and this distribution refers to the Chi² distributions with $n - k$ degrees of freedom [13]. As the estimate of standard deviation is $s = \sqrt{s^2}$, the PDF

$p_s(s|\sigma)$ of estimate s can be presented in the form of Chi distribution [27]:

$$p_s(s|\sigma) = \frac{1}{\sigma} \cdot \frac{2 \cdot \left(\frac{n}{2}\right)^{\left(\frac{n-k-1}{2}\right)}}{\Gamma\left(\frac{n-k-1}{2}\right)} \left(\frac{s}{\sigma}\right)^{n-k-2} e^{-\frac{n \cdot s^2}{2\sigma^2}}. \quad (15)$$

Due to independence of estimates b_m and s , their joint PDF $p_{b,s}(b_m, s|\beta_m, \sigma)$ is a product of the PDF $p_b(b_m|\beta_m, \sigma)$ (14) and PDF $p_s(s|\sigma)$ (15):

$$p_{b,s}(b_m, s|\beta_m, \sigma) = \frac{\sqrt{n}}{\sigma^2 \sqrt{2\pi} \sqrt{\mathbf{D}n_{m,m}}} \frac{2 \cdot \left(\frac{n}{2}\right)^{\left(\frac{n-k-1}{2}\right)}}{\Gamma\left(\frac{n-k-1}{2}\right)} \cdot \left(\frac{s}{\sigma}\right)^{n-k-2} \cdot e^{-\frac{n}{2\sigma^2} \left(s^2 + \frac{(b_m - \beta_m)^2}{\mathbf{D}n_{m,m}}\right)}. \quad (16)$$

On the other hand, in solving the uncertainty problem, inversely to the previous case, b_m and s are constants and β_m and σ are the random variables for which there is a joint PDF $p_{\beta}(\beta_m|\beta_m, s)$. Taking into account that both problems (estimation and uncertainty evaluation) relate to the same measurement experiment, only from different sides, we can assume that the probability elements $p_{b,s}(b_m, s|\beta_m, \sigma) \cdot db_m \cdot ds$ in the estimate problem and $p_{\beta,\sigma}(\beta_m, \sigma|b_m, s) \cdot d\beta_m \cdot d\sigma$ in the uncertainty problem should be consistent with each other, *i.e.*:

$$p_{\beta,\sigma}(\beta_m, \sigma|b_m, s) \cdot d\beta_m \cdot d\sigma = p_{b,s}(b_m, s|\beta_m, \sigma) \cdot db_m \cdot ds. \quad (17)$$

As can be seen from previous relationship (14), the estimate b_m and coefficient β_m appear as the difference $(b_m - \beta_m)/\sigma$ (β_m and σ are independent), while in (15) the estimate s and standard deviation σ occur as the ratio: s/σ . Therefore, elementary increments $db_m \cdot ds$ and $d\beta_m \cdot d\sigma$ depend on each other:

$$\frac{d}{db_m} \left(\frac{b_m - \beta_m}{\sigma}\right) = \frac{1}{\sigma}; \quad \frac{d}{d\beta_m} \left(\frac{b_m - \beta_m}{\sigma}\right) = -\frac{1}{\sigma}; \quad \frac{d}{ds} \left(\frac{s}{\sigma}\right) = \frac{1}{\sigma}; \quad \frac{d}{d\sigma} \left(\frac{s}{\sigma}\right) = -\frac{s}{\sigma^2}. \quad (18)$$

From (18), it follows that $db_m = -d\beta_m$ and $ds = -\frac{s}{\sigma} d\sigma$, therefore $db_m \cdot ds = \frac{s}{\sigma} \cdot d\beta_m \cdot d\sigma$, and the substitution of it in (17) gives the desired joint a posteriori PDF:

$$p_{\beta,\sigma}(\beta_m, \sigma|b_m, s) = \frac{s}{\sigma} \cdot p_{b,s}(b_m, s|\beta_m, \sigma). \quad (19)$$

It means that the joint PDF $p_{\beta,\sigma}(\beta_m, \sigma|b_m, s)$ of the coefficients $\boldsymbol{\beta} = (\beta_0, \beta_1, \dots, \beta_k)^T$ and σ (in the uncertainty problem) can be obtained by multiplying the joint PDF $p_{b,s}(b_m, s|\beta_m, \sigma)$ of the estimators $\mathbf{b} = (b_0, b_1, \dots, b_k)^T$ and s (in the estimation problem) by the ratio s/σ . After substituting (16) in (19), the required a posteriori joint PDF takes the form of:

$$p_{\beta,\sigma}(\beta_m, \sigma|b_m, s) = \frac{\sqrt{n}}{s^2 \sqrt{2\pi} \sqrt{\mathbf{D}n_{m,m}}} \frac{2 \cdot \left(\frac{n}{2}\right)^{\left(\frac{n-k-1}{2}\right)}}{\Gamma\left(\frac{n-k-1}{2}\right)} \cdot \left(\frac{s}{\sigma}\right)^{n-k+1} e^{-\frac{n \cdot s^2}{2\sigma^2} \cdot \left(1 + \frac{(b_m - \beta_m)^2}{s^2 \mathbf{D}n_{m,m}}\right)}. \quad (20)$$

It follows from (20) that, although the variables b_m and s in the estimation problem are independent, in the uncertainty problem the random variables β_m and σ are not independent.

Integrating (20) with σ gives the a posteriori PDF $p_{\beta}(\beta_m | b_m, s)$ of the coefficients:

$$p_{\beta}(\beta_m | b_m, s) = \frac{1}{s\sqrt{\pi}\sqrt{\mathbf{D}n_{m,m}}} \cdot \frac{\Gamma\left(\frac{n-k}{2}\right)}{\Gamma\left(\frac{n-k-1}{2}\right)} \cdot \frac{1}{\left[1 + \left(\frac{\beta_m - b_m}{s \cdot \sqrt{\mathbf{D}n_{m,m}}}\right)^2\right]^{\frac{n-k}{2}}} \quad (21)$$

This PDF is exactly the same as PDF (10), which was obtained using the Bayesian approach, but (21) was derived without assuming any a priori distributions. In addition, it should be noted that to obtain the standard uncertainty of the function, you also need the covariance of the coefficients β_m and β_l , which in this case is equal to: $\text{cov}(\beta_m, \beta_l) = \frac{s^2}{n-k-3} \mathbf{D}n_{m,l}$.

2.2. Analysis of Type A standard uncertainty

The comparison of (8) and (12) shows that standard uncertainty calculated according to (8) is only an approximation of the exact value of standard uncertainty (12) for a large number of observations: $n \gg k+4$ or $d \gg 3$. For a small number of degrees of freedom, standard uncertainty (8) may differ significantly from the exact value (12). This is shown in Fig. 1, which gives the dependence of standard uncertainty underestimation

$$\delta u_A(d) = \left| \frac{u_A(\beta_m | s)}{\sigma(b_m | s)} - 1 \right| \cdot 100\% = \left| \sqrt{\frac{d}{d-2}} - 1 \right| \cdot 100\% \quad (22)$$

of approximate value (8) on the number of degrees of freedom d .

