1. Introduction

The requirements of the Measuring Instruments Directive 2014/32/EU (MID) [1] form the basis of the legislation of Ukraine on conformity assessment of measuring instruments (MI). According to the new version of the Law of Ukraine “On metrology and metrological activity” (came into force on 01.01.2016), MI intended for application in...
the field of legally regulated metrology must undergo a procedure of conformity assessment with the requirements of technical regulations (TR). This procedure is the process of proving that the essential requirements of the TR regarding the MI have been fulfilled. The conformity assessment procedure covers specialized MI software. Such MI must meet the essential requirements, including those related to the software, namely: suitability for use and protection against unauthorized interference.

The rules and procedures for testing MI software are established by the document [2] of the International Organization of Legal Metrology (OIML), as well as documents and recommendations of regional metrology organizations. Testing procedures for MI software are governed by the recommendation [3] of the Euro-Asian Cooperation of National Metrological Institutions (COOMET), document [4] and guidelines [5, 6] of the European Cooperation in Legal Metrology (WELMEC). At the national level, appropriate lists of national standards have been established, which, in particular, give the presumption of MI compliance with the essential requirements of the TR. The analysis of the state of the regulatory framework for testing MI software at the international, regional and national levels has been the subject of previous research [7–10].

National metrology institutes and conformity assessment bodies test MI software according to previously established methods and algorithms. From January 2016 to the present, the number of MI software tests has been more than 500 in Ukraine and is constantly growing at a significant rate. The latter are based on the requirements of the national standard [11] with the additional use of the OIML D 31 [2] and WELMEC 7.2 requirements [5]. This approach contributes to the consideration of all the elements necessary to achieve the presumption of compliance of the software with the essential requirements of the TR. However, these approaches do not answer the question of the quality of software compliance assessment.

The urgency of the work is confirmed by the urgent need for conformity assessment of legally regulated MI in accordance with the requirements of national legislation, TR or European directives. National metrology institutes and conformity assessment bodies are interested in effective testing methods for MI software and in assessing the risks associated with the application of such MI software. Given this, the pressing issue is to validate the MI software test results.

2. Literature review and problem statement

The analysis of various aspects of MI software testing has been the subject of previous research of the authors [7–10].

In [7], the peculiarities of the regulatory support of MI software testing are investigated. The main stages of MI software testing and features in accordance with the requirements [2, 5, 6] are discussed in [8]. The main factors and algorithms for MI software testing in accordance with OIML and WELMEC requirements are considered, a universal algorithm for MI software testing is proposed in [9]. However, these studies did not analyze quality assessment indicators of MI software regarding the effect on the overall test results of MI software.

In [10], the main differences are identified and the necessary elements are established to achieve the presumption of software compliance with the essential TR requirements when conformity assessment of MI. However, the methods and algorithms described in [10] make it impossible to determine the validity of the results obtained from conformity assessment of MI software.

In [12], approaches to software quality requirements and software testing methods are compared, and different approaches to software quality assessment in different international standards and guidelines, in particular on issues related to the quality assessment of MI software, are determined.

In [13], issues related to the validation of MI software covered by the MID [1] are considered. A methodology is presented that can be extended not only to software on MI categories that fall under the MID, but also to most other MI categories. The paper [14] also discusses the validation of MI software based on risk classes for MI software and some possible testing methods.

In [15–18], a method for assessing the risks and current threats posed by MI software, including those integrated into open networks is presented. The method uses a structural and combines elements of specialized international standards and may be useful for conformity assessment bodies and industry.

However, [11–15] do not provide a comparative analysis of the importance of the impact of specific characteristics of MI software on the overall result of the software quality assessment. The requirements of the international document [2] and the possibility of software application for local MI are not taken into account in [16–18].

In [19], a system architecture is considered that could eliminate the risks of general-purpose operating systems. This is achieved both through the use of custom software and control of communication between major software components and the environment.

Thus, it can be concluded that the above studies did not analyze the influential indicators and results of quality assessment of MI software, and did not apply methods of their validation. Therefore, the analytic hierarchy process (AHP) was chosen to investigate such complex objects as MI software: the basic method [20, 21] and its modifications [22–24]. This method allows to structure a complex decision-making problem in the form of a hierarchy in a clear and rational way, to compare and quantify alternative solutions. Recently, AHP has been actively used in practice in various fields of activity. The AHP mathematical apparatus is described in detail in [25]. It is therefore necessary to conduct research and identify the most influential indicators that are analyzed in assessing the suitability of MI software, with a view to improving MI software testing methods.

