



MANAGEMENT OF MEASUREMENT UNCERTAINTY AS A SYSTEM IN AN ACCREDITED LABORATORY

Tibor GREGÁSZ¹, Veronika PÁL²

¹⁻²Óbuda University RKK

Abstract:

The main process of laboratories produces a distinctive "product" such as the measurement result. Consequently, the uncertainty of the measurement procedure determines one of the most important quality requirements. For accredited laboratories, their standard requirements (in this particular case: ISO/IEC 17025:2017) request measurement uncertainty ("u") to be evaluated and determined for each procedure.

Laboratories intend to meet this requirement, but in case of certain procedures the complex process of determination, unknown uncertainty components and subjective test methods make it difficult to achieve this goal.

In order to fulfill this obligation, it is necessary for a quality-controlled laboratory to systematically monitor which tests have already defined their method's uncertainty, which calculations still need to be done, and which should be reduced by improvements.

The textile testing laboratory in question performs hundreds of different test methods, which has led the management to set the objective of systematically monitoring and reporting measurement uncertainties. This includes evaluating the test methods and their respective 'u' components, the determination method of 'u', plans for their reduction and their implementation.

Throughout this work, we present the development of this system and the process of assigning the "u" for a color fastness test.

Keywords:

measurement uncertainty, colour fastness testing, uncertainty components, measuring equipment supervision, monitoring.

1. INTRODUCTION

One of the laboratory's quality improvement concepts is to assess the existence of measurement uncertainty values for all test procedures used. After filling in the gaps found during the evaluation, the measurement uncertainty will be known. Although the standard defining general requirements of testing laboratories clearly states that contributions to uncertainty shall be known and evaluated, we do not have any information or references on this from laboratories in the same field. In other words, there are certain textile testing methods where this requirement of the standard is not currently fulfilled.

In our view, the documented result of the evaluation would be more than a simple two-column list, where the numeric value next to the name of the test method would tell us if it is already known or - in case of an empty cell - not. According to the distant image in the figure below, in four blocks a total of 15 aspects of information are assigned to each test method in the context of knowledge of measurement uncertainty. Creating this register has initiated a process in the laboratory that, in addition to simply filling in measurement uncertainties, provides a systematic review of procedures, risk-based priority treatment, accurate numerical input, and, as an essential element, monitoring of continuous progress.



2. MEASUREMENT UNCERTAINTY IN GENERAL

1.1 The meaning of measurement uncertainty

The most common definition of this phenomenon and this attribute of measurement methods is the following: "A parameter associated with a measurement result that describes the variation of values reasonably attributable to the quantity to be measured". It can be given by the variation characteristic calculated from the measurement results (Type A determination) or by other scientifically valid information (Type B determination). This concept is intrinsically linked to all measurement results existing. A general principle of metrology is that the true value of the measured characteristic cannot be entirely defined, and various types of measurement errors may occur under the most stringent conditions, and in most cases may also accumulate. Therefore, it is important to realize that a measurement result, due to its many influencing factors, can never be completely identical and will deviate from its true value, namely encumbered with error. [1]

1.2 The role of measurement uncertainty in the work of testing laboratories

Because measurement uncertainty is a "quantitative indication of the quality of a result" [2], in practice it defines one of the most important quality requirements for the "product" – i.e. a measurement result - produced by the main process of laboratories. The accuracy of a testing laboratory is not only demonstrated by its ability to measure results close to the reference value, but also by its recognition and handling of its measurement uncertainties and the factors affecting it. In addition to being involved in interlaboratory round robin tests, this is a quality indicator that can and must be improved. It is very important to pay attention and dedicate energy to this concept, although its profitability is not self-evident. For accredited laboratories, their applicable standard requirements (e.g. ISO 17025:2017) require that for each procedure, the components of measurement uncertainty and their estimated or specific value (u) be determined. [3] As many test standards fail to do so, there is a need for a system that monitors the definition and development potential of the uncertainty factors and values for each test method.

3. MANAGEMENT OF MEASUREMENT UNCERTAINTY IN A TEXTILE LABORATORIES

One of the laboratory's quality improvement ideas is to assess the existence of measurement uncertainty values for all test procedures applied. To remedy the deficiencies found during the evaluation, the measurement uncertainty must be known. Although the standard defining the general requirements for the competence of testing laboratories clearly states that this aspect of the testing method is to be known, we do not have any information or references on this from laboratories in the same field. In other words, there are processes in textile testing where this requirement of the standard is not currently fulfilled.

Therefore, the worksheet created during the research will be significantly more advanced than a simple two-column list, in which the numeric value next to the name of the test method would tell us whether we know the uncertainty value of the particular test or not. It is also the technical appearance of a methodology of a registry and monitoring system.