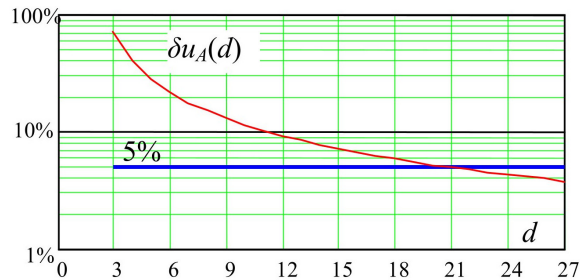


Fig. 1. Dependence of underestimation ($\delta u_A(d)$, %) of the approximate standard uncertainty (8) on the number of degrees of freedom $d = n - k - 1$.

When $n < k + 4$ or $d < 3$, the correct values of standard uncertainties $u_A(\beta_m)$ (8) cannot be determined, therefore in such a case standard uncertainty (8) makes no sense. When $n \geq k + 4$ or $d \geq 3$, the approximate value (8) may be much lower than the exact value (12). For example, if the number of observations is minimal *i.e.*, $n = k + 4$ or $d = 3$, then the approximate standard uncertainty $\tilde{u}_A(\beta_m)$ in (8) is smaller by $\sqrt{3}$ than the exact value $u_A(\beta_m)$, *i.e.*, it is over $\approx 73\%$ underestimated. Even if $n = k + 10$ or $d = 9$, then the approximate standard uncertainty is over 13% smaller than the exact value.

Here the question arises for what values of the number of degrees of freedom the standard uncertainty determined with the classical method may substitute the exact value of this uncertainty? The answer to this question can be obtained from [14], where **Specific Guidance 7.2.6** indicates

that the numerical values of the estimate y and its standard uncertainty $u_c(y)$ or expanded uncertainty U_p should not be given with an excessive number of digits. It usually suffices to quote $u_c(y)$ and U_p [as well as the standard uncertainties $u(x_i)$ of the input estimates x_i] to at most two significant digits, although in some cases it may be necessary to retain additional digits to avoid round-off errors in subsequent calculations. If the uncertainty is expressed with two significant digits, then the maximum rounding error has values not exceeding 4.76%. Therefore, we assume a limit value of $\approx 5\%$ for which an approximate uncertainty value can replace the exact value. It follows from this that only at approximately $n - k - 1 = d > 21$ the underestimation of the approximate standard uncertainty versus the exact one is less than 5% (Fig. 1).

In order to highlight the problems in the classical assessment of standard uncertainty related to the limited number of experimental points, we will analyze some examples from various literature sources. In numerical example 15.1.1 given in [28], the linear function ($k = 1$) is built using the $n = 4$ experimental points. As in this example $k = 1$ and $n = 4$, the number of degrees of freedom $d = 4 - 1 - 1 = 2 < 3$, then due to (10) and (11), the standard uncertainties $u_A(\beta_0)$, $u_A(\beta_1)$ of the coefficients β_0 , β_1 of linear function cannot be determined correctly. In another example – 16.1.1 in [28], the quadratic function ($k = 2$) is also built for the input data of $n = 4$ points. In this example $k = 2$ and $n = 4$, therefore $d = 4 - 2 - 1 = 1 < 3$, standard uncertainties $u_A(\beta_0)$, $u_A(\beta_1)$ and $u_A(\beta_2)$ of coefficients β_0 , β_1 , β_2 of the parabolic function cannot be determined correctly either. One can find the same problem in Example 1.1 in [11] which refers to the treatment of data from a study of effect of ozone pollution on soybean yield with a linear function ($k = 1$). Here, the number of observations is also $n = 4$, therefore also $d = 4 - 1 - 1 = 2 < 3$, thus it is impossible to determine correct values of the standard uncertainties.

In next examples the number of measurement points is sufficient: $n > k + 3$ or $d \geq 3$, but the standard uncertainties of coefficients and functions are determined with significant underestimation. The first one refers to Example H.3 Calibration of a thermometer in the Guide [14]: Linear calibration curve ($k = 1$). The $n = 11$ pairs of the data are given in the second and third columns of Table H.6 [14], *i.e.*, here $d = n - 1 - 1 = 9$. Because the standard uncertainties of the intercept and slope and also of the value predicted at $t = 30^\circ\text{C}$ in this example are determined by (8), these values are underestimated (12) by $(\sqrt{9/7} - 1) \cdot 100\% \approx 13.4\%$.

The problem arises to a greater extent in Example 8.5 in [29]. In this example the linear dependence ($k = 1$) of temperature T on pressure P is constructed on the basis of $n = 5$ pairs of temperature measurement results at given pressure values. Here $d = 3$, therefore the correct value (12) of the standard uncertainty of the coefficients is $\sqrt{3} \approx 1.73$ times bigger than presented in the example (underestimation is about 73%).

The next example is a linear function going through the origin that describes the relationship of the relative risk of individuals exposed to different levels of dust [11]. The linear function going through the origin ($\beta_0 = 0$) is built by $n = 9$ points. The number of degrees of freedom is $d = 9 - 1 = 8$, therefore in this example the standard uncertainty is underestimated by about $\approx 15\%$. In the last example 8.1 in [11], the cubic polynomial model ($k = 3$) is built. This example presents time serial analysis and relates to the treatment of algae density measurements over time. The $n = 14$ solutions were randomly assigned for measurement to each of 14 successive days of the study. In this example the number of degrees of freedom is $d = 14 - 3 - 1 = 10$, therefore, according to (12), the standard uncertainties of measurands are underestimated by about $\approx 11.8\%$.

After analysis of these examples, we can find that even when order k of the function is low but the number of experimental points is not enough large, the classical evaluation of Type A uncertainty is not fully correct. Namely, when the number of degrees of freedom $d < \approx 20$, the classical approach provides significant (over than 5% up to 73%) underestimation of standard uncertainties.

3. Evaluation of Type B uncertainty caused by systematic effects in both input and output quantities

In the following analysis it is assumed that the mathematical models of systematic effects mainly consist of additive errors Δx_0 in input quantity and Δy_0 in output quantity and multiplicative δ_{gx} , δ_{gy} in input and output quantities. Therefore, the results x_i , y_i of measurement of input and output quantities are:

$$x_i = X_i + \Delta x_0 + \delta_{gx} \cdot X_i, \quad y_i = Y_i + \Delta y_0 + \delta_{gy} \cdot Y_i. \quad (23)$$

In (23), the value of the input quantity can be approximated as:

$$X_i = \frac{x_i - \Delta x_0}{1 + \delta_{gx}} \approx x_i - \Delta x_0 - \delta_{gx} \cdot x_i. \quad (24)$$

Here, product $\delta_{gx} \cdot \Delta x_0$ is negligibly small, as of the second order of smallness, *i.e.*, it can be neglected: $\delta_{gx} \cdot \Delta x_0 \approx 0$. Therefore, having taken into account (24), the result of measurement y_i of output quantity Y_i can be presented in the form:

$$y_i = Y_i + \Delta y_0 + \delta_{gy} \cdot Y_i = F(x_i - \Delta x_0 - \delta_{gx} \cdot x_i) + \Delta y_0 + \delta_{gy} \cdot F(x_i - \Delta x_0 - \delta_{gx} \cdot x_i). \quad (25)$$

