

MEASUREMENT SYSTEM RELIABILITY ASSESSMENT

Ryszard Kłos

Polish Naval Academy, Department of Underwater Work Technology in Gdynia, Poland

ABSTRACT

Decision-making in problem situations is based on up-to-date and reliable information. A great deal of information is subject to rapid changes, hence it may be outdated or manipulated and enforce erroneous decisions. It is crucial to have the possibility to assess the obtained information. In order to ensure its reliability it is best to obtain it with an own measurement process. In such a case, conducting assessment of measurement system reliability seems to be crucial. The article describes general approach to assessing reliability of measurement systems.

Key words: analysis of Gage R&R, Gage Reliability, Gage Repeatability.

ARTICLE INFO

PolHypRes 2015 Vol. 51 Issue 2 pp. 31-46

ISSN: 1734-7009 eISSN: 2084-0535

DOI: 10.1515/phr-2015-0009

Pages: 16, figures: 6, tables: 1

page **www** of the periodical: www.phr.net.pl

Original article

Delivery date: 18.04.2015r.

Date of approval for print: 27.05.2015r.

Publisher

Polish Hyperbaric Medicine and Technology Society

INTRODUCTION

The article is a continuation¹ of a discussion commenced already quite some time ago within Polish Hyperbaric Research magazine on the method of assessment of measurement systems reliability without the use of reference materials – *Gage R&R*² [1,2]. *Gage R&R* methods are an attempt to utilise stochastic methods³ in search of sources of variability in measurement results obtained with particular measurement systems and later having an effect on their reliability, which should be suited to quality requirements *CTQ*⁴.

Literature on *Gage R&R* methods distinguishes two main trends. The first is formed by utilitarian studies [3], whereas the second provides a grouping of exhaustive statistical analyses [4,5]. Not many works on *Gage R&R* methods make reference to reliability assessment of the measurement process by comparison with reference materials, and such methods are commonly attributed to works on metrology [6].

The article presents an approximate assessment of measurement system reliability with *Gage R&R* methods utilising etalons and constituting standard reference materials for correct values $\hat{\mu}$ for various physicochemical quantities X .

The correct value $\hat{\mu}$ should be understood as a quantitative estimation of the value reproduced by etalon, burdened with negligible systematic errors⁵ ε_s from the perspective of the defined applications. This means that although the reproduced correct value $\hat{\mu}$ is burdened with residual systematic errors ε_s , in particular applications⁶ it may be seen as sufficient⁷ estimation of the true value⁸ μ .

METHOD

Measurement system

In previous works it was mentioned that in the assessment of a measurement system⁹ it is highly important to evaluate both the reliability of a measuring instrument and equipment¹⁰ as well as that of a *SOP*¹¹ [1,2]. Despite a significant progress, measuring systems usually use the same physiochemical phenomena. Without sensor replacement, using modernised electronic support systems for the measuring process will only have a limited impact on the thus far encountered limitations of measurement methods.

Knowledge of physiochemical phenomena underlying operation of the used sensors is crucial in the assessment of possibilities to increase reliability of measurement systems. What is continuously important is the traditional metrological knowledge¹², constantly modernised and extended.

Accurate measurements require relatively frequent calibration of measurement systems with the use of etalons¹³. Calibration¹⁴ may moreover constitute a prospective method of assessing the capacity of a measurement process¹⁵ while finding its main use in legal metrology¹⁶, in the calibration of measuring instruments¹⁷, for applications in responsible measurements¹⁸, etc.

We distinguish several types of methods using etalons: comparisons performed with a single standard, comparisons with two standards, following the calibration curve method, etc. [7].

Methods utilising standard reference materials are not always profitable¹⁹ or applicable, as it is the case with retrospective methods²⁰. A post factum check-up of the correctness of operation of a measurement system by comparing with standards should not be used in making assumptions on the previous operation of the system.

Calibration process is carried out with the use of etalons of particular physiochemical quantity X characterised by an assigned and recognised²¹ value for that physiochemical quantity established with an acceptable uncertainty for concrete applications²² – the correct value $\hat{\mu}$.

Calibration on the basis of etalons of correct values $\hat{\mu}$ results in minimisation of systematic errors ε_s , thus allowing to maintain²³ quality requirements *CTQ*²⁴ for the measurement system. *Gage R&R* methodologies, which do not use etalons, cannot afford the same possibilities of assessment of various sources of systematic errors.

Preparation of measurement systems without the use of etalons guarantees only a reduction in the deviation connected with reproducibility and repeatability of measurements – fig. 1. Deviations connected with scale linearity²⁵ or stability in time²⁶ are determined via metrological tests with the use of etalons and, most commonly, are to be found in the examination certificate of metrological measurement system.

Their determination is usually possible only with the use of etalons. Compensation of various types of deviations is most commonly established by the use of good laboratory practice of *SOP*, which may be found in manufacturers' recommendations, metrological manuals or scientific studies [8,7].