1.3 Tests to be included in the measurement uncertainty register

As the first step of classification, we separate tests providing qualitative results, which give a logical Yes/No as an output. Essentially, the answer given for the question imposed on the test object is the fulfillment or non-fulfillment of a numerically inexpressible criterion. (For example, the sample is soluble or not in a given solution; melted or not; the test material is permeable or not.) It is also important to know the uncertainty of these tests, but our current research and development is only suitable for measuring methods. In their case, we plan to use qualitative (so-called attributive) Gage R&R, where we will give a measure of the consistency of the results of approx. 10 to 20 repeated tests. As it has already been the case with measurements, here the recommendations of the automotive measurement system analyzers' reference manual (MSA) will be used in greater proportion. When defining uncertainties, we often emphasize an analytical technique, indicator, or a comparison methodology.

The output of a measurement test is typically a number multiplied by a unit of measurement, to which measurement uncertainty characterizes the range of deviations in the quality of the procedure. Thus, for measurement test results, the uncertainty of measurement is given by some quantity of variance.

1.4 General description and significance of the uncertainty register

The basic concept of managing measurement uncertainty is outlined in the table below. However, filling it in carefully and competently and keeping it up to date is more than a documented piece of information (as it is mentioned in today's management standards) that can only support the fact in an audit that a criterion is met.

The creation of this spreadsheet triggered a process that, in addition to simply filling in measurement uncertainty values, provides a systematic review of procedures, risk-based prioritization, accurate numerical input, and, as an essential element, monitoring of continuous progress. Thus, the Excel spreadsheet seems to be only a visual representation of the method and somewhat capable of recording facts, but much more than that. In fact, a metrology-sensitive expert group has been formed in the lab, which meets regularly to review the current state of the lab's tests and assign tasks. Alterations are made as equipment change, new experiences in measurement procedures are obtained, sample preparation techniques, and standards change. As a result of the changes just described, one of the components of uncertainty is likely to change in the right direction and ultimately require recalculation of the "U". Therefore, there is a need for keeping this "data" for measurement procedures as accurate as possible, also reflecting laboratory conditions.

The spreadsheet (based on the number of tests and the amount of information known about them as a database) will not be complete at the time of its creation and will never be one that always has all cells filled for the following reasons. Upload is labour-intensive and logically the most frequent tests are prioritized. Uncertainties in the past were not defined namely that often, so most of them are waiting the working group's meeting. Entering Type B-uncertainty may require experiments and calculations that need to be planned into the daily work schedule, thus slowing down the completion of the chart. This group should be expected to fill in the column(s) regarding development issues.

The specific uncertainty figures may not yet be known, but for metrology-related duties this chart will provide baseline information and help to schedule tasks.

1.5 Detailed description of the uncertainty spreadsheet

The measurement uncertainty spreadsheet represents the following functions in its rows and columns:

- In its ranks it summarizes the tests within – and possibly outside as well of - the scope of the accreditation. Obviously, it is of paramount importance in the field of accredited examinations. As mentioned above, in the first part of the work we've been including quantitative tests, but later on they can be extended to qualitative tests or to those made by other structures. In some cases, it is possible to standardize tests. This occurs when one or more of the phases or conditions of the test methods do not differ or are shown to be irrelevant to the "U".
- The columns were divided into four thematic blocks
 - I. the first one is to handle the uncertainty components;
 - II. the second block is able to rank the significance of the measurement results on a risk basis; and
 - III. the third is to clarify and calculate the value of the assignment "U".
 - IV. a block that records the development plan and is specifically designed to reduce measurement uncertainty.

Uncertainty contributions			Risk of measurement error					Determination of U			Development of test method		
Known contributions to uncertainty	Significant contributions to uncertainty	Suspected (based on practical knowledge) but not yet proven contributions	Frequency of test (monthly)	Rating of frequency (1 - lowest; 4 - highest)	Importance of test	Rating of importance (1 - lowest; 4 - highest)	Risk rating	Type of U	Subtype	Amount of measurement uncertainty	What kind of development is planned?	Responsible person	Proposed deadline
defined in the standard		sampling	2-5x - rare	2	test for regulatory purposes (REACH and EU Regulation 1007/2011 on fiber composition)	3	6	8	defined from test standard		internal training	KV	30.04.2020

Figure 1: Image of the contents of the blocks

The first block is handling the uncertainty components along the columns:

		Uncertainty contributions		
		Known contributions to uncertainty	Significant contributions to uncertainty	Suspected (based on practical knowledge) but not yet proven contributions
Laboratory test	Test method			
Determination of the raw material composition. Mixtures of silk and wool or hair	MSZ EN ISO 1833-18	defined in the standard		sampling
Determination of the raw material composition. Mixtures of polyester and certain other fibres	MSZ EN ISO 1833-25	defined in the standard		expiry date of chemicals & sampling

Figure 2: Image of block and details of measurement uncertainty

- 1) In case they are known, the first column records all realistic sources of uncertainty components. As knowledge and testing techniques evolve, new resources may be involved. It should be noted

that the amount of uncertainty factors listed does not directly correlate with the calculated uncertainty value as many insignificant ones are not comparable to one significant

- 2) In the second column, the list that really plays a significant role in the uncertainty value is listed as a significant factor in the second column. This second cell is important because it is necessary to refer to these in the description of the measurement procedure (standard or validation in the description of the measurement process) as they result in a significant distortion effect. With the advancement of the test technique, the uncertainty factor can be removed here by downgrading.
- 3) In the third column, we record information from an investigation which factors are suspected but not currently proven to be influential.

