MATEC Web of Conferences **94**, 04012 (2017) DOI: 10.1051/matecconf/20179404012
CoSME'16

Comparison of the methods for determination of calibration and verification intervals of measuring devices

Pavlina Toteva[1,*], *Stoyan* Slavov[1], and *Dimka* Vasileva[1]

[1]Technical University of Varna, Department of Technology of Machine Tools and Manufacturing, 9000 Varna, 1 Studentska str., Bulgaria

Abstract. The paper presents different determination and optimisation methods for verification intervals of technical devices for monitoring and measurement based on the requirements of some widely used international standards, e.g. ISO 9001, ISO/IEC 17020, ISO/IEC 17025 etc., maintained by various organizations implementing measuring devices in practice. Comparative analysis of the reviewed methods is conducted in terms of opportunities for assessing the adequacy of interval(s) for calibration of measuring devices and their optimisation accepted by an organization – an extension or reduction depending on the obtained results. The advantages and disadvantages of the reviewed methods are discussed, and recommendations for their applicability are provided.

1 Introduction

Many organizations use technical devices for monitoring and measurement (TDMM) of processes, products and services. This includes most of the enterprises involved in processing and manufacturing industry, as well as testing laboratories and authorities for technical control. In order to prove a result of measurement reliable, it is necessary to use only a calibrated TDMM. This is a basic prerequisite in widely applied standards with requirements for management systems, such as ISO 9001, ISO/IEC17020, ISO/IEC17025, ISO14001, OHSAS 18001 (ISO 45001) in terms of ensuring adequate management and maintenance of TDMMs [1-6].

According to the Legal Metrology in the Republic of Bulgaria, to ensure the accuracy and reliability of measurements in health care service and public safety measurements, environmental protection, state and municipal receivables and trade payments metrological control should also be carried out. For this purpose the State Agency for Metrology and Technical Surveillance has been founded to regulate the frequency of verification of measuring instruments in use that are subject to metrological control.

The accuracy of the technical devices used in each accredited laboratory is regulated in clause 5.5.2 of ISO/IEC 17025: 2006, namely: *"Technical devices and related software used*

* Corresponding author: pavlina_toteva@tu-varna.bg

for testing, calibrations and sampling must provide the accuracy required and must be in compliance with the requirements for these tests and/or calibrations."

In organizations, technical devices for monitoring and measurement used to check the required conformity of products and services must provide valid and reliable results [1].

Calibration of measuring instruments should be carried out at appropriate intervals, the calibration periods depending on different factors, such as required uncertainty, frequency of use, method of use, stability of measuring instruments, the specific guidance of the manufacturer of TDMM, etc. [8].

Before being put into operation a technical device should be calibrated or checked in order to be determined whether it satisfies the requirements set out by the laboratory and whether it is in accordance with the relevant essential requirements [4].

The periods of verification of measuring instruments are determined in the Ordinance on measuring instruments subject to metrological control [7] and by order of the Chairman of the State Agency for Metrology and Technical Surveillance. For other types of TDMMs Bulgarian legislation does not specify periods for calibration. The organizations using TDMMs are responsible for determining the time intervals themselves to ensure reliable results. Moreover, it requires the presence of documented methodology by which the organization determines the periods for calibration [1, 5] relevant for the applied TDMM.

This leads to difficulties in small and middle-size companies because of the use of a large number of TDMMs undergoing calibration. Furthermore, it should be taken into account that TDMMs are usually obtained from different manufacturers (suppliers); they are introduced in operation at different time points, so frequency and duration of use in process monitoring and control is not identical. This often requires an individual approach to the metrological maintenance of each TDMM (or a group of TDMMs) which further complicates the task of implementing appropriate methodology that is adequate for all TDMMs used in the organization.

Inadequate definition of calibration intervals carries risks for the organization in two ways: too short an interval results in higher costs for calibration of the respective TDMM for the organization. If a longer than necessary period is considered, one faces the risk of getting unreliable results using TDMMs in processes for monitoring and/or production evaluation. Consequently, the organization can also suffer significant losses limited not only to the financial aspect.

For these reasons, the issue of choosing one or another method for determining the periods to calibrate TDMMs is not unequivocal and it is relevant for all organizations. Therefore, the paper presents an overview of the most common methods as well as their advantages and disadvantages and their practical applications.