In the assessment of a measurement system we need to consider the reproducibility and repeatability connected with operation in the case of various allowable modifications in the measurement system²⁷, including changes in operational modes – fig. 1.

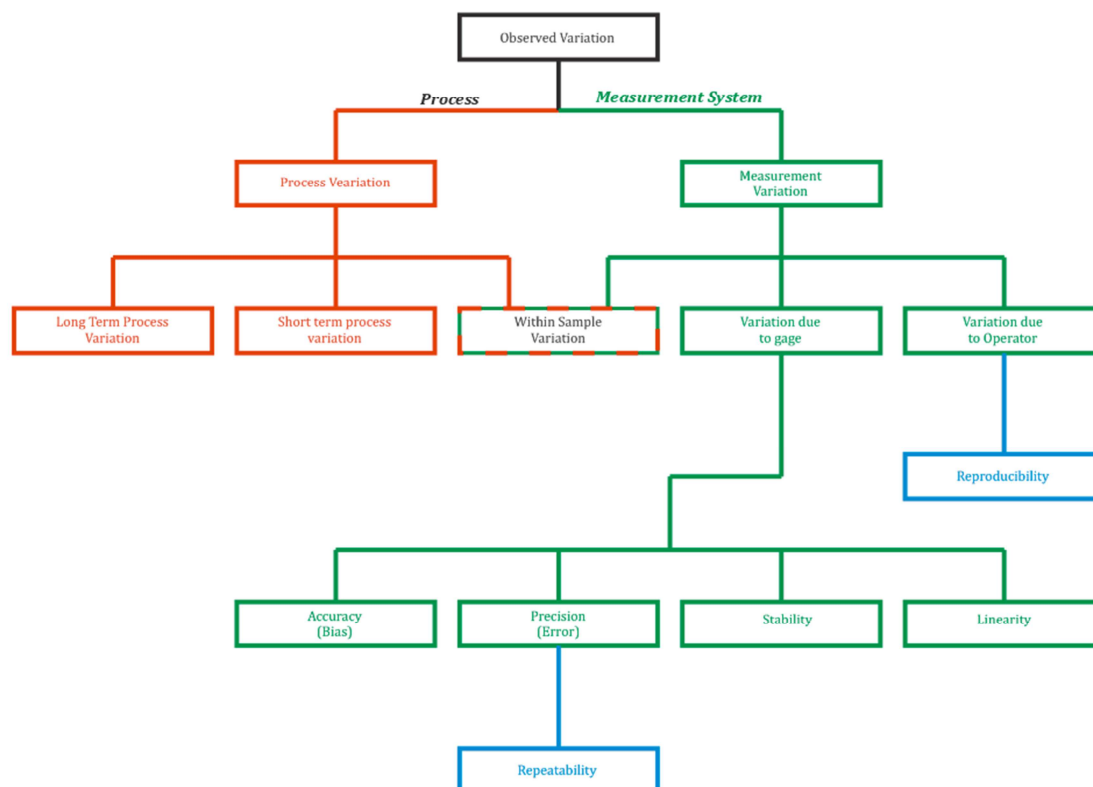


Fig. 1 . Types of the observed process variability and overlapping variability of the measurement system itself [9].

Reproducibility and repeatability

Measurement system reliability assessment, both with and without the use of etalons²⁸ is possible on the basis of an assessment of uncertainty connected with reproducibility²⁹ and repeatability³⁰ of measurements. Reproducibility and repeatability may be estimated by conducting an analysis of historical measurement data³¹ – the retrospective method.

In such a case *reproducibility* characterises the potential possibilities of a measurement system, whereas *repeatability* characterises the practically achievable process capacity with the use of a particular instrument, standard SOP and GLP³².

Variability connected with *reproducibility* is observed in the measurements of the same sample performed by different analysts or/and modifications of the measurement system. *Repeatability*, on the other hand, constitutes an observed changeability in the measurement results of the same sample, with measurements performed by the same analyst, and with the use of the same measurement instrument and at not too long time intervals³³ – fig. 1.

Methods of assessment of a measurement system

The need for development of methods of reliability assessment³⁴ of a measurement process is mainly called for by large-scale and mass production industry [3]. The most important methods include R&R methodologies of process capacity assessment³⁵ [1].

Methods used in metrology relate to the use of reference of measurement results to a chain of etalons of correct values. Thanks to this, with their help it is possible to estimate propagation of nearly all constituents of residual systematic uncertainty³⁶ expressed by systematic errors³⁷. Systematic error ε_s constitutes a difference between the defined value of the measurand $X \leftarrow x$ and the adopted for comparison correct value³⁸ $\hat{\mu}$, most commonly expressed as a difference between the mean value from the conducted measurements \bar{x} and the correct value $\hat{\mu}$: $\varepsilon_s = \bar{x} - \hat{\mu}$.

Often, *reproducibility* is identified with systematic uncertainty. It is required to distinguish between repeatability determined for an etalon of the correct value $\hat{\mu}$, which should not be identified with reproducibility, connected with l –fold measurement repetition $X \leftarrow x_l$ on the same sample X with unknown or only roughly estimated correct value $\sim \hat{\mu}$.