Therefore, for the **linear function** ($k = 1$): $Y1(x) = \beta_{10} + \beta_{11} \cdot X$, using (25), the observed values of output quantity can be presented in the form:

$$\begin{aligned} y_{1i} &\approx \beta_{10} + \beta_{11} \cdot x_i + \Delta y_0 + \delta_{gy} \cdot \beta_{10} - \beta_{11} \cdot \Delta x_0 + \beta_{11} (\delta_{gy} - \delta_{gx}) \cdot x_i \\ &= \beta_{10} + \beta_{11} \cdot x_i + \Delta 1_0 + \Delta 1_1 \cdot x_i. \end{aligned} \quad (26)$$

Here and below, parts of expressions relating to second and higher order components are assumed as negligibly small. The components which caused coefficient uncertainties in (26) are approximately equal:

$$\Delta 1_0 \approx \beta_{10} \cdot \delta_{gy} - \beta_{11} \cdot \Delta x_0 + \Delta y_0; \quad \Delta 1_1 \approx \beta_{11} \cdot (\delta_{gy} - \delta_{gx}). \quad (27)$$

From (27), even when the function is linear, the systematic additive and multiplicative effects in the results of measurement of input quantity are processed with different coefficients. Namely, in (27) the influence of additive component Δx_0 of the input quantity is determined by coefficient β_{11} , but the influence of additive component Δy_0 on the output quantity is determined by coefficient 1. Besides, the multiplicative component δ_{gy} of the output quantity affects the additive component by coefficient β_{10} ($\beta_{10} \cdot \delta_{gy}$), but it also affects multiplicative component by the other coefficient *i.e.*, β_{11} ($\beta_{11} \cdot \delta_{gy}$). More complicate situations appear for high order functions.

For the **quadratic function** ($k = 2$): $Y2(X) = \beta_{20} + \beta_{21} \cdot X + \beta_{22} \cdot X^2$ using (25), the observed values of output quantity can be presented in the form of:

$$\begin{aligned} y_{2i} &\approx \beta_{20} + \beta_{21} \cdot x_i + \beta_{22} \cdot x_i^2 + (\Delta y_0 + \delta_{gy} \cdot \beta_{20} - \beta_{21} \cdot \Delta x_0) + (\beta_{21} \cdot (\delta_{gy} - \delta_{gx}) - 2\beta_{22} \cdot \Delta x_0) \cdot x_i \\ &\quad + \beta_{22} \cdot (\delta_{gy} - 2\delta_{gx}) \cdot x_i^2 = \beta_{20} + \beta_{21} \cdot x_i + \beta_{22} \cdot x_i^2 + \Delta 2_0 + \Delta 2_1 \cdot x_i + \Delta 2_2 \cdot x_i^2. \end{aligned} \quad (28)$$

Therefore, the components, which caused uncertainties of corresponding function coefficients, are equal:

$$\begin{aligned} \Delta 2_0 &\approx \beta_{20} \cdot (\delta_{gy} - 0 \cdot \delta_{gx}) - 1 \cdot \beta_{21} \cdot \Delta x_0 + \Delta y_0; \\ \Delta 2_1 &\approx \beta_{21} \cdot (\delta_{gy} - 1 \cdot \delta_{gx}) - 2 \cdot \beta_{22} \cdot \Delta x_0; \quad \Delta 2_2 \approx \beta_{22} \cdot (\delta_{gy} - 2 \cdot \delta_{gx}). \end{aligned} \quad (29)$$

Using a similar procedure, error expressions for high order (k) functions can be derived:

$$\begin{aligned} \Delta_0 &\approx \beta_0 \cdot \delta_{gy} - \beta_1 \cdot \Delta x_0 + \Delta y_0; \\ \Delta_m &\approx \beta_m \cdot (\delta_{gy} - m \cdot \delta_{gx}) - m(k) \cdot \beta_{m+1} \cdot \Delta x_0, \quad m = 1, \dots, k, \end{aligned} \quad (30)$$

where $m(k) = \text{mod}(m + 1, k + 1)$.

From (27), (29) and (30) we can see that each error expression (except the last one when $m = k$) consists of the both additive and multiplicative systematic components. Besides, in each expression two different coefficients β_m, β_{m+1} ($m = 0, 1, \dots, k-1$) are presented. Due to [14] the type B standard uncertainty usually is determined using the MPE ($\pm\Delta_{\text{MPE}}$) of a verified instrument which often are given in manufacturer's specifications traditionally as: $\pm c \%$ (or *ppm*) of instrument indication (x) $+ \pm d\%$ (or *ppm*) of instrument range (R). In a given digital measuring instrument the characteristic curve of the errors must have values within $\pm\Delta_{\text{MPE}}$. Under such assumptions, the sum of additive and multiplicative error components satisfies the condition:

$$-\Delta_{\text{MPE}} = -(c \cdot x + d \cdot R) \leq \approx \Delta_0 + \delta_g \cdot x \approx \leq +(c \cdot x + d \cdot R) = +\Delta_{\text{MPE}}. \quad (31)$$

More often the distribution of variation of indication of measuring instrument within low level $-\Delta_{\text{MPE}} = -(c \cdot x + d \cdot R)$ and high level $+\Delta_{\text{MPE}} = +(c \cdot x + d \cdot R)$ is assumed as uniform [14]. In the case of two instruments of the same type (equal $\pm\Delta_{\text{MPE}}$ values), each has different values of systematic effects within the same limits $\pm\Delta_{\text{MPE}}$. For example, Fig. 2a shows the relationships $\Delta_1(x) = \Delta_{01} + \delta_{g1} \cdot x$ and $\Delta_2(x) = \Delta_{02} + \delta_{g2} \cdot x$ for a single-polarity measuring instrument. Both $\Delta_1(x)$ and $\Delta_2(x)$ are within $\pm\Delta_{\text{MPE}}$. From the physical point of view, based on the knowledge of MPE, there is no reason to assume in (31) a complete lack of correlation between the additive Δ_0 and multiplicative $\delta_g \cdot x$ components. Namely, when $\Delta(x) = \Delta_0 + \delta_g \cdot x$, assuming rectangular distribution and corresponding limits $\pm d \cdot R$ for Δ_0 and $\pm c$ for δ_g , without taking into account their correlation, standard deviation is:

$$\sigma_y(x) = \sqrt{\frac{d^2 \cdot R^2 + c^2 \cdot x^2}{3}}. \quad (32)$$

When $x = 0$, $\sigma(\Delta(0)) = d \cdot R / \sqrt{3}$, *i.e.*, it is correct, but when $x = R$ then standard deviation is $\sigma(\Delta(R)) = R \cdot \sqrt{\frac{d^2 + c^2}{3}} \neq R \cdot \frac{c + d}{\sqrt{3}}$. The last may be true when both components (offset and gain errors) are strongly correlated, which in praxis does not take place.