3. The aim and objectives of the study

The aim of the study is to develop approaches to improve methods of testing and conformity assessment of special MI software at the national level.

To achieve this goal, the following objectives were set:
– to carry out a comparative analysis of the MI software testing results by all indicators using the chosen method;
– to determine the quality indicators of both built-in and universal computer MI software, which have the greatest impact on the overall assessment results.
4. Materials and methods of research for measuring instrument software quality assessment

The problem of comparative analysis of the MI software testing results using AHP is solved by means of three hierarchy levels:

– the first level of the hierarchy corresponds to the aim – to define the most preferred MI software;

– the second – contains criteria (indicators) to define the most preferred MI software;

– the third is the specific MI software that should be compared.

In general, the list of indicators should be such that the most comprehensive evaluation of each MI software is made. Each generalized indicator can be estimated by partial indicators contained in software documents or other available sources. For a relevant comparison, when assessing a particular MI software, it is necessary to consider all the elements compared. Therefore, they are grouped into generalized indicators, each of which is evaluated separately, and pairwise comparisons and all other stages of assessment, using the AHP, are performed on the basis of generalized indicators.

The main stages of comparative quality assessment of MI software based on AHP are as follows [25]:

1. Perform the following actions:
   – to compile a list of M of MI software that will be compared;
   – to carry out an analysis of available information about MI software (software description, user manual, etc.);
   – to determine the list of indicators for comparative assessment, which should contain a sufficient number of indicators (not more than 9) in order to fully reflect all the essential features of MI software.

2. Determine the numerical values of the relative importance of the indicators \( a_{ij} (i, j = 1, 2, ..., N) \), included in the pairwise comparison matrix (PCM) \( A \) of assessment indicators, established by the evaluation expert, directly during each specific MI software comparison.

3. Determine:
   – Normalized eigenvector \( A_i \) of the matrix \( A \);
   – consistency index of output data \( I_c \);
   – maximum eigenvalue \( \lambda_{max} \).

4. Determine:
   – numerical values of the elements of the PCM \( B_k \) \( b_{ij}^k \) \( (i, j = 1, 2, ..., M) \), using available information from various sources to establish the array of \( B_k \) pairwise comparison matrices of MI software, according to each indicator;
   – normalized eigenvectors for each constructed matrix \( B_k \).

5. Implement the consistency check of the local priorities that are included in the \( B_k \) matrices using the consistency relation \( C_I \) \( (C_I \leq 0.1 \text{ – consistency condition}) \). Determine the consistency index of the original data \( I_c^k \) and maximum eigenvalues \( \lambda_{max}^k \).

6. Determine the generalized priorities \( G_n \) for each of \( M \) of MI software compared by:

\[
G_n = \sum_{i=1}^{N} B_{i,n} \quad n = 1, 2, ..., M
\]  

with the further ranking of global priorities \( G_n \) for all MI software and definition of MI software that has the greatest advantage – the software with the maximum value \( G_n \).

The structure of the model of quality assessment of MI software using AHP is shown in Fig. 1.

According to the results of the analysis of the requirements for testing the MI software of the WELMEC 7.2 guidelines [4], the following generalized indicators can be distinguished for assessing the quality of the MI software:

\( K_p \) – built-in computer software characteristic (P);

\( K_U \) – universal computer software characteristic (U);

\( K_I \) – test indicator of storage devices (L);

\( K_T \) – test indicator of data transfer devices (T);

\( K_R \) – reading test indicator (S);

\( K_D \) – test indicator of software separation levels (D);

\( K_I \) – specific test indicator of software for specific MI (I).

The structure of links between the requirements for quality assessment of MI software according to WELMEC 7.2 [4] is shown in Fig. 2.

For pairwise comparison of all generalized indicators with the help of AHP, including quantitative and qualitative, with the submission of the comparison result in quantitative form, one should use the Saati scale [20].

The list of partial indicators, which make up the generalized indicators \( K_p, K_U, K_I, K_T, K_R, K_D, K_I \) and expressions to obtain a numerical value for each generalized indicator are determined.

The numerical value of the built-in computer software characteristic \( K_p \) is determined by

\[
K_p = \sum_{i=1}^{N} P_i a_i^p / N_p,
\]  

where \( N_p \) – the total number of estimated MI software indicators \( (N_p = 7) \);

\( P_i \) – constituent estimates of the generalized indicator \( K_p \) (numerical characteristics of partial indicators with certain weight coefficients \( a_i^p \)):  

\[
G_n = \sum_{i=1}^{N} B_{i,n} \quad n = 1, 2, ..., M
\]  

with the further ranking of global priorities \( G_n \) for all MI software and definition of MI software that has the greatest advantage – the software with the maximum value \( G_n \).