The use of mathematical statistics methods allowed working out of evaluation methods³⁹ for measurement systems on the basis of measurement of the same samples originating from a stabilised or low frequency process⁴⁰ by estimation and separation of variability of a measurement process and measurand⁴¹ variability⁴² and thus determining reproducibility and repeatability of a measurement system.

R&R analysis is a simplified measurement system assessment method as compared with methods used in metrology⁴³ [6,7,1]. With its use it is possible to evaluate a measurement system in the absence of standards or impossibility of their application [7]. However, such an evaluation no longer allows estimation of all sources of residual systematic errors ε_s in a measurement system – tab. 1 [6].

Tab. 1

Suggested method used in measurement system reliability assessment depending on measurement type and availability of a standard correct value $\hat{\mu}$, on the basis of ⁴⁴ [6].

Access to standard correct value	Type of measurement	
	Destructive	Non-destructive
Available	Comparison with standard correct value	Comparison with standard correct value
Unavailable	<i>crossed R&R*</i> method)	<i>nested R&R**</i> method)

*)method does not allow to estimate all constituents of a systematic error of the measurement system ε_s and only enables estimation of uncertainty related to reproducibility and repeatability
 **)method does not allow to estimate variability related to reproducibility and repeatability and only enables estimation of measurement uncertainty as a whole, for which reason it could be formally omitted

DISCUSSION

General assessment of measurement system reliability

General approach to the assessment of measurement system reliability with the use of R&R procedures with and without etalons consists in determining the proportion of the

$$PTR = k \cdot \frac{s}{USL - LSL} < 0,1 \quad \left| \quad s = \sqrt{\frac{\sum_i (\bar{x}_i - \hat{\mu}_i)^2}{n - 1}} \right. \quad (1)$$

obtained measurement precision to the scope of tolerance TR ⁴⁵, defined as [5]:

where: PTR – the ratio of the obtained measurement precision to tolerance representing quality requirements CTQ for the measurement system, k – coverage factor, most commonly two values are adopted $k \in \{5, 15; 6\}$, s – standard deviation from correct value, USL and LSL – respectively, the upper and lower specification limit consisting quality requirement CTQ for the measurement system, i – number of standards, \bar{x}_i – mean value from l measurements for i -th standard, $\hat{\mu}_i$ – correct value for i -th standard, n – total number of replications consisting of all combinations of the number of system modifications m , the number of standards i and replications l for each standard i : $n = \sum_i l_i \cdot m_i$.

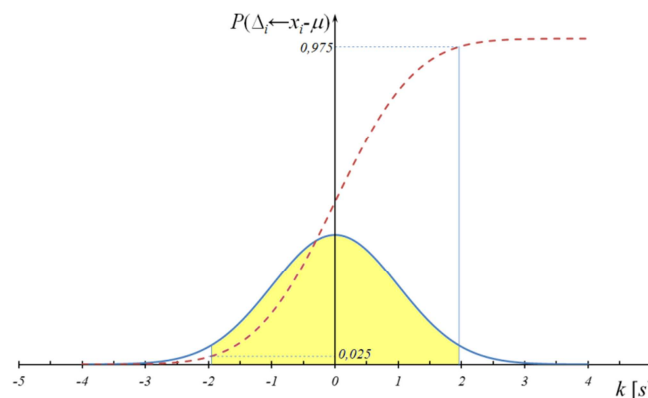


Fig. 2. Theoretical coverage interval for deviations $\Delta_i = \bar{x}_i - \mu_i$ in the mean \bar{x}_i from $n_i = l_i \cdot m_i$ from replications connected with repetitions l and system modifications m in measurement implementation for the standard i at the technical accuracy level corresponding to probability at the level of ca. $P(x_i - \mu) \cong 95\%$ of overall probability constituting a symmetrical interval for values $k \in (-1.96; 1.96)s$.

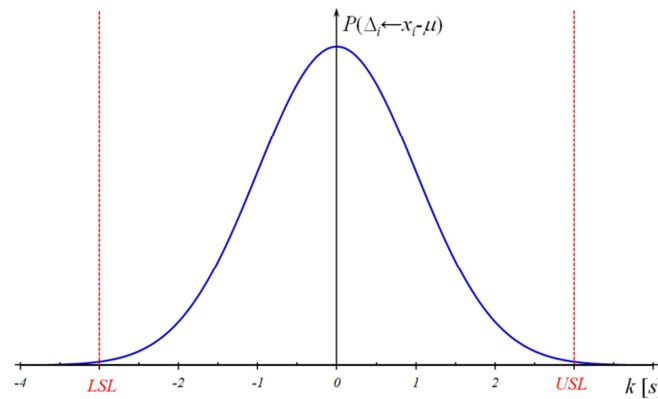


Fig. 3. Theoretical coverage interval for deviations $\Delta_i = \bar{x}_i - \mu_i$ in the mean \bar{x}_i from $n_i = l_i \cdot m_i$ replications connected with repetitions l and system modifications m in measurement implementation for the standard i as measurement coverage interval for sextuple value $k = 6 \cdot s$ of the observed standard deviation s constituting ca. 99.73% of the probability of variability occurrence Δ_i : $P(x_i - \mu) \cong 99.73\%$.