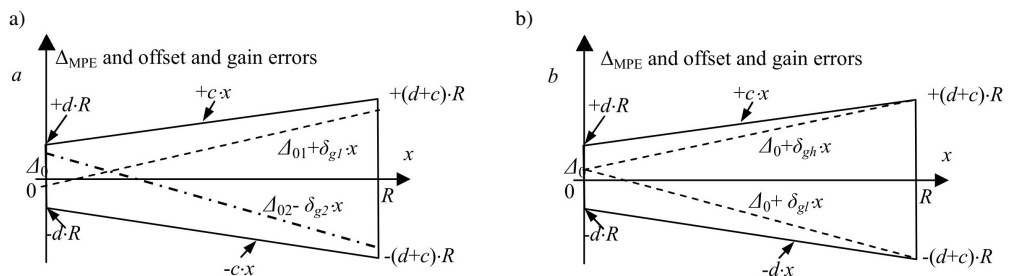


Fig. 2. Offset Δ_0 and gain $\delta_g \cdot x$ error components versus indication x for the known maximum permissible errors $\pm\Delta_{\text{MPE}}$: a) two different sums of both components; b) the ranges of changes between low δ_{gl} and high δ_{gh} values of multiplicative components when an additive component takes the value Δ_0 .

From (31) and Fig. 2b it follows that for a single-polarity instrument when $x = R$, the low level is $-\Delta_{\text{MPE}} = -R \cdot (d + c)$ and high level is $+\Delta_{\text{MPE}} = +R \cdot (d + c)$. Therefore, for an actual value Δ_0 : $-d \cdot R \leq \Delta_0 \leq +d \cdot R$, the value of δ_g must be within the limits:

$$\delta_{gl} = -\left(c + d + \frac{\Delta_0}{R}\right) \leq \delta_g \approx \leq \left(c + d - \frac{\Delta_0}{R}\right) = \delta_{gh}. \quad (33)$$

Thus, from (33) it follows that additive and multiplicative components should be partially dependent. Namely, when $\Delta_0 = +d \cdot R$, the maximum range of the gain error component is $-(c + 2d) \leq \delta_g \approx \leq c$. Similarly, when $\Delta_0 = -d \cdot R$, the maximum range of the gain error component is $-c \leq \delta_g \approx \leq (c + 2d)$. It means that the range of the gain error components depends on the actual value of the offset error component, and inversely. In other words, the value of the gain error is negatively correlated with the offset error. It is possible to show that under these assumptions the standard deviation of the multiplicative component and covariation of both components (δ_g and $\delta_0 = \Delta_0/R$) are:

$$\sigma_g = \sqrt{\frac{(c+d)^2 + d^2}{3}}, \quad \text{cov}(\delta_g, \Delta_0) = -\frac{d^2 \cdot R}{3}. \quad (34)$$

Therefore, taking into account that $\sigma_s(\Delta_0) = d \cdot R/\sqrt{3}$ and (34), depending on the indication of x , standard deviation caused by systematic components can be given as:

$$\sigma_s(x) = \frac{1}{\sqrt{3}} \sqrt{d^2 R^2 + ((c+d)^2 + d^2) x^2 - 2d^2 R \cdot x} = \frac{1}{\sqrt{3}} \sqrt{d^2 (R-x)^2 + (c+d)^2 x^2}. \quad (35)$$

In (35), when $x = 0$, standard uncertainty takes the value $\sigma_s(0) = d \cdot R/\sqrt{3}$, and when $x = R$, it takes the value $\sigma_s(R) = (c+d) \cdot R/\sqrt{3}$. Therefore, assuming that MPEs of both input and output errors are determined by c_x, d_x, x, R_x and c_y, d_y, y, R_y respectively, the covariance components, which the determined impact of systematic effects in (30), can be described analytically as:

$$u_{c_B}^2(\beta_0) = u_{c_B}^2(\beta_0, \beta_0) \approx \frac{b_0^2 \cdot (d_y + c_y)^2 + b_1^2 \cdot d_x^2 R_x^2 + d_y^2 (R_y - b_0)^2}{3}, \quad (36)$$

$$u_{c_B}^2(\beta_0, \beta_m) \approx \frac{1}{3} \left[b_0 \cdot b_m \left[d_y^2 + (d_y + c_y)^2 \right] - b_m R_y d_y^2 + b_1 R_x d_x^2 \left[m(k) \cdot b_{m+1} R_x - m \cdot b_m \right] \right], \quad (37)$$

$$m = 1, \dots, k,$$

$$u_{c_B}^2(\beta_m) = u_{c_B}^2(\beta_m, \beta_m) \approx \frac{b_m^2 \cdot \left[d_y^2 + (d_y + c_y)^2 + m^2 (d_x^2 + (d_x + c_x)^2) \right]}{3} + \frac{m(k) \cdot b_{m+1} R_x d_x^2 \cdot \left[m(k) \cdot b_{m+1} \cdot R_x - 2m \cdot b_m \right]}{3}, \quad (38)$$

$$u_{c_B}^2(\beta_l, \beta_m) \approx \frac{1}{3} \left[\begin{array}{l} b_l \cdot b_m \cdot \left[d_y^2 + (d_y + c_y)^2 + l \cdot m (d_x^2 + (d_x + c_x)^2) \right] \\ + R_x d_x^2 \cdot \left[(l+1) \cdot m(k) \cdot b_{l+1} \cdot b_{m+1} R_x \right. \\ \left. - l \cdot m(k) b_l \cdot b_{m+1} - (l+1) \cdot m \cdot b_{l+1} \cdot b_m \right] \end{array} \right] \quad \begin{array}{l} l = 1, \dots, m-1, \\ m = 2, \dots, k. \end{array} \quad (39)$$

4. Combined standard uncertainties of the functions

The combined standard uncertainties of functions can be determined by standard procedure [14], *i.e.*, as the square root of the sum of the squares of both components. Using the

Type A component (12) and Type B components (36)–(39), the combined standard uncertainties of coefficients and functions of order k are given as:

$$u_c(\beta_m) = \sqrt{\frac{s_k^2}{n-k-3} \cdot \mathbf{D}n_{m,m} + u_{cB}^2(\beta_m)}, \quad m = 0, 1, \dots, k, \quad (40)$$

$$u_c(y(x)) = \sqrt{\frac{s_k^2}{n-k-3} \cdot \sum_{m=0}^k \sum_{l=0}^k \mathbf{D}n_{m,l} x^{m+l} + \sum_{m=0}^k \sum_{l=0}^k u_{cB}^2(\beta_m, \beta_l) x^{m+l}}. \quad (41)$$

It should be noted that if the nonlinearity of the function (1) is not significant, *i.e.*, the influence of its second- and higher-order components compared to the linear component is small, then the second component of the combined uncertainty in (41) mainly depends on the constant and linear components, *i.e.*, it consists of only three terms:

$$u_c(y(x)) \approx \sqrt{\frac{s_k^2}{n-k-3} \cdot \sum_{j=0}^k \sum_{i=0}^k \mathbf{D}n_{m,l} x^{m+l} + u_{cB}^2(\beta_0) + 2u_{cB}^2(\beta_0, \beta_1) \cdot x + u_{cB}^2(\beta_1) \cdot x^2}. \quad (42)$$

5. Result verifications by Monte-Carlo method

The *Monte Carlo method* (MCM) is a practical alternative to the GUM uncertainty framework and mainly can be applied when a) linearization of the model provides an inadequate representation, or b) the PDF for the output quantities and also for the input ones departs appreciably from Gaussian distribution [30]. Since in the investigated problem of systematic effects in the measurement results of the input quantity cannot be linearly introduced into the model for calculating the coefficients of a polynomial function and their uncertainties, and in addition, these effects are usually characterized by uniform or other distributions, *i.e.*, not Gaussian, the Monte Carlo method is used to verify the results obtained in this paper. The object of verification is the nominal quadratic function:

$$Y_n(X) = \beta_{0n} + \beta_{1n} \cdot X + \beta_{2n} \cdot X^2 = 100 + 0.39702 \cdot X - 5.8893 \cdot 10^{-5} \cdot X^2, \quad (43)$$

with coefficients: $\beta_{0n} = 100.00$, $\beta_{1n} = 0.39702$, $\beta_{2n} = -5.8893 \cdot 10^{-5}$.