The structure of the model of quality assessment of MI software using AHP is shown in Fig. 1.

According to the results of the analysis of the requirements for testing the MI software of the WELMEC 7.2 guidelines [4], the following generalized indicators can be distinguished for assessing the quality of the MI software:

\( K_p \) – built-in computer software characteristic (P);

\( K_U \) – universal computer software characteristic (U);

\( K_I \) – test indicator of storage devices (L);

\( K_T \) – test indicator of data transfer devices (T);

\( K_R \) – reading test indicator (S);

\( K_D \) – test indicator of software separation levels (D);

\( K_I \) – specific test indicator of software for specific MI (I).

The structure of links between the requirements for quality assessment of MI software according to WELMEC 7.2 [4] is shown in Fig. 2.

For pairwise comparison of all generalized indicators with the help of AHP, including quantitative and qualitative, with the submission of the comparison result in quantitative form, one should use the Saati scale [20].

The list of partial indicators, which make up the generalized indicators \( K_p, K_U, K_I, K_T, K_R, K_D, K_I \) and expressions to obtain a numerical value for each generalized indicator are determined.

The numerical value of the built-in computer software characteristic \( K_p \) is determined by

\[
K_p = \sum_{i=1}^{N} P_i a_i^p / N_p,
\]  

where \( N_p \) – the total number of estimated MI software indicators \( (N_p = 7) \);

\( P_i \) – constituent estimates of the generalized indicator \( K_p \) (numerical characteristics of partial indicators with certain weight coefficients \( a_i^p \)):  

\[
G_n = \sum_{i=1}^{N} B_{i,n} \quad n = 1, 2, ..., M
\]  

with the further ranking of global priorities \( G_n \) for all MI software and definition of MI software that has the greatest advantage – the software with the maximum value \( G_n \).

The structure of the model of quality assessment of MI software using AHP is shown in Fig. 1.

According to the results of the analysis of the requirements for testing the MI software of the WELMEC 7.2 guidelines [4], the following generalized indicators can be distinguished for assessing the quality of the MI software:

\( K_p \) – built-in computer software characteristic (P);

\( K_U \) – universal computer software characteristic (U);

\( K_I \) – test indicator of storage devices (L);

\( K_T \) – test indicator of data transfer devices (T);

\( K_R \) – reading test indicator (S);

\( K_D \) – test indicator of software separation levels (D);

\( K_I \) – specific test indicator of software for specific MI (I).

The structure of links between the requirements for quality assessment of MI software according to WELMEC 7.2 [4] is shown in Fig. 2.

For pairwise comparison of all generalized indicators with the help of AHP, including quantitative and qualitative, with the submission of the comparison result in quantitative form, one should use the Saati scale [20].

The list of partial indicators, which make up the generalized indicators \( K_p, K_U, K_I, K_T, K_R, K_D, K_I \) and expressions to obtain a numerical value for each generalized indicator are determined.

The numerical value of the built-in computer software characteristic \( K_p \) is determined by

\[
K_p = \sum_{i=1}^{N} P_i a_i^p / N_p,
\]  

where \( N_p \) – the total number of estimated MI software indicators \( (N_p = 7) \);

\( P_i \) – constituent estimates of the generalized indicator \( K_p \) (numerical characteristics of partial indicators with certain weight coefficients \( a_i^p \)):  

\[
G_n = \sum_{i=1}^{N} B_{i,n} \quad n = 1, 2, ..., M
\]  

with the further ranking of global priorities \( G_n \) for all MI software and definition of MI software that has the greatest advantage – the software with the maximum value \( G_n \).

The structure of the model of quality assessment of MI software using AHP is shown in Fig. 1.

According to the results of the analysis of the requirements for testing the MI software of the WELMEC 7.2 guidelines [4], the following generalized indicators can be distinguished for assessing the quality of the MI software:

\( K_p \) – built-in computer software characteristic (P);

\( K_U \) – universal computer software characteristic (U);

\( K_I \) – test indicator of storage devices (L);

\( K_T \) – test indicator of data transfer devices (T);

\( K_R \) – reading test indicator (S);

\( K_D \) – test indicator of software separation levels (D);

\( K_I \) – specific test indicator of software for specific MI (I).