Coverage factor k is a number expressed in the units of the observed standard deviation $[k] = s$ related to expanded uncertainty u estimated as: $u = k \cdot U$, where U stands for standard uncertainty estimated on the basis of variance of measurement results obtained by a measurement system. Expanded uncertainty defines the confidence interval $\Delta(P) \equiv \pm u$ at the level of confidence P for measurement results. The thus constructed confidence interval is a typical method of estimating uncertainty for a replicated measurement [1].

In measuring reference values reproduced by etalon i the deviations $\Delta_i(P)$ are estimated from observable deviations in mean values \bar{x}_i from measurements⁴⁶ x_i from the correct value $\hat{\mu}_i$ reproduced by etalon. If the variability resulting from systematic errors ε_s is significantly smaller $\varepsilon_s \ll \varepsilon_r$ than the variability due to random errors ε_r , then the deviations $\Delta_i = \bar{x}_i - \mu_i$ in the mean \bar{x}_i z $n_i = l_i \cdot m_i$ from replications connected with repetitions l and system modifications m in measurement implementation $X \leftarrow x_i$ for the standard i from its correct value $\hat{\mu}_i$ should have a normal distribution, since by a significant majority they represent variability connected only with random error propagation ε_r . In order to ensure central value of this interval at the point $k = 0$ the variabilities resulting from systematic

errors ε_s and having a constant impact causing a shift of the centre from value $k = 0$ in the form of an adjustment – see fig. 6.

Coverage factor k is selected accordingly to the preferred method. In metrology, the field representing probability of occurrence of a variance $x_i - \mu$ constituting $P(x_i - \mu) \cong 99\%$ of overall field below the Gauss curve representing a certain event, is commonly adopted as satisfactory for the class of laboratory measurements. In such an approach, the coverage factor is at the level of $k \cong 5.15 \cdot s$. For technical measurements, confidence interval is adopted at the level of $P(x_i - \mu) \cong 95\%$ with coverage factor k at the level of $k \cong 3.92 \cdot s$ – fig. 2. Factor $k = 6 \cdot s$ is compliant with the approach referred to as *Six - Sigma*[10] – fig. 3.

It is possible to assume that the thus construed confidence intervals, sometimes called coverage intervals of measurement capacity of the measurement system, constitute significantly weak reliability for such a system. The greater the interval around the observed divergence in the measurement method the easier it is to differentiate and classify the measured properties with its use.

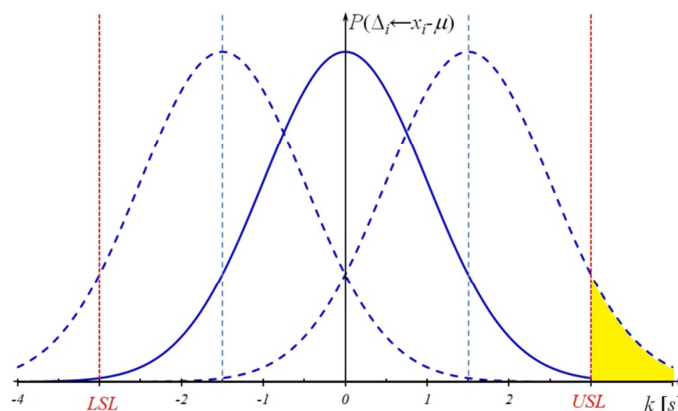


Fig. 4. Illustration of probability connected with surpassing of upper tolerance limit USL caused by the adopted maximum fluctuations of a systematic error at the level of $\varepsilon_s^{max} = \pm 1.5 \cdot \sigma$

Experience has shown, however, that in stabilised processes it is still possible to observe fluctuations connected with the occurrence of systematic errors ε_s causing maximum shift in the reliability of measurement results from the mean \bar{x} by the adopted maximum value of $\varepsilon_s^{max} = 1.5 \cdot s$. With such an

assumption, the mean value may deviate by the value ε_s^{max} towards both directions $\bar{x} \pm 1.5 \cdot s$. Leaving tolerance limits $[LSL; USL]$ at the same level as shown in fig. 3 will cause that with the maximum value of a systematic error ε_s^{max} in the part being beyond one of the tolerance limits $[LSL; USL]$ probability will be

reduced by: $P(x_i - \mu) \cong 6.67\%$ – fig. 4

In experimental determination of the ratio of the obtained measurement precision to tolerance PTR it is needed to head towards an even coverage with standards i of the required measurement scope⁴⁷ $[USL; LSL]$. Measurement of each standard should be realised through its multiple replication $l > 5$, with several possible modifications⁴⁸ in the measurement system $m > 2$, including work of different operators. The thus designed plan of action should be randomised⁴⁹.