A basic analysis of the impact of type B uncertainty components was performed with the following measurement instrument's data:

- measurements of input quantity X : main parameters are: Range $R_x = 300$, MPE: $c_x = \pm 0.025\%$ of Reading (x), $d_x = \pm 0.033\%$ of range R_x ;
- measurements of output quantity Y : main parameters are: Range $R_y = 1000$, MPE: $c_y = \pm 0.017\%$ of Reading (y), $d_y = \pm 0.001\%$ of range R_y .

The study was carried out in was stages.

Stage 1. At this stage the correctness of the derived formulas (30) was examined, according to which the influence of uncorrected systematic effects on the measurement results of the input and output quantities was evaluated. At this stage, the random effects in the measurement results were not taken into account (in MCM simulation $\sigma_n = 0.00001$).

The $n = 13$ values of the input quantity were calculated, of which the results are: $x_{0,i} = 0 \quad 25 \quad 50 \quad 75 \quad 100 \quad 125 \quad 150 \quad 175 \quad 200 \quad 225 \quad 250 \quad 275 \quad 300$.

For these values, the nominal values of the output quantity were determined according to the nominal function (43): $y_{n,i} = 100.000 \ 109.889 \ 119.704 \ 129.445 \ 139.113 \ 148.707 \ 158.228 \ 167.675 \ 177.048 \ 186.348 \ 195.574 \ 204.727 \ 213.806$.

According to the instrument's MPE values (c_x , d_x and c_y , d_y) given above, the ranges R_x and R_y and indications within their respective limits, the $M = 30\,000$ sets ($j = 0, 1, \dots, M - 1$) of values of additive (offset error) $\Delta_{0x,j}$, $\Delta_{0y,j}$, and multiplicative (gain error) $\delta_{gx,j}$, $\delta_{gy,j}$ effects were generated, which determine the corresponding effects in the measurement results of both quantities:

$$\Delta(X)_{i,j} = \Delta_{0x,j} + \delta_{gx,j} \cdot X_i, \quad \Delta(Y)_{i,j} = \Delta_{0y,j} + \delta_{gy,j} \cdot Y_i. \quad (44)$$

The simulated values of the output quantity taking into account these effects were determined according to (25) ($i = 0, 1, \dots, 12$; $j = 0, 1, \dots, M - 1$):

$$y_{i,j} = \left[\beta_{0n} + \beta_{1n} \cdot \left[x_{0,i} (1 - \delta_{gx,j}) - \Delta_{0x,j} \right] + \beta_{2n} \left[x_{0,i} (1 - \delta_{gx,j}) - \Delta_{0x,j} \right]^2 \right] \cdot (1 + \delta_{gy,j}) + \Delta_{0y,j}. \quad (45)$$

Using a vector of the $n = 13$ values of the input quantity from (3) and (4) the matrices \mathbf{M}_n and \mathbf{D}_n are calculated:

$$\mathbf{M}_n = \begin{pmatrix} 1 & 150 & 31250 \\ 150 & 31250 & 7.313 \cdot 10^6 \\ 31250 & 7.313 \cdot 10^6 & 1.824 \cdot 10^9 \end{pmatrix}, \quad (46)$$

$$\mathbf{D}_n = \begin{pmatrix} 6.71 & -0.086 & 2.29 \cdot 10^{-4} \\ -0.086 & 1.61 \cdot 10^{-3} & -4.99 \cdot 10^{-6} \\ 2.29 \cdot 10^{-4} & -4.99 \cdot 10^{-6} & 1.66 \cdot 10^{-8} \end{pmatrix}.$$

After substituting the values of $y_{i,j}$ (45) into (4), the M sets of values of the estimated coefficients ($\mathbf{b} = \mathbf{D}_n \cdot \mathbf{Y}_n$) were determined:

$$\mathbf{b}_j^T = (b_{0,j}; b_{1,j}; b_{2,j}), \quad j = 0, 1, \dots, M-1. \quad (47)$$

Then, based on the given values of coefficients (43) and determined \mathbf{b}_j (47), M set error values $\Delta b_{m,j}$ were calculated:

$$\Delta b_{0,j} = b_{0,j} - \beta_0; \quad \Delta b_{1,j} = b_{1,j} - \beta_1; \quad \Delta b_{2,j} = b_{2,j} - \beta_2. \quad (48)$$

At the same time, the expected error values $\Delta 2_0$, $\Delta 2_1$ and $\Delta 2_2$ of these coefficients were calculated directly according to approximate expressions (29). The values of the first 6 ($j = 0, 1, \dots, 5$) sets of estimated errors (48) and directly calculated errors (29) are given in Tables 1 and 2. From a comparison of the error values given in these tables, we see their complete convergence. It means that the simplifications assumed in (24)–(30) for determining the impact of systematic effects are quite correct.

Table 1. Values of estimated errors (48) of coefficients, $m = 0, 1, 2$, $j = 0, 1, \dots, 5$.

	0	1	2	3	4	5
$\Delta b =$	$-4.409 \cdot 10^{-3}$	0.023	$5.267 \cdot 10^{-3}$	-0.02	-0.043	-0.027
	$1.979 \cdot 10^{-4}$	$-1.538 \cdot 10^{-4}$	$-9.806 \cdot 10^{-5}$	$9.574 \cdot 10^{-5}$	$5.628 \cdot 10^{-5}$	$1.236 \cdot 10^{-4}$
	$-6.127 \cdot 10^{-8}$	$3.681 \cdot 10^{-8}$	$2.659 \cdot 10^{-8}$	$-3.022 \cdot 10^{-8}$	$-1.738 \cdot 10^{-8}$	$-4.463 \cdot 10^{-8}$

Table 2. Values of directly determined errors (29) of coefficients.

	0	1	2	3	4	5
$\Delta_2 =$	$-4.415 \cdot 10^{-3}$	0.023	$5.256 \cdot 10^{-3}$	-0.02	-0.043	-0.027
	$1.98 \cdot 10^{-4}$	$-1.538 \cdot 10^{-4}$	$-9.787 \cdot 10^{-5}$	$9.565 \cdot 10^{-5}$	$5.625 \cdot 10^{-5}$	$1.235 \cdot 10^{-4}$
	$-6.155 \cdot 10^{-8}$	$3.676 \cdot 10^{-8}$	$2.602 \cdot 10^{-8}$	$-2.986 \cdot 10^{-8}$	$-1.732 \cdot 10^{-8}$	$-4.396 \cdot 10^{-8}$

Stage 2. At this stage, the simultaneous influence of both systematic and random components on the combined uncertainty of the parabolic function was examined.