2 Existing methods for determining the intervals for the calibration of TDMMs

It is well known that in meteorology it is practically impossible to predict the exact time point at which the indications of a TDMM will go beyond the margin of error of measurement. Therefore, multiple methods for determining the calibration interval of a TDMM are developed aiming at the guarantee that the proportion of measurements that fall outside the margin of error remains as small as possible. The methods for determining calibration intervals are introduced by the international organizations OIMLH and ILAC [10, 11].

2.1 Methods for determining the intervals for the calibration of TDMMs recommended by International Organization of Legal Metrology (OIML) and International Laboratory Accreditation Cooperation (ILAC)

OIML and ILAC offer general guidelines for the criteria and methods for determination of calibration intervals [10]. In general, the methods for determining the time interval between calibrations of TDMMs can be divided into two main groups - statistical and algorithmic techniques. Statistical methods are based on the statistical modelling of the relationship between the probabilities of indications of a TDMM to be in the tolerance of the metrological characteristics for a certain period after the last calibration. It aims at predicting the duration of the interval to the next calibration based on the probability of reaching a percentage of measurements within acceptable limits at the end of this period. They usually require a significant amount of data from previous calibrations and they are relatively challenging in terms of their practical application (see Table 1).

Table 1. Comparison of methods for determining ranges of calibration of TDMMs [12].

Main indicators	**1st method – Automatic adjustment or "staircase" (based on calendar time)**	**2nd method - Checklist (based on calendar time)**	**3rd method – Duration of measurement device usage**	**4th method - Control during operation or testing with "black box"**	**5th method - Statistical approaches**
Reliability	medium	high	medium	high	medium
Difficulty of application	low	high	medium	high	high
Balanced load	medium	medium	low	medium	bad
Applicability in specific measurement devices	medium	low	high	high	low
Availability of measurement devices	medium	medium	medium	high	medium

The second group of methods is based on both relatively simple and complex logical algorithms for determining the calibration interval depending on the results and the conditions of TDMMs during calibration. Usually these approaches contain logically reasoned "decision" to extend or shorten the calibration interval depending on observations of current or previous calibrations. Due to their nature, these methods are called algorithmic methods [13]. Currently, algorithmic methods are relatively widely used in practice because of their uncomplicated character and the low cost of implementation. However, they have some shortcomings, some of the most important ones being the following:

-Change in the calibration interval can be induced by even a single evidence of a TDMM outside the margin of error, which can be caused by random factors, i.e. by random fluctuations in readings.

-Algorithmic methods usually do not include the impact of climate uncertainty of measurement of a TDMM in determining the calibration interval.

-Algorithmic methods cannot be easily adapted to the determination of target reliability related to the quality of measurements. The level of reliability achievable with algorithmic methods can only be determined experimentally or by simulation.

-If the prescribed interval is consistent with the desired level of reliability, the results of next calibration or next few calibrations often lead to moving away from the optimal interval.

DOI: 10.1051/matecconf/20179404012

-The time span necessary to achieve reliable determination of time intervals could reach up to 40-50 years.
-Automation of algorithmic methods is usually demanding. Therefore, they are often implemented manually by calibration specialists which increases the time for their realization, leading to increased costs of the process.

2.2 Comparative analysis of the methods for determining and optimizing the intervals for the calibration of TDMMs given in Recommendations on Interstate Standardization (RMG 74-2004), State Measurement System (GSI)

In RMG 74-2004 GSI "Methods for determining the intervals of verification and calibration of measuring devices" [11] six methods for determining and optimizing calibration intervals of TDMMs are described. Table 2 depicts the comparative analysis of the advantages and disadvantages of these methods for the optimization of calibration intervals of TDMMs carried out in different organizations [11]. The results of the analysis indicate that none of the discussed methods provides comprehensive information about the duration of time interval between calibrations of TDMMs. To obtain the most accurate time interval it is necessary to conduct empirical research within the life-cycle of a TDMM. Empirical tests should be carried out in not too long intervals of the operation period as these tests are labour-intensive and they are associated with increased expenditure of time and human resources. Regarding this, it is appropriate to conduct accelerated testing of a TDMM allowing prediction of the importance of starting parameters that determine the performance of a TDMM.

Based on the performed comparison and taking into account the data from the source [14] the authors propose a combined method for statistical check and update (if necessary) of calibration intervals of TDMMs using an algorithmic approach.

Table 2. Methods for determining verification intervals of TDMMs.