Should a measurement system not require operation, it is necessary to group measurements and implement them by applying significant time intervals between the groups in allowable, yet different conditions, with various charging levels of the integral battery pack, with the use of different available software versions, allowable system modifications, different times of day,

locations⁵⁰, etc., also with consideration of randomisation process for such modifications.

Relation [1] reflects the generally adopted assumption that the measurement system should enable differentiation of at least 10 different values for the predicted scope of the expected measurand variability, as one of the quality requirements CTQ regarding the measurement process.

Fig. 5 presents an example assessment of the value of ratio PTR of the reached measurement precision to tolerance in measurement performance for two standards and application of *Six – Sigma* approach.

The observed etalon divergence was used to define the coverage factor for tolerance interval $[LSL; USL]$.

Nearly half of the field below the Gauss curve

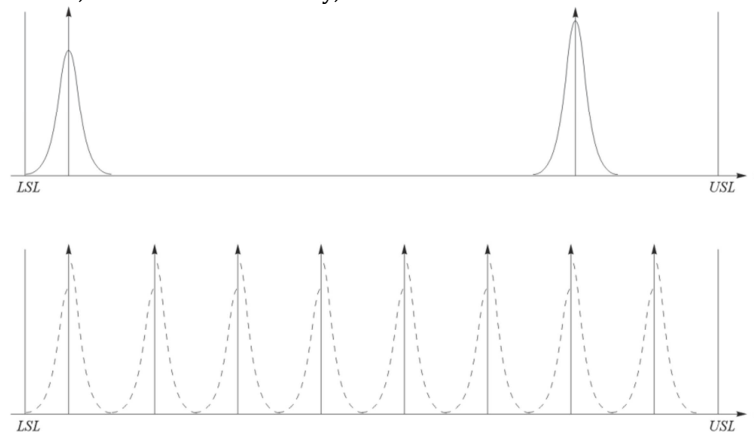


Fig. 5. Representation of an assessment of precision to tolerance ratio PTR for the measurement system (bottom figure) with coverage factor $k = 6 \cdot s$ and the tolerance scope $[LSL; USL]$ for which the measurements were performed, with two standards (top figure). However, the conducted calculations show only approximately octuple coverage of the tolerance interval $[LSL; USL]$, hence the system failed to meet the requirements of relation [1].

is found within the distance from the centre of ca. $k \cong 3 \cdot s$. Therefore, the mean for the two standards will be depicted by combination of the halves of distribution curves, one for each standard respectively. The thus obtained mean precision interval of a measuring method constituted the measure to determine the scope of tolerance. The calculations carried out on fig. 5 show that the tolerance interval was covered over eight times. In concord with relation [1] however, the system does not meet the adopted quality requirements CTQ for the adopted measurand variability range.

Relation [1] depicts the familiar principles of mnemotechnics stipulating that the measuring precision should be at least by an order of magnitude larger than the anticipated measurand variability range for which the measurements need to be conducted. Hence, the value of the ratio of the reached measurement precision to tolerance PTR from relation [1] should be lower than $PTR < 0.1$.

Measurement system assessment based on a standard

Metrology commonly utilises estimation of an absolute error Δ in the measurement result in the form of an arithmetic mean \bar{x} constituting the estimator⁵¹ of the true value μ from the performed n measurements: $\Delta = k \cdot \frac{s}{\sqrt{n}} = k \cdot u \Rightarrow \frac{\Delta}{u} = k$, where s is a standard deviation calculated on the basis of the results obtained in the measuring cycle, whereas u is a standard error. The choice of values k for technical measurements is

illustrated in fig. 2.

Generally, deviations $\Delta = \bar{x} - \mu$ in the mean \bar{x} from l measurement replications $X \leftarrow x_l$ from true value μ should have a normal distribution, as they should represent only random errors. Coverage factor k is selected on the basis of this distribution. However, often what is available are only results obtained for a small sample $n < 30$, and in such a case the value of the coverage factor k selected on the basis of t –Student distribution. For sample multiplicity $n \geq 30$ it is assumed that differences between the t –Student and normal distribution may be omitted in the majority of problem situations with factor k adopted on the basis of normal distribution. In technical measurements and with normal distribution the most commonly adopted value of the coverage factor of $k = 2$.

There is the mnemotechnic principle stipulating that the result⁵² \bar{x} of a measurement cycle⁵³ should be at least by an order of magnitude larger than its uncertainty expressed with the use of absolute error Δ : $\bar{x} \gg \Delta$.

$\bar{x} \gg 10$.

By the assumption that a measurement is a value of weighted absolute error $\bar{\Delta}$ from measurements i standards calculated as mean $\bar{\Delta} = \frac{\Delta}{n}$ from the values of absolute differences $\Delta_i = |\bar{x}_i - \hat{\mu}_i|$ in the mean \bar{x}_i of l measurement replications $X \leftarrow x_n$ for all modifications in the measurement system n , then analogously to the above representation, the absolute error for this value $\bar{\Delta}$ will be the product $k \cdot u$, where u is standard uncertainty $u = \frac{1}{\sqrt{n}} \cdot s$.