- For this purpose, there were generated uncorrelated random $un_{i,j}$ ($i = 0, 1, \dots, 12; j = 0, 1, \dots, M-1$) values with normal distribution with zero expected value and standard deviation σ_n , which were added to the values $y_{n,i}$ of the output quantity. The study was carried out for four values of standard deviation of random component: $\sigma_n = 0.01; 0.0316; 0.1; 0.316$, in the following order.
- For each value of σ_n there were determined M sets values of:
 - coefficients: $b_{0,j}, b_{1,j}, b_{2,j}$ as in (4);
 - errors $\Delta b_{0,j}; \Delta b_{1,j}; \Delta b_{2,j}$ of these coefficients as in (48);
 - errors of estimated quadratic function:

$$\Delta y_{i,j} = \Delta b_{0,j} + \Delta b_{1,j} \cdot x_i + \Delta b_{2,j} \cdot x_i^2; \quad (49)$$

- estimated standard deviations of the coefficients:

$$s(b_m) = \sqrt{\frac{1}{M-1} \sum_{j=0}^{M-1} (b_{m,j} - \bar{b}_m)^2}, \quad \bar{b}_m = \frac{1}{M} \sum_{j=0}^{M-1} b_{m,j}, \quad m = 0, 1, 2; \quad (50)$$

- combined standard uncertainties of the coefficients as in (40):

$$u_c(\beta_{m,j}) = \sqrt{u_A^2(\beta_{m,j}) + u_{cB}^2(\beta_{m,j})}, \quad m = 0, 1, 2, \quad j = 0, \dots, M-1. \quad (51)$$

- The standard deviations $s(b_m)$ (50) (including the coefficient $\sqrt{\frac{n-3}{n-5}}$ (13)) were compared to the mean values $\overline{u_c(\beta_m)}$ of standard uncertainties (51). The results of the comparison are given in Table 3. They show a very good convergence of standard deviations $s(b_m)$ and mean standard uncertainty $\overline{u_c(\beta_m)}$ values.
- Next, there were determined the estimates of expanded (for a confidence level $p = 0.95$) uncertainties:
 - of the coefficients:

$$U_{p,e}(\beta_m) = k_p \cdot u_e(\beta_m), \quad m = 0, 1, 2; \quad (52)$$

- of the quadratic function:

$$U_{p,e}(Y_i) = k_p \cdot u_e(Y_i), \quad i = 1, 2, \dots, n. \quad (53)$$

Table 3. Estimated standard deviations $s(b_m)$ of coefficients and the mean values of standard uncertainties $\overline{u_c(\beta_m)}$ and also the values of estimated probabilities $p_e(\beta_m)$ and $p_e(y)$, when coefficients errors and function errors are within the corresponding limits determined by expanded uncertainties $U_{0.95,e}(\beta_m)$ and $U_{0.95,e}(Y(x))$, as functions of standard deviation σ_n of random effect.

σ_n	0.01	0.0316	0.1	0.316
$\frac{s(b_0) \cdot \sqrt{\frac{n-3}{n-5}}}{u_c(\beta_0)}$	$\frac{0.030}{0.027}$	$\frac{0.038}{0.036}$	$\frac{0.086}{0.084}$	$\frac{0.253}{0.256}$
$\frac{s(b_1) \cdot \sqrt{\frac{n-3}{n-5}}}{u_c(\beta_1)}$	$\frac{2.23 \cdot 10^{-4}}{2.05 \cdot 10^{-4}}$	$\frac{4.33 \cdot 10^{-4}}{4.17 \cdot 10^{-4}}$	$\frac{1.29 \cdot 10^{-3}}{1.25 \cdot 10^{-3}}$	$\frac{3.87 \cdot 10^{-3}}{3.95 \cdot 10^{-3}}$
$\frac{s(b_2) \cdot \sqrt{\frac{n-3}{n-5}}}{u_c(\beta_2)}$	$\frac{4.08 \cdot 10^{-7}}{4.02 \cdot 10^{-7}}$	$\frac{1.32 \cdot 10^{-6}}{1.27 \cdot 10^{-6}}$	$\frac{4.03 \cdot 10^{-6}}{3.99 \cdot 10^{-6}}$	$\frac{1.24 \cdot 10^{-5}}{1.27 \cdot 10^{-5}}$
$p_e(\beta_0)$	0.955	0.955	0.952	0.952
$p_e(\beta_1)$	0.952	0.952	0.952	0.952
$p_e(\beta_2)$	0.949	0.950	0.951	0.955
$p_e(y)$	0.947	0.951	0.953	0.951

5. The values $\Delta b_{0,j}$; $\Delta b_{1,j}$; $\Delta b_{2,j}$ of the calculated errors were compared with the expanded uncertainty values $U_{p,e}(\beta_m)$ and estimates $p_e(\Delta b_m)$ of the probability of not exceeding these errors of the respective expanded uncertainty values $U_{p,e}(\beta_m)$ were determined:

$$p_e(\Delta b_m) = \frac{\sum_{j=0}^{M-1} f_{m,j}}{M}, \quad f_{m,j} = \begin{cases} 1, & \text{if } |\Delta b_{m,j}| \leq U_{p,e}(\beta_m), \\ 0, & \text{otherwise,} \end{cases} \quad m = 0, 1, 2. \quad (54)$$

The values of $p_e(\Delta b_m)$ are presented in Table 3. We can see that estimated probability values $p_e(\Delta b_m)$ are very close to the given value $p = 0.95$.

6. On the basis of estimated coefficients, the standard uncertainties of the quadratic function at each value x_i of the input quantity were determined:

$$u_e(Y_i) = \sqrt{\frac{1}{M-1} \sum_{j=0}^{M-1} \left(b_{0,j} - \overline{b_0} + (b_{1,j} - \overline{b_1}) \cdot x_i + (b_{2,j} - \overline{b_2}) \cdot x_i^2 \right)^2}, \quad i = 1, 2, \dots, n. \quad (55)$$

7. Taking into account analytical components of matrix \mathbf{Dn} in (4), the analytical Type A standard uncertainty of quadratic function was determined:

$$u_A^2(x) = \frac{\sigma_n^2}{n} \left[1 + 3 \frac{n-1}{n+1} \left(\frac{x-\bar{x}}{Vx} \right)^2 + \frac{5}{4} \cdot \frac{n^2-1}{n^2-4} \left(3 \frac{n-1}{n+1} \left(\frac{x-\bar{x}}{Vx} \right)^2 - 1 \right)^2 \right], \quad (56)$$

$$\text{where } \bar{x} = \frac{1}{n} \sum_{l=1}^n x_l = 150, \quad Vx = \frac{x_n - x_1}{2} = 150.$$

8. Based on parameters c_x , d_x and c_y , d_y and the ranges R_x , R_y of instruments, the components (36), (37) and (38) of Type B combined standard uncertainties of coefficients were determined:

$$u_{B,a}^2(\beta_0) = \frac{\beta_0^2 \cdot (d_y + c_y)^2 + \beta_1^2 \cdot d_x^2 R_x^2 + d_y^2 (R_y - \beta_0)^2}{3} \approx 6.324, \quad (57)$$

$$u_{B,a}^2(\beta_0, \beta_1) = \frac{\beta_0 \cdot \beta_1 \left[d_y^2 + (d_y + c_y)^2 \right] - \beta_1 R_y d_y^2 - \beta_1 R_x d_x^2 (\beta_1 - 2\beta_2 R_x)}{3} \approx -0.0147, \quad (58)$$

$$u_{B,a}^2(\beta_1) = \frac{\beta_1^2 \cdot \left[d_y^2 + (d_y + c_y)^2 + d_x^2 + (d_x + c_x)^2 \right] - 4\beta_2 R_x d_x^2 (\beta_1 - \beta_2 R_x)}{3} \approx 2.648 \cdot 10^{-4}. \quad (59)$$