Methods for setting the verification intervals	Advantages	Disadvantages
1. Based on statistical implicit and explicit rejections.	High degree of reliability of the test results.	A large amount of experimental data for the process of change over time; the survey is labour-intensive.
2. Economic criteria.	Minimize operating costs of a TDMM; remove the consequences of possible errors caused by measurement errors.	Use of approximate models which relate mostly to one type of a TDMM.
3. Arbitrarily set the initial period of calibration with subsequent adjustments during operation of a TDMM.	Minimal financial and time costs.	Incorrect determination of the initial interval; there are no recommendations for all types of TDMMs about the initial interval; lack of reliability data of TDMMs.
4. Set the intervals for the calibration of a TDMM based on analogy.	Lack of financial and time costs.	The results obtained for analogues are not always applicable to newly developed or purchased TDMMs. There are no data on the reliability of the elements of analogue.

Methods for setting the verification intervals	Advantages	Disadvantages
5. Determination of the intervals for the calibration of a TDMM on indicators of reliability (failure rates or time limit for metrological operation).	High degree of reliability of test results.	A large amount of experimental data for the process of change over time; the survey is labour-intensive.
6. Determination of calibration intervals for analysis of the progressive component of the inaccuracy of a TDMM.	Low financial costs and high reliability of test results.	Need data availability for the longest period of inspection as well as costs for constant monitoring of the need for verification.

3 A combinatorial method for determining the calibration interval of TDMMs

3.1 The essence of the method

The methodology described [13-15], aims at minimizing most of the discussed shortcomings of the algorithmic methods in determining or verifying a calibration interval of a TDMM.

In the analysis of calibration intervals, the term "reliability" of indications of a TDMM is linked to the probability of a parameter measurement of a TDMM within the tolerance range for the measured value. The observed reliability is defined as the proportion of measurements that have been found to be within the tolerances when performing measurements, calibration or other checks.

In this method, the observed reliability is introduced as the variable R_{obs}. The verification of a calibration interval is based on the comparison between the observed values of reliability (R_{obs}) and the specified target reliability of indications of a TDMM (R_{targ}). R_{obs} is defined as the total number of tests (n) carried out during the current calibration against the number of tests (x) where the readings were within tolerance. For example, if n = 24 and x = 19, the value of the observed reliability R_{obs} will be:

$$R_{obs} = \frac{x}{n} = \frac{19}{24} = 0.792 = 79.2\% \tag{1}$$

The current interval is tested by comparing the value obtained for R_{obs} and the target reliability R_{targ} set to verify whether there is a significant difference between these two parameters. The size of difference and confidence levels determines the level of significance. If the evaluation of the observed reliability R_{obs} and the set target reliability R_{targ} reveals a significant difference, the selected current calibration interval of a TDMM is rejected and a shorter or longer interval is recommended (Fig. 1).

The algorithm used for verifying the current calibration interval calculates whether R_{obs} is within the confidence limits set by the value of the target reliability R_{targ}. If so, the current interval is considered to be properly defined and a recommendation for its preservation should be done. Confidence limits for Q are determined based on the function of the binomial distribution. With the pre-set total number of tests (n) and the number of tests within the tolerances of the measured value (x), the upper confidence limit p_U and the lower confidence limit p_L are determined by the dependencies below:

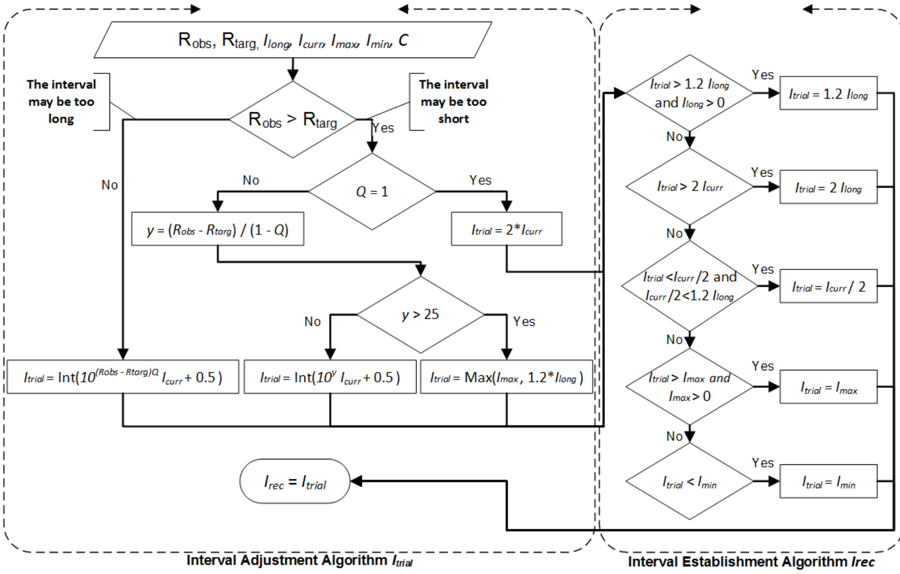


Fig. 1. Algorithm to check and determine the recommended interval I_{rec}.