Since it was previously mentioned that the reasonable minimum value of the coverage factor k is $k = 2$, then going back to the principle stating that the measured value should be by an order of magnitude greater than the absolute error for this value, we may write down the following: $\forall n \geq 30 \Rightarrow k=2 \quad \frac{\bar{\Delta}}{k \cdot u} \gg 10 \Rightarrow \frac{\bar{\Delta}}{u} \gg 5$. This allows to estimate the value of measurement

system reliability assessment index for the method based on the use of etalons, as a parameter constituting the ratio of mean absolute error $\bar{\Delta}$ to the estimated mean value of its standard deviation \bar{s} : $\frac{\bar{\Delta}}{\bar{s}}$ [6].

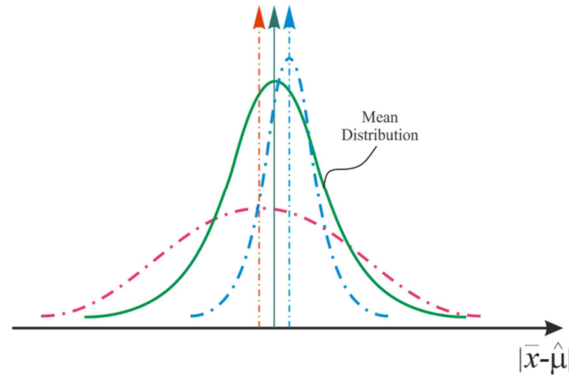


Fig. 6. Deviations $|x_i - \hat{\mu}_i|$ for $i = 2$ standards of correct values $\hat{\mu}_i$ and the distribution of their mean value $\frac{1}{n} \cdot \sum_i |x_i - \hat{\mu}_i|$.

$$\frac{\bar{\Delta}}{u} \gg 5 \quad \left| \quad \bar{\Delta} = \frac{\sum_i |x_i - \hat{\mu}_i|}{n}; u = \frac{1}{\sqrt{n}} \cdot s = \frac{1}{\sqrt{n}} \cdot \sqrt{\frac{\sum_i (x_i - \hat{\mu}_i)^2}{n-1}} \right. \quad (2)$$

where: $\bar{\Delta}$ – standard deviation from correct values $\hat{\mu}_i$ for i standards; u – standard uncertainty calculated for all standards i .

In its nature, principle [2] constitutes a variance analysis carried out in the same manner as estimation of measurement uncertainty, where the measurement results are absolute values of the deviation $|x_{n,i} - \hat{\mu}_i|$ of measuranda $\hat{\mu}_i$.

Averaging of deviations is justified in this case, as the variables accumulate only the allowable differences in the modifications of the measurement which may be neglected from the point of view of quality requirements *CTQ* for the measurement process – fig. 6. Such choice of variables should in the majority of cases ensure their normal distribution. If in the course of an obligatory check-up of deviation distribution $|x_{n,i} - \hat{\mu}_i|$ it proves to be disturbed, we will be bound to suspect presence of an excessive number of systematic errors⁵⁴.

The deterministic disturbance caused by presence of systematic errors should be first identified and minimised before carrying out another analysis on the reliability of a measurement system. Seeking the sources of systematic errors is only possible with etalons reproducing correct values $\hat{\mu}_i$ for the measurands.

CONCLUSIONS

Relation [1] suggests the capacity of a measurement process to cover the scope of the expected measurand variations, whereas inequality [2] implies fulfilment of quality requirements of each measurement. For a measurement system which meets the quality requirements *CTQ*, conditions [1] and [2] should be fulfilled simultaneously.

Fulfilment of only one condition [1] expressed through precision to tolerance ratio *PTR*, which may fall below the value $PTR < 0.1$ and thus meet the quality requirements *CTQ* concerned with the possibility of distinguishing of the measured value is insufficient.

Along with the fulfilment of condition [1], the value of ratio [2] of mean absolute error $\bar{\Delta}$ to its standard uncertainty u may fall below the value $\frac{\bar{\Delta}}{u} < 5$ and the system will not guarantee the required precision with approved reliability. Similarly, meeting of condition [2] will ensure measurement precision, however at the same the measurement process may not ensure coverage of the required interval and thus fail to fulfil condition⁵⁵ [1].

Information concerning reliability of a measurement system should be provided by manufacturer after filing a request for proper declaration. Often, values declared by the manufacturer may be improved by applying proper operational procedures *SOP*, as usually manufacturers provide in their declarations average values selected according to the worst case analysis method. In the determination of true metrological values of a measurement system it is necessary to carry out metrological tests.

These tend to be costly, not only in terms of expenditures related to carrying them out, but in majority of cases also require covering investment costs by purchasing a measurement system. Companies rarely agree to lend measurement systems for the purpose of conducting metrological tests. For this reason, this stage of works is often omitted, thus resulting in an expensive disappointment during the operation of a measurement system.