9. After substituting both Type A (56) and Type B (57)–(59) components in (42), the analytical expression for the expanded ($p = 0.95$) uncertainty $U_{p,a}(Y(x))$ of the function takes the form:

$$U_{0.95,a}(Y(x)) = 1.96 \sqrt{\frac{\sigma_n^2}{13} \left[1 + \frac{18}{7} \left(\frac{x-150}{150} \right)^2 + \frac{14}{11} \left(\frac{18}{7} \left(\frac{x-150}{150} \right)^2 - 1 \right)^2 \right] + \frac{6.324 - 2 \cdot 0.0147 \cdot x + (2.648 \cdot 10^{-4})^2 \cdot x^2}{3 \cdot (100\%)^2}}. \quad (60)$$

The $U_{0.95,e}(Y(x_i))$ values of the estimated expanded uncertainty (53) and $U_{0.95,a}(Y(x_i))$ values of the analytical uncertainty (60) as a function of x_i values are presented in Fig. 3.

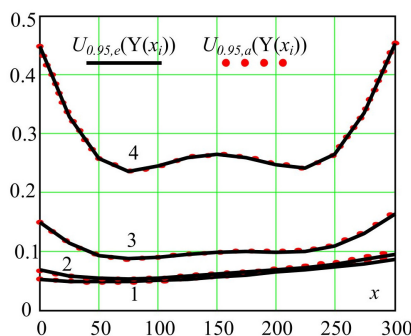


Fig. 3. Estimated expanded uncertainty $U_{0.95,e}(Y(x_i))$ (solid black line) and a priori expanded uncertainty $U_{0.95,a}(Y(x_i))$ (red dots) for $p = 0.95$ and $\sigma_n = 0.010$ (1), $\sigma_n = 0.0316$ (2), $\sigma_n = 0.10$ (3), $\sigma_n = 0.316$ (4).

10. The values $\Delta y_{i,j}$ of the calculated errors (49) were compared with the expanded uncertainty values $U_{0.95,e}(Y_i)$ (53) and estimates $p_e(\Delta y)$ of the probability of not exceeding these errors of the expanded uncertainty values $U_{0.95,e}(Y_i)$ were determined:

$$p_e(\Delta y) = \frac{\sum_{j=0}^{M-1} \sum_{i=0}^{n-1} d_{i,j}}{M}, \quad d_{i,j} = \begin{cases} 1, & \text{if } |\Delta y_{i,j}| \leq U_{p,e}(Y_i), \\ 0, & \text{otherwise.} \end{cases} \quad (61)$$

The values of estimated probability $p_e(\Delta y)$ are given in Table 3. We can see that the estimated values of $p_e(\Delta y)$ also are very close to the given value $p = 0.95$.

Fig. 3 also shows a very good convergence between the estimated $U_{0.95,e}(Y(x_i))$ and analytical $U_{0.95,a}(Y(x_i))$ expanded uncertainties. Namely, the maximum relative values of the differences between these uncertainties are about: 6% when $\sigma_n = 0.01$, 6% when $\sigma_n = 0.0316$, 2% when $\sigma_n = 0.1$, and 0.3% when $\sigma_n = 0.316$. In addition, we can see from this figure that at a small level of random component ($\sigma_n = 0.01$, $2\sigma_n/Y_{\max} \approx 0.009\%$), the main factors determining the level of uncertainty are systematic effects in the measurement results. In contrast, if the level of the random component increases significantly ($\sigma_n = 0.316$, $2\sigma_n/Y_{\max} \approx 0.3\%$), then the random factor determines the level of uncertainty.

6. Conclusions

The conclusions formulated below refer to the evaluation of the uncertainties caused by random independent noises distorting the output quantity and the uncorrected additive and multiplicative systematic effects in the measurements of both the input and output quantities of the systems, which are described by a polynomial function whose parameters are determined by the experimental data.

1. According to definition of uncertainty in the Guide [14], the standard uncertainty $u_A(\beta_m | s)$ of the polynomial function coefficient β_m can be determined correctly only on the basis of its PDF $p_\beta(\beta_m | b_m, s)$. It was shown that this PDF can be derived directly from integrating a joint PDF $p_{\beta,\sigma}(\beta_m, \sigma | b_m, s)$ which can be obtained from PDF $p_{b,s}(b_m, s | \beta_m, \sigma)$ (of estimates b_m of coefficient and s of standard deviation σ) multiplied by a ratio s/σ . In the presented method, no a priori distribution is required.

Because the values of type A standard uncertainty depend on $\sqrt{n-k-3} = \sqrt{d-2}$ (where $d = n - k - 1$ is the number of degrees of freedom) when the number of degrees of freedom is low (less than about 20), determination of standard uncertainty with classical approach provides its underestimation from about 5% when $d = 20$ up to 73% when $d = 3$. Based on the analysis of a number of examples, taken from literature sources, it was found that in these examples, due to the relatively small number of experimental points, the standard uncertainty values determined with the classical approach are less than correct between ten and several tens of percent or cannot be determined.

2. A very simple way of estimating the component of type B uncertainty caused by uncorrected systematic effects in the results of measurements of input and output quantities has been proposed. For this purpose, analytical expressions of coefficients errors dependent on the additive Δ_{0x} , Δ_{0y} and multiplicative δ_{gx} , δ_{gy} systematic components in measurements of both input X and output Y quantities were derived. It has been shown that the impact of these components of errors on error Δ_m of coefficient b_m depends on the values of the function coefficients β_m, β_{m+1} ($m = 0, 1, \dots, k-1$). On the basis of these analytical expressions of errors, formulas for type B standard uncertainties were derived in the function of declared in the manufacturer's specification values of the MPE of the measuring instruments used. Namely, the type B standard uncertainty components are determined as a function of parameters: c_x, c_y of instruments indications (x, y) and d_x, d_y of instruments ranges (R_x, R_y).
3. Monte Carlo simulation studies carried out for the estimation (with both Type A and Type B methods) of the corresponding uncertainty components of coefficients of the quadratic function showed a very good convergence with the derived analytical relationships.

Namely, from the results obtained by the MCM we can see that: (i) the derived relationships (30) for calculating coefficient errors caused by the influence of systematic additive and multiplicative effects on the measurement results of input and output quantities are fully correct, and (ii) and formulas (36)–(39) based on them, used for the calculation of the components of type B uncertainty, are also correct.

Studies of the simultaneous impact of both factors (random and systematic) were carried out in the form of a comparison of determined errors of coefficients and functions with the estimated values and their analytically determined expanded uncertainties. Namely, the maximum relative error between the estimated and analytically determined expanded uncertainties ranges from a few tenths of a percent to several percent.