$$\sum_{k=0}^{x}\binom{n}{k}p_U^k(1-p_U)^{n-k} = \begin{cases} \frac{1-C}{2}, 0<x<n \\ 1, x=n \\ 1-C, x=0 \end{cases} \tag{2}$$

$$\sum_{k=0}^{x}\binom{n}{k}p_L^k(1-p_L)^{n-k} = \begin{cases} \frac{1-C}{2}, 0<x<n \\ 1-C, x=n \\ 0, x=0 \end{cases} \tag{3}$$

where C is the given confidence level of changing the examination interval. If the requirement $p_L \leq R_{targ} \leq p_U$ is fulfilled, the selected current interval satisfies the check and may be kept unchanged. Otherwise, it does not satisfy the verification and should be adjusted (increased or decreased).

3.2 Calculating the confidence level Q for the rejection of the selected interval

The value of Q is determined by a numerical iterative method based on the pre-set confidence level C and the resulting values for p_1 (lower limit) and p_2 (upper limit) upon satisfaction of the following equations:

$$\sum_{k=0}^{x}\binom{n}{k}p_1^k(1-p_1)^{n-k} = \frac{1-C}{2} \tag{4}$$

and

$$\sum_{k=0}^{x}\binom{n}{k}p_2^k(1-p_2)^{n-k} = \frac{1-C}{2} \tag{5}$$

The process of the iterative calculation begins by setting the variable C to a value close to 1.0 (for example, 0.9999999). Iterations continue until the lower or upper limit (p_1 or p_2)

satisfy one of the following:

$$p_1 < R_{targ} \text{ and } R_{targ} - p_1 < \varepsilon, \tag{6}$$

$$p_2 > R_{targ} \text{ and } p_2 - R_{targ} < \varepsilon \tag{7}$$

where ε is the pre-set level of accuracy (ε = 0.00001). If any of the conditions (6) or (7) is satisfied in the course of iterations, the confidence level Q for the rejection of the inspected interval takes the value of C. For the calculation of parameter C, an iteration method of interval bisection is used, the so called "bisection method" [16]. Gradually, approximate values for the variables p_1 and p_2 are set. These are obtained as follows: First, the variables P and y_p are defined by the expression P = (1 - *C*) / 2 and $y_P = \Phi^{-1} \cdot (1 - P)$, where Φ is the inverse function normal distribution. Next, the variable p is calculated by the equation:

$$p \cong \frac{a}{a + b \cdot e^{2 \cdot w}}, \tag{8}$$

where:

$$w = \frac{y_p (h + \lambda)^{1/2}}{h} - \left(\frac{1}{2b-1} - \frac{1}{2a-1} \right) \left(\lambda + \frac{5}{6} - \frac{2}{3h} \right)$$

$$h = 2 \cdot \left(\frac{1}{2b-1} + \frac{1}{2a-1} \right)^{-1} \text{ and } \lambda = \frac{y_P^2 - 3}{6}$$

Once the variable p is determined, the preliminary values p_1 and p_2 are defined as:

$$p_1 = p - dp \text{ and } p_2 = p + dp \tag{9}$$

where *dp* is set conditionally (e.g. *dp* = p / 100). During each iteration step, the values of p_1 and p_2 are recalculated by:

$$I_{P1}(x, n - x + 1) = 1 - C \text{ for } p_1\text{, and} \tag{10}$$

$$I_{P2}(x + 1, n - x) = C \text{ for } p_2. \tag{11}$$

Upon receipt of the recalculated values for p_1 and p_2, they are tested and the iterative process for the calculation of *C* is terminated, (6) or (7) are fulfilled and Q takes the value of C for the current iteration step.