The proposed indicators [1]-[2] relating to the fulfilment of quality requirements *CTQ* of a measurement system are useful, albeit their application is rather conceptually complicated. A natural, self-explanatory method in the assessment of measurement systems consists in the establishment of confidence intervals for the fractions of observed divergence[4] estimated with variance analysis. Nonetheless, this approach is rarely used in the analysis of measurement systems or statistical inference [11].

This is probably due to the necessity to consciously adopt the level of the power of inference involving the awareness of relatively large, impossible to accept values of errors of the second type for small

samples. Should metrological tests indicate an improvable value of errors of the second type, they need to be repeated for a greater population, thus significantly increasing the costs and exposing researchers to their own dissatisfaction, as well as that of sponsors, management, etc.

Often, due to lack of knowledge or for the sake of peace they are altogether omitted or undisclosed by researchers. In designing life support system such cynical omissions or lack of sufficient knowledge may have very dramatic repercussions.

The publication presents effects of the research carried out by the Naval Academy of Gdynia, financed from the educational fund for the period of 2013–2015 within the developmental project entitled "Designing decompression in MCM dives", contract no. DOBR/0047/R/ID1/2012/03.

BIBLIOGRAPHY

1. Allen T.T. 2006. Introduction engineering statistics and Six Sigma. London : Springer Ltd., 2006.
2. Automotive Industry Task Force. 2010. Measurement systems analysis Reference Manual. Fourth Edition. No place indication : Chrysler Group LLC, Ford Motor Company, General Motors Corporation, 2010. ISBN 978-1-60-534211-5.
3. Burdick R. K., Graybill F. A. 1992. Confidence Intervals on Variance Components. New York : Marcel Dekker, 1992. ISBN 0-8247-8644-0.
4. Burdick R.K., Borror C.M., Montgomery D.C. 2005. Design and Analysis of Gauge R&R Studies: Making Decisions with Confidence Intervals in Random and Mixed ANOVA Models. Philadelphia : Society for Industrial and Applied Mathematics, 2005. ISBN 0-89871-588-1.
5. GUM. 1984. Międzynarodowy słownik podstawowych i ogólnych terminów metrologicznych. GUM 1996. Geneva : NIPM, IEC, ISO, OIMC, 1984. translation: International vocabulary of basic and general terms in metrology.
6. Jakus B., Kłos R. 2007. Measurement process capacity. Polish Hyperbaric Research. 2007, Vol. 19, ISSN 1734-7009, pp. 33-46.
7. Kłos R. 1990. Method of determination of the composition of breathing mixes in diving hyperbaric complexes. Gdynia : Polish Naval Academy, 1990. Doctoral dissertation.
8. —. 2007. Application of statistical methods in the diving technique - Handbook. Gdynia : Polish Society of Hyperbaric Medicine and Technique, 2007. ISBN 978-83-924989-26.
9. Montgomery D.C. 2009. Statistical quality control. 6th Edition. New York : John Wiley & Sons Inc., 2009. ISBN 978-0-470-16992-6.
10. Collective work. 1990. Metrology textbook. [ed.] P.H. Sydenham. Warsaw : Wydawnictwo Komunikacji i Łączności, 1990. Vol. I and II ISBN 83-206-0681-0; ISBN 83-206-0812-0.
11. Pyzdek T. 2003. The Six Sigma handbook. New York : The McGraw-Hill Companies, Inc., 2003. ISBN 0-07-141596-3; DOI: 10.1036/0071415963.
12. Riding K. 2001. The Book of Knowledge. no place : GE Power Systems University, 2001. training materials. ver. 1.3.

dr hab. inż. Ryszard Kłos, prof. nadzw.

AMW

Akademia Marynarki Wojennej im.
Bohaterów Westerplatte
Zakład Technologii Prac Podwodnych
81 – 103 Gdynia 3, ul. Śmidowicza 69
Tel.: +58 626 27 46, Fax.: +58 626 27 61

¹previous article presented a practical application of *R&R* analysis in the evaluation of newly implemented analysers, whereas here I offer a general theoretical introduction,

²gage repeatability and reproducibility,

³i.e. random; this term is used for the purpose of differentiation from the term 'statistical' which is also used outside the mathematical statistics,

⁴critical to quality,

⁵systematic errors are minimised to the accepted level by introducing alterations or using special measurement methods,

⁶if in a given problem situation it is acceptable to assume accuracy at a determined level from the point of view of purposefulness of undertaken action,

⁷acceptable in particular conditions,

⁸correct value is an estimator of true value,

⁹Current understanding of a measurement system points to a set providing in a simple way easily interpretable results, without complicated operating procedures requiring presence of highly qualified operators. In the past, a measuring instrument was equipment dedicated to special types of measurements. At present, due to substantial progress in the available methods and electronisation of measurements it is difficult to encounter a measuring device that would not constitute a complicated electro-mechanical system,

¹⁰measuring instrument is a device designed to perform independent measurements or in combination with one or many auxiliary devices [5],