The structure of links between the requirements for quality assessment of MI software according to WELMEC 7.2 [4] is shown in Fig. 2.
where

\[ K_{U} = \sum_{i=1}^{9} U_i w_i^U / N_{U}, \]  

where \( N_U \) – the total number of estimated MI software indicators (\( N_U = 9 \));

\( U_i \) – constituent estimates of the generalized indicator \( K_U \) (numerical characteristics of partial indicators with certain weight coefficients \( w_i^U \)):

\( U_1 \) – documentation;
\( U_2 \) – software identification;
\( U_3 \) – influence through user interfaces;
\( U_4 \) – impact through transmission interfaces;
\( U_5 \) – modification protection;
\( U_6 \) – software modification protection;
\( U_7 \) – parameter protection;
\( U_8 \) – software authentication and results transfer;
\( U_9 \) – impact of other software.

The numeric value of the test indicator of storage devices \( K_L \) is determined by

\[ K_L = \sum_{i=1}^{8} L_i w_i^L / N_{L}, \]  

where \( N_L \) – the total number of estimated MI software indicators (\( N_L = 8 \));

\( L_i \) – constituent estimates of the generalized indicator \( K_L \) (numerical characteristics of partial indicators with certain weight coefficients \( w_i^L \)):

\( L_1 \) – completeness of stored data;
\( L_2 \) – protection against accidental or conscious modification;
\( L_3 \) – data integrity;
\( L_4 \) – authenticity of stored data;
\( L_5 \) – conference keys;
\( L_6 \) – recovery of stored data;
\( L_7 \) – automatic storage;
\( L_8 \) – storage capacity and sequence.

The numeric value of the test indicator of data transfer devices \( K_T \) is determined by

\[ K_T = \sum_{i=1}^{8} T_i w_i^T / N_{T}, \]  

where \( N_T \) – the total number of estimated MI software indicators (\( N_T = 8 \));

\( T_i \) – constituent estimates of the generalized indicator \( K_T \) (numerical characteristics of partial indicators with certain weight coefficients \( w_i^T \)):

\( T_1 \) – completeness of the transmitted data;
\( T_2 \) – protection against accidental or conscious modification;
\( T_3 \) – data integrity;
\( T_4 \) – authenticity of transmitted data;
\( T_5 \) – conference keys;
\( T_6 \) – spoiled data processing;
\( T_7 \) – transmission delay;
\( T_8 \) – suitability of transmission services.

The numeric value of the reading test indicator \( K_S \) is determined by

\[ K_S = \sum_{i=1}^{9} S_i w_i^S / N_{S}, \]  

where \( N_S \) – the total number of estimated MI software indicators (\( N_S = 3 \));

\( S_i \) – constituent estimates of the generalized indicator \( K_S \) (numerical characteristics of partial indicators with certain weight coefficients \( w_i^S \)):
comparative quality assessment of MI software, then they can be modified as necessary.

**Table 1**

<table>
<thead>
<tr>
<th>Numerical values of the elements of the PCM of indicators for comparative quality assessment of MI software</th>
</tr>
</thead>
<tbody>
<tr>
<td>$K_P$</td>
</tr>
<tr>
<td>-------</td>
</tr>
<tr>
<td>1.014</td>
</tr>
</tbody>
</table>

In the case of the data in Table 1, the consistency condition for the consistency ratio ($C_R < 0.1$) is satisfied ($C_R = 0.02$). The consistency index $I_c = 0.028$, and the largest eigenvalue of the vector is $\lambda_{\text{max}} = 7.17$.

The weight coefficients $w_i$ for the selected MI software quality assessment indicators are defined in Table 2.

**Table 2**

<table>
<thead>
<tr>
<th>Weight coefficients for the selected MI software quality assessment indicators</th>
</tr>
</thead>
<tbody>
<tr>
<td>$i$</td>
</tr>
<tr>
<td>------</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td>4</td>
</tr>
<tr>
<td>5</td>
</tr>
<tr>
<td>6</td>
</tr>
<tr>
<td>7</td>
</tr>
<tr>
<td>8</td>
</tr>
<tr>
<td>9</td>
</tr>
<tr>
<td>10</td>
</tr>
</tbody>
</table>

At the beginning of the assessment, it is necessary to determine the basic configuration of software: with a built-in computer $P$ or with a universal computer $U$. Then, a complete set of requirements relating to the corresponding basic configuration must be used.

Eight software were selected for comparative quality assessment:
- six built-in computer MI software: for electricity meters (SW1, SW2, SW6), gas distribution columns (SW3), gas chromatograph (SW4), heat meter (SW5);
- two universal computer MI software: for cardiac monitoring (SW7) and liquid chromatograph (SW8).

The numerical values of the MI software quality indicators were converted into the numbers needed for assessment by the AHP in the range from 1 to 9 using the Saati scale. The results of assessment of MI software quality indicators in accordance with the presented methodology are shown in Table 3. Specialized software “AHP Competence 1.2” (Ukraine), which implements AHP, was used for the necessary calculations.