3.3 Study of the influence of "target reliability" (R_{targ}) and "set confidence level" (C) parameters

To examine the results of the method, it is necessary to set specific values of the confidence level C and the target reliability R_{targ}. Upon further consideration it is assumed that methodology is set at nonzero current interval $I_{curr} \neq 0$, and at zero minimum permitted interval $I_{min} = 0$, for the maximum acceptable interval I_{max} and for the longest interval I_{long} the set values being 10 000.

When verifying statistical significance, the level of confidence is generally assumed to be 95%. If the value of the confidence level C is set to be 95% by methodology, the confidence limits p_L and p_U determine 95% confidence level for the parameter Robs, and the tested interval will be rejected with 95% confidence if R_{targ} is out of the calculated tolerance (p_L, p_U).

It is important to take into consideration that the inspected interval will be rejected if R_{obs} significantly differs from R_{targ}, as the level of importance is assigned by the chosen level of confidence (95%). Therefore, if the maximum possible value of R_{obs} is 1.0 (i.e. 100%) and R_{targ} is high, no extension of current interval will be recommended. This will hold true until enough experimental data are collected, resulting in its rejection. Furthermore, the following should be taken into account:

Assuming that R_{targ} = 95% and confidence level is 95%, then the methodology will not recommend an extension of the interval until the following condition is not satisfied, i.e. 59 measurements with a TDMM should be within the permissible scope of measurement parameter's dispersion for 59 measurements. If R_{targ} = 90%, then the recommendation to extend the interval will occur if 29 measurements are within the permissible dispersion scope of measurement parameter from all 29 measurements. If R_{targ} = 85%, then the condition for extension will be fulfilled in 19 measurements in the permissible field of all 19 measurements.

Changing the confidence level also affects the conditions for the extension of the inspected interval. For example, if R_{targ} = 95% and the confidence level C is set to be 80%, a recommendation for an extension of the current interval will occur in 32 measurements within the permissible limits of 32 measurements. If the confidence level is 70%, a recommendation to extend the interval will occur if 24 measurements are within the permissible limits of 24 measurements carried out with a TDMM.

As can be seen above, the higher the target level of reliability R_{targ}, the more difficult it is to obtain an extension of the interval for calibration. This also holds true for shortening the interval. For example, if R_{targ} = 95% and confidence level C is also 95% recommendation to shorten the current interval will occur if there are two measurements that are outside the tolerance of the measured parameter. If the confidence level is 95% and R_{targ} = 64% recommendation to shorten the current interval will occur if three measurements are outside the tolerance of the measured parameter.

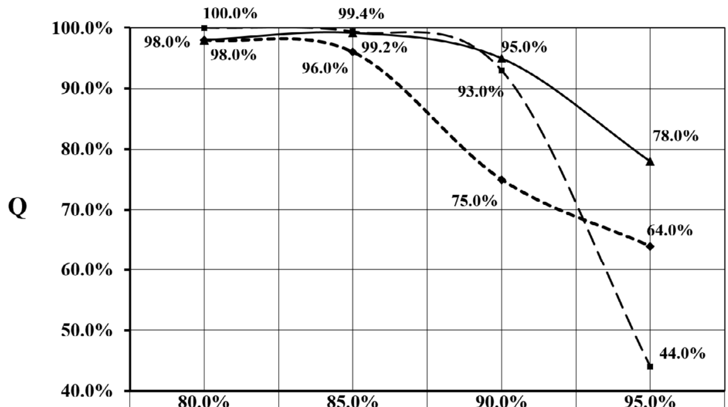
Similar to the extension of the interval, shortening of the verified calibration interval is recommended more often at a lower confidence level C. For example, assuming that R_{targ} = 90% and the confidence level C is set to be 95%, a recommendation to shorten the current interval calibration will occur at two measurements outside the tolerance parameter measurements. However, if the confidence level is set to 89%, then only one measurement outside the tolerance is enough to recommend a shorter current interval.

These results demonstrate that a reduction in the confidence level C leads to a more frequent rejection of current interval and to recommendations for its change. In addition, for higher values of R_{targ}, a decrease in current interval is more often recommended than its increase.

Since the test intervals are likely to decrease rather than increase, the confidence level C can be raised in order to achieve higher target reliability R_{targ} for the inspected TDMM, bearing in mind that the excessive increase more frequently leads to the recommendation for changing the calibration interval.

Under similar additional conditions, it could be concluded that higher values of target reliability (R_{targ}) lead to a lower level of confidence (Q) in the change of the interval.

Figure 2 shows the relationship between the parameters' alterations "confidence level" Q and "target reliability" R_{targ} - Q = f (R_{targ}) for three types of tests: Q1 - for 20 measurements; Q2 - for 30 measurements and Q3 - for 50 measurements.