¹¹Standard Operating Procedure – *SOP*,

¹²the achievements of metrology are still an indispensable element in the knowledge of an analyst of measurement systems – quite often negligence of this fact may lead to serious problems,

¹³a standard unit of measurement (etalon) is a measuring instrument, reference material or measurement system designed to define, realise, preserve or reproduce units of measurement or one of many values of a certain quantity and acting as a reference[5],

¹⁴calibration is a set of defined operations carried out in defined conditions, relationship between the values of a measured quantity indicated by a measuring instrument or measurement system or values represented by standard measurement or reference material and proper values of the quantity realised by standard units of measurement [5],

¹⁵the capacity of a measurement process should be understood as the level of meeting quality requirements *CTQ*,

¹⁶For instance:

Type test constituting a set of activities aimed at determining whether a measuring instrument of a certain type meets the requirements for approval in accordance with the act on measurement units.

Legalisation comprises a number of activities including check-up, confirmation and certification with a proof of legalisation that a measuring instrument meets the requirements stipulated in the act on measurement units. Authentication resulting from calibration which may be confirmed with a document known as a calibration certificate or protocol,

¹⁷Universal measurement systems require adjustment and calibration to the conditions of measurement performance, as depending on the choice of working parameters they may be utilised for different types of measurements, e.g. a gas chromatograph, depending on column selection, its working parameters, detector type and *SOP* may be used in qualitative or quantitative measurements of various liquid and gaseous mixes.,

¹⁸sometimes it is necessary to aim at precise estimation of the current reliability of measurements, e.g. in security technique, health protection, working conditions, legal metrology, etc.,

¹⁹sometimes its use is impossible not to due technical reasons but rather due to absence of an economic justification,

²⁰sometimes it is difficult to apply calibration on an operating production line as it may require interrupting a production process, shutting off or performing a switch-over to a parallel instrument, which quite often is an element of an automation system – of course properly designed measurement systems need to allow calibration, however it may be unjustified or problematic at a given time,

²¹for instance, on the basis of a certificate issued by a recognised company,

²²the best, commonly available in the market,

²³Checking,

²⁴systematic errors have been compensated to such a degree that they are an order of magnitude smaller than random errors, gross errors are subject of statistical analysis of outlying points,

²⁵relation of the response of a measurement system O in the function of the measurand, e.g. the standard W : $O = f(W)$,

²⁶stability of the response O of a measurement system in the function of time τ : $O = f(\tau)$,

²⁷for instance, it may be possible to use two types of software to operate the measurement system,

²⁸however in such a case it will be impossible to estimate all of the constituents of a systematic error in the measurement method,

²⁹reproducibility means the level of compliance of measurement results of the same measured value, performed in varying conditions [5],

³⁰repeatability means the level of compliance of measurement results of the same measured value, performed in the same conditions [5],

³¹the choice of the time segment for performing such estimates is crucial as it has a significant effect on the values of reproducibility and repeatability – a shorter time usually means better values of the evaluated parameters

³²Good Laboratory Practice,

³³repeating of measurement,

³⁴also frequently applied is determination of the capability of a measurement process which is concerned with the confirmation of fulfilment of quality requirements *CTQ* of a measurement process for particular applications, whereas reliability assessment underpins the formal confirmation of such capacity,

³⁵reliability and reproducibility,

- ³⁶traditionally, and correctly, it is taught that systematic errors should be minimised to an acceptable level, thus perhaps implying the possibility of their complete elimination, whereas in fact in many cases it is impossible to achieve[7,8],
- ³⁷systematic error constitutes the absolute value of deviation, whereas uncertainty is understood as an extension of this value with the boundaries of accuracy of its determination,
- ³⁸sometimes a greater value of a systematic error is adopted taking into account also different modification in the measurement system, such as: different software versions, adjustment methods, operational performance, etc., most commonly in the form of manufacturer's declaration regarding maximum systematic error for a given measurement system,
- ³⁹estimation of the level of fulfilment of CTQ quality requirements by a measurement system,
- ⁴⁰e.g. a production process,
- ⁴¹natural fluctuations in the measurand,
- ⁴²measured physical quantity, property, state, etc.,
- ⁴³with the use of a standard correct value,
- ⁴⁴lack of complete compliance with the quoted source,
- ⁴⁵also called *PTR* – precision to tolerance ratio,
- ⁴⁶being a realisation of random variable $X \leftarrow x_i$,
- ⁴⁷*USL*– Upper Specification Limit; *LSL*– Lower Specification Limit,
- ⁴⁸Settings,
- ⁴⁹random order of realisation of various sample combinations and measurement system settings,
- ⁵⁰for a mobile device,
- ⁵¹function used to estimate parameter distribution value,
- ⁵²usually provided in the form of arithmetic mean \bar{x} ,
- ⁵³understood as a result of a cycle of research activities leading to the estimation of the final result in the form of research conclusions,
- ⁵⁴presence of a deterministic factor disturbing analysis,
- ⁵⁵frequently the scope of a given measurement system fails to cover the required.