4. The results obtained and presented in the paper are very important for practical applications because when determining the parameters of a polynomial (and other) function on the basis of the measurement results of both input and output quantities, it is not always possible to omit the instrumental components of uncertainty, and in many cases these components may be dominant in the combined uncertainties of coefficients and functions.

Acknowledgements

This publication is supported by Polish Ministry of Education and Science under the program “Regional Initiative of Excellence” in 2019–2023 (project number 027/RID/2018/19).

References

- [1] Meijer, G. (Ed.). (2008). *Smart sensor systems*. John Wiley & Sons.
- [2] Doebelin, E. O. (1990). *Measurement systems. Application and design*. McGraw-Hill Publish.
- [3] Piotrowski, J. (1992). *Theory of physical and technical measurement*. Elsevier.
- [4] Islam, T., & Mukhopadhyay, S. C. (2019). Linearization of the sensors characteristics: A review. *International Journal on Smart Sensing and Intelligent Systems*, 12(1), 1–21. <https://doi.org/10.21307/ijssis-2019-007>
- [5] Tomczyk, K. (2016). Problems in modelling charge output accelerometers. *Metrology and Measurement Systems*, 23(4), 645–659. <https://doi.org/10.1515/mms-2016-0045>
- [6] Hajiyev, C. (2016). Sensor calibration design based on D-optimality criterion. *Metrology and Measurement Systems*, 23(3), 413–424. <https://doi.org/10.1515/mms-2016-0029>
- [7] Lira, I., & Grientschnig, D. (2017). Error-in-variables models in calibration. *Metrologia*, 54(6), S133. <https://doi.org/10.1088/1681-7575/aa8f02>
- [8] Zakharov, I., Neyezhnikov, P., Semenikhin, V., & Warsza, Z. L. (2022). Measurement Uncertainty Evaluation of Parameters Describing the Calibrated Curves. In *Automation 2022: New Solutions and Technologies for Automation, Robotics and Measurement Techniques* (pp. 391–398). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-031-03502-9_38
- [9] Sachin Date. (2022). *Time Series Analysis, Regression and Forecasting – With tutorials in Python*. <http://timeseriesreasoning.com>
- [10] Silva, M. A., Amado, C., Ribeiro, Á., & Loureiro, D. (2022). Uncertainty evaluation in time-dependent measurements. *Measurement*, 196, 111196. <https://doi.org/10.1016/j.measurement.2022.111196>

- [11] Rawlings, J. O., Pantula, S. G., & Dickey, D. A. (1998). *Applied Regression Analysis: A Research Tool, Second Edition*. Springer-Verlag.
- [12] Draper, N. R., & Smith, H. (1998). *Applied regression analysis*. John Wiley & Sons.
- [13] Kendall, M. G., & Stuart, A. (1966). *The advanced theory of statistics. Inference and relationship. Second Edition*. Chares Griffin and Company Limited.
- [14] Joint Committee for Guides in Metrology. (2008). *Evaluation of measurement data – Guide to the expression of uncertainty in measurement* (JCGM 100:2008). http://www.bipm.org/utis/common/documents/jcgm/JCGM_100_2008_E.pdf
- [15] Chunovkina, A., & Stepanov, A. (2019, May). Estimation of linear regression confidence bands in case of correlated noise. In 2019 12th International Conference on Measurement (pp. 58–61). IEEE. <https://doi.org/10.23919/MEASUREMENT47340.2019.8779916>
- [16] Warsza, Z. L., & Puchalski, J. (2021). Uncertainty bands of the regression line for autocorrelated data of dependent variable Y. In *Automation 2021: Recent Achievements in Automation, Robotics and Measurement Techniques* (pp. 364–386). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-74893-7_33
- [17] Klauenberg, K., Martens, S., Bošnjaković, A., Cox, M. G., van der Veen, A. M., & Elster, C. (2022). The GUM perspective on straight-line errors-in-variables regression. *Measurement*, 187, 110340. <https://doi.org/10.1016/j.measurement.2021.110340>
- [18] Puchalski, J. G. (2021). A new algorithm for generalization of least square method for straight line regression in Cartesian system for fully correlated both coordinates. *International Journal of Automation, Artificial Intelligence and Machine Learning*, 2(2), 20–54.
- [19] Forbes, A. B. (2009). Parameter estimation based on least-squares methods. In F. Pavese and A. B. Forbes (Eds.), *Data modeling for metrology and testing in measurement science*. Birkauer-Boston. https://link.springer.com/chapter/10.1007/978-0-8176-4804-6_5
- [20] Warsza Z. L., & Puchalski, J. (2021). Uncertainty Bands of the Regression Line for Data with Type A and Type B Uncertainties of Dependent Variable Y. In R. Szewczyk et al. (Eds.), *Automation 2021: Recent Achievements in Automation, Robotics and Measurement Techniques* (vol. 1390, pp. 342–363). Springer. https://doi.org/10.1007/978-3-030-74893-7_32
- [21] Kacker, R., & Jones, A. (2003). On use of Bayesian statistics to make the Guide to the Expression of Uncertainty in Measurement consistent. *Metrologia*, 40(5), 235. <https://doi.org/10.1088/0026-1394/40/5/305>
- [22] Weise, K., & Woger, W. (1993). A Bayesian theory of measurement uncertainty. *Measurement Science and Technology*, 4(1), 1. <https://doi.org/10.1088/0957-0233/4/1/001>
- [23] Bich, W., Cox, M., & Michotte, C. (2016). Towards a new GUM – an update. *Metrologia*, 53(5), S149. <https://doi.org/10.1088/0026-1394/53/5/149>
- [24] Cox, M., & Shirono, K. (2017). Informative Bayesian Type A uncertainty evaluation, especially applicable to a small number of observations. *Metrologia*, 54(5), 642. <https://doi.org/10.1088/1681-7575/aa787f>
- [25] Jeffreys H., (1983). *Theory of probability (3rd edition)*. Oxford University Press.
- [26] Dorozhovets, M. (2020). Forward and inverse problems of Type A uncertainty evaluation. *Measurement*, 165, 108072. <https://doi.org/10.1016/j.measurement.2020.108072>
- [27] Evans, M., Hastings, N., & Peacock, B. (2000). *Statistical Distributions 3rd Edition*. Wiley.

M. Dorozhovets: UNCERTAINTY OF THE CONVERSION FUNCTION CAUSED BY SYSTEMATIC EFFECTS IN MEASUREMENTS...

- [28] Pollard, J. H. (1977). *Handbook of Numerical and Statistical Techniques*. Cambridge University Press.
- [29] Taylor, J. R. (1982). *An introduction to error analysis*. University Science Books Mill Valley.
- [30] Joint Committee for Guides in Metrology. (2008). Evaluation of measurement data – Supplement 1 to the Guide to the Expression of Uncertainty in Measurement’ – propagation of distributions using a Monte Carlo method.



Mykhaylo Dorozhovets received the Ph.D. degree from Lviv Polytechnic National University, Ukraine, in 2001. He is currently Full Professor at the Department of Metrology and Diagnostic Systems of Rzeszów University of Technology, Poland and the Department of Information Measuring Technology of Lviv Polytechnic National University, Ukraine. He has authored or coauthored 12 books, over 100 journal and 100 conference publications. His current research interests include measurement

data processing, uncertainty of measurement and industrial tomography.