	80.0%	85.0%	90.0%	95.0%
Q1 (20 tests)	0.98	0.992	0.95	0.78
Q2 (30 tests)	0.98	0.96	0.75	0.64
Q3 (50 tests)	1	0.994	0.93	0.44

R_{targ}

Fig. 2. Interrelationship between "level of confidence" Q and "target reliability" R_{targ} parameters.

The figure illustrates clearly for all three tests that an increase in the value of R_{targ} beyond a certain threshold (R_{targ}> 80-85%) leads to abrupt decrease in the confidence level Q regarding the change of calibration interval. Therefore, using this methodology it is recommended that the target reliability should be selected within 80% ≥ R_{targ} ≤ 90%.

4 Conclusion

The implementation of a large number of methods for the determination of calibration intervals within a company is not recommended since this complicates the supporting process of TDMMs and results in higher costs on the one hand, and requires highly qualified human resources on the other.

The methods described in [10, 11] could be problematic for application in small and medium-size enterprises.

Despite its belonging to the so called algorithmic methods, the proposed methodology for verifying and modifying the calibration interval of a TDMM overcomes some of their shortcomings.

The main advantage of the method is that there are no changes in the calibration interval unless it is rejected based on statistical data checks in the calibration of a TDMM. Furthermore, in case of rejection of the inspected interval, the methodology contains a simple algorithm for calculating the recommended calibration interval with regard to the restricted conditions pre-set by the user (I_{min}, I_{max}).

The methodology uses relatively simple logical and mathematical algorithms which make it manageable for a wider range of specialists, at the same time allowing easy automation with modern software products.

This makes it possible for the proposed methodology to be implemented as an independent tool for the control and regulation of calibration intervals in organizations that use and manage a few TDMMs and to be integrated as a module of more complex systems for metrological analysis.

In the future, we will focus on the methods for determining calibration intervals applied in sectorial management systems, e.g. in automotive and aerospace industry and medical devices manufacturing. Our research work will be based on standards such as 16949, 9100, 13485, VDA standards, OEM CSR, etc. regarding the management of the measuring and monitoring devices. We will also elaborate on the specific approaches for metrology confirmations like Gauge repeatability and reproducibility (GRR) and Capability of Gauges (Cg & Cgk) for new and repaired ones.

References

1. Quality Management Systems - Requirements (ISO 9001:2015), ISO, 44
2. Environmental Management Systems - Requirements with Guidance for use (ISO 14001:2015), ISO, 48
3. Conformity assessment - Requirements for the operation of various types of bodies performing inspection (ISO/IEC 17020:2012), ISO, 28
4. General requirements for the competence of testing and calibration laboratories (ISO/IEC 17025:2005/Cor.1:2006), ISO, 39
5. Measurement management systems - Requirements for measurement processes and measuring equipment (ISO 10012:2003), ISO, 28
6. ISO/DIS 45001, Occupational health and safety management systems - Requirements with guidance for use, ISO, (to be published)
7. Ordinance on measuring instruments subject to metrological control, promulgated in the State Gazette of the Republic of Bulgaria **22** (2015)
8. A. Williams, SLR Ellison, M Roesslein, *EURACHEM/CITAC Guide Quantifying Uncertainty in Analytical Measurement* (LGC, UK, 2000)
9. *EURACHEM/CITAC Guide: Use of uncertainty information in compliance Assessment* (LGC, UK, 2007)
10. ILAC-G24/OIML D 10: Guidelines for the determination of calibration intervals of measuring equipment used in testing laboratories
11. Recommendations on interstate standardization, RMG 74-2004. State system for ensuring the uniformity of measurements, State measurement system. Methods for determining the intervals of verification and calibration of measuring instruments
12. Radev Hr., Dyukendzhiev G., Zakharov I. et al., Metrology and Measuring Equipment, 1, SOFTRADE, **489**, (2008)
13. http://www.isgmax.com - The Integrated Sciences Group Website
14. Establishment and Adjustment of Calibration Intervals, Recommended Practice RP-1, National Conference of Standards Laboratories, (1996)
15. Jackson, D., Ferling, J. and Castrup, H., Concept Analysis for Serial Number Based Calibration Intervals, Proc. Meas. Sci. Conf., **23** (1986)
16. *Numerical Recipes* (Cambridge University Press, 1992)