The comparison of the global priorities $G_n$ of the MI software under consideration with the ranking by their reduction (AHP 1.2 Competence, Ukraine) is shown in Fig. 3.

**Table 3**

<table>
<thead>
<tr>
<th>Results of assessment of MI software quality indicators</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>-----------------</td>
</tr>
<tr>
<td>Indicators</td>
</tr>
<tr>
<td>$K_P$</td>
</tr>
<tr>
<td>$K_U$</td>
</tr>
<tr>
<td>$K_L$</td>
</tr>
<tr>
<td>$K_T$</td>
</tr>
<tr>
<td>$K_S$</td>
</tr>
<tr>
<td>$K_D$</td>
</tr>
<tr>
<td>$G_n$</td>
</tr>
</tbody>
</table>
The analysis of the obtained results shows the advantage of MI software in the following sequence: SW1>SW4>SW5>SW7>SW8>SW6>SW2>SW3. It should be noted that the universal computer MI software has an average level of quality. The built-in computer MI software for the SW1 and SW2 meters has significantly different levels of quality.

6. Discussion of the results of comparative software quality assessment

The consistency relation $C^*_f$ for all MI software with a large margin meets established requirements ($C^*_f \leq 0.1$). This means that the results obtained are acceptable and that further analysis is possible. Maximum eigenvalues $\lambda_{max}$ vary within a wide range from 7.985 (for the $K_D$ indicator) to 8.099 (for the $K_T$ indicator) for all quality indicators of MI software. It is important to analyze the values of the MI software quality indicators to identify the weight of their impact on the overall software quality assessment. Comparative analysis of the priority vectors of the MI software quality indicators shown in Table 3 showed the following.

Without the submission of documentation on software with MI with built-in and universal computers and its identification, it is not possible to start the quality assessment process in accordance with the established requirements. That is, built-in computer ($K_P$) and universal computer ($K_U$) software indicators are important by default, although their maximum eigenvalues are average (8.000 and 7.994, respectively).

The test indicator of storage devices ($K_L$) and the test indicator of data transfer devices ($K_T$) are some of the important quality indicators (maximum eigenvalues of 8.073 and 8.099, respectively).

At the same time, reading test indicator ($K_S$), specific test indicator of software for specific MI ($K_I$), and test indicator of software separation levels ($K_D$) have a small contribution to the overall quality indicator (maximum eigenvalues of 8.000, 7.994 and 7.985, respectively).

The diagram of the weight of MI software quality indicators, which allows to present clearly the obtained results, is shown in Fig. 4.

It should be noted that the reading test indicator ($K_S$) and the test indicator of data transfer devices ($K_T$) are practically impossible to apply to software with a built-in computer that is used in simple MI (such as simple electricity meters). This is because such MI lack reading and data transfer devices. At the same time, these MI software quality indicators are important for software with a versatile computer used in complex MI, such as cardio monitors and liquid chromatographs.

Thus, the main quality indicators of MI software with a built-in and versatile computer that have the greatest impact on the results of conformity assessment, are identified.

The results of the study can later be used to modify and improve the algorithm and methodology for conformity assessment of MI software.

7. Conclusions

1. A comparative analysis of the testing results of software for MI with built-in and universal computers using AHP is carried out. Based on the analysis of the requirements of the
WELMEC 7.2 guidelines for MI software testing, generalized and partial indicators are identified to assess the quality of MI software. Expressions to obtain the numerical value of each partial indicator by each generalized indicator are generated. The results of the comparison showed the AHP suitability for pairwise comparisons of all quantitative and qualitative indicators of quality assessment of MI software.

2. The quality indicators of MI software with built-in and universal computer, which have the greatest impact on quality assessment results are determined. It is found that without the submission of documentation and identification of MI software with built-in and universal computer, it is not possible to start the conformity assessment procedure as required. That is, the quality indicators of built-in computer (Kp) and universal computer (KU) software are important by default. Also, the test indicator of storage devices (Kl) and the specific test indicator of software for specific MI (Kl) are some of the important indicators. It is determined that the reading test indicator (Kr) and the test indicator of software separation levels (Ko) are practically inapplicable and can be neglected.

References
1. Introduction

The Quanton diagnostic and health complex is a technical innovation designed to perform a triple function: non-invasive diagnostics of people, gaining information about the characteristics of the desired health normalizing effect on them and implementation of this effect. Non-invasive diagnostics combines spectral and binary methods. The spectral method with a certain level of reliability allows identifying organs that have deviations from the standards. The binary