The functions of the second block measure the risk associated with the test result and try to give a numerical risk value to it. Functions along the columns:

		Risk of measurement error				
		Frequency of test (monthly)	Rating of frequency (1 - lowest; 4 - highest)	Importance of test	Rating of importance (1 - lowest; 4 - highest)	Risk rating
Laboratory test	Test method					
Determination of the raw material composition. Mixtures of silk and wool or hair	MSZ EN ISO 1833-18	2-5x - rare	2	test for regulatory purposes (REACH and EU Regulation 1007/2011 on fiber composition)	3	6
Determination of the raw material composition. Mixtures of polyester and certain other fibres	MSZ EN ISO 1833-25	2-5x - rare	2	test for regulatory purposes (REACH and EU Regulation 1007/2011 on fiber composition)	3	6
Determination of formaldehyde. Part 1: Free and hydrolised formaldehyde (water extraction method)	MSZ EN ISO 14184-1	5-20x - frequent	3	test for regulatory purposes (REACH and EU Regulation 1007/2011 on fiber composition)	3	9

Figure 3: Image of the risk analysis block and its details

		Risk of measurement error	
		Monthly frequency of test	Importance of test
1	1x - memory refresh	informative	
2	2-5x - rare	provides important results for distinguishing/categorization purposes on the market	
3	5-20x - frequent	test for regulatory purposes (REACH and EU Regulation 1007/2011 on fiber composition)	
4	20x< - large scale	test for regulatory purposes (PPE)	

Figure 4: Scale values for risk factors

- 4) The categorization of the frequency of the test is important as the difference resulting from the uncertainty is of course accompanied by a decision error (Type 1 and Type 2). For a larger number of cases, the economic impact of uncertainty (as interpreted by any actor) may be more significant, which should be considered in the fourth block for the development option.
- 5) The drop-down selection list categorizes by the estimated number of tests per month. In our view, more testing can result in more false results and more miscalculations.
- 6) Choosing the test quantity generates a risk value between 1 and 4 ("O" - opportunity) in the cell next to it, known from the FMEA risk analysis method.
- 7) Another crucial factor in risk assessment is importance. This indicates the importance of the test result and the decision made on the basis thereof, from simple information relevant for decision-

making to the safety-related accident or poisoning hazard. It also automatically follows our programmed responses to the built-in selection list

- 8) The program also calculates a severity ("S") rating between 1 and 4, also according to the FMEA notation.
- 9) This column multiplies the number of the two automatically generated risk factors, which technically changes to a green-yellow-red colour for conditional formatting according to a built-in 1-16 risk scale.

The features of the third block record the circumstances under which the measurement uncertainty is given.

		Determination of U		
		Type of U	Subtype	Amount of measurement uncertainty
Laboratory test	Test method			
Determination of the raw material composition. Mixtures of silk and wool or hair	MSZ EN ISO 1833-18	B	defined from test standard	
Determination of the raw material composition. Mixtures of polyester and certain other fibres	MSZ EN ISO 1833-25	B	defined from test standard	
Determination of formaldehyde. Part 1: Free and hydrolised formaldehyde (water extraction method)	MSZ EN ISO 14184-1	B	defined from test standard	

Figure 5: The "U" block and its details of the measurement uncertainty

- 10) First of all, the type of measurement uncertainty reported according to the guideline, Type A (standard deviation based on statistics of actual measurements), and Type B evaluation (derived from other information and assessments) can be chosen.
- 11) According to the Type A assessment, the choice of calculation and reporting methods is automatic here. The options are: s - standard error, repeatability standard deviation, u - standard uncertainty, U - expanded uncertainty, or q - confidence intervals that are less common in metrology, but generally used in statistics.
- 12) For Type-B evaluation, certain methods of measurement uncertainty not included here are also available, according to the built-in selection list.
- 13) In this column the measurement uncertainty value, possibly estimated, using one of the methods of determination described above is given

The fourth block is in the service of continuous improvement, as this area includes the recording of development related to the improvement of the measurement process, including the reduction of measurement uncertainty. The elements below should be clarified following quality assurance practice:

		Development of test method		
		What kind of development is planned?	Responsible person	Proposed deadline
Laboratory test	Test method			
Determination of the raw material composition. Mixtures of silk and wool or hair	MSZ EN ISO 1833-18	internal training	KV	30.04.2020
Determination of the raw material composition. Mixtures of polyester and certain other fibres	MSZ EN ISO 1833-25	internal training & further practical experiments	KV	30.04.2020
Determination of formaldehyde. Part 1: Free and hydrolised formaldehyde (water extraction method)	MSZ EN ISO 14184-1	purchase of new equipment	KV	31.12.2019

Figure 6: Image of the block of development (reduction of uncertainty) and its details

- 14) What should be the purpose of the development, i.e. what do we want to achieve (what reduction method) with which factor of uncertainty.
- 15) Who should be the leader/responsible person for the innovative action?
- 16) When are the results of the development expected?

1.6 Systematization of the components of uncertainty when examining the color fastness to perspiration [4]

To determine measurement uncertainty, we first break down the process leading to the measurement result, since during these steps the measurement result may be distorted and groups of measurement uncertainty components become distinguishable. These factors will be included or discarded in our later model depending on their degree. The steps of the process are illustrated in the rows of a chart. For each process elements, there are additional parameters that can increase uncertainty, so all of them must be subject to at least a critical analysis, but the significant ones should be quantitatively assessed and defined.

The columns of the chart group the mapped parameters according to the elements of the measurement system, and the colour codes represent the classification made during the critical analysis, thus meeting the requirements of ISO 17025:2017 clause 7.6.1.

In addition to providing an accurate picture of uncertainty factors, the spreadsheet illustrates and exemplifies the proportions of the three pillars of the measurement system (namely instrument accuracy, accredited suppliers and operator proficiency). Therefore, in the event of any of these being compromised, it will provide a clearer understanding of the causal relations and helps appropriate risk management.



As a long-term development goal, the laboratory would like to use this method to assess uncertainty in similarly complex test methods using a spreadsheet-based template, taking into account local conditions. By creating additional worksheets, critical components can also be calculated. This way, over time, these tables can become a complete, documented monitoring system about the uncertainty factors of accredited laboratory tests, which will also be subject to continuous improvement. In addition, they can provide effective support for resource allocation and even pricing by accurately specifying the parameters required by the standard.

1.7 Design and implementation of experiments on critical factor

Primarily because of the frequency of examinations, we decided to launch a series of practically determined “U” values by evaluating the specific uncertainty factors for colour fastness to perspiration. The numerical values of the uncertainty components obtained for each module during the experiments can be applied to other tests later on.

Uncertainty due to sampling of textiles

The first common element of the colour fastness testing processes, which is sampling, was identified as one of the suspicious factors of uncertainty to be evaluated. Comparing the results of the sample pairs treated in the exact same way, we obtained a value from the statistical characteristics of the differences, which can finally be used as a numerical factor of uncertainty. After averaging the ranges between the results for a total of 75 pairs of specimens, the standard deviation was estimated. This allows the sampling procedure to extend the measurement uncertainty to

$$U = k \cdot s \quad (1)$$

using the formula **above**. Calculated at a 95% **confidence interval (where k=2)**, the expanded measurement uncertainty ($U_{\text{Sampling95\%}}$) is **0.266 grades; and for 99% (k=3)**, $U_{\text{Sampling99\%}} = 0.399$ grades.

Uncertainty caused by solution preparation

Due to the lack of tolerance limits and a large amount of practical experiences of the measurement method, the concentration of artificial perspiration solutions was chosen. Using a balance accurate to one hundredth grams during measurement, and based on empirical data, a dose error of $\pm 0.1\text{g}$ per component can occur during daily work. Therefore, we prepared a solution that causes the most severe and mildest exposure to dyes and examined the statistical difference between the results of the treated samples. With the difference of the average scores and the corrected empirical standard deviation of the two experimental settings, we used two test groups to see if there was any random difference between the results of the samples of the two different test groups. Otherwise, it would be proven that the results come from two different sets of basic properties. Due to the design of the experiment, this could only happen because of different pre-treatment solutions, since no other parameters have been changed in the process.



First, a statistical F-test was used to prove that a difference between variances is not statistically significant ($F = 1.042$; $F_{\text{crit}} = 1.11$), that is, from an identified group of results. Since both of our groups have a sample size greater than 30, the following formula was used:

$$t = \frac{|\bar{x} - \bar{y}|}{\sqrt{\frac{s_x^2}{n_x} + \frac{s_y^2}{n_y}}} \quad (2)$$

Based on this:

$$t_{\text{vegyszer}} = \frac{|3,74 - 3,72|}{\sqrt{\frac{1,312^2}{96} + \frac{1,339^2}{96}}} = 0,109$$

The values at 100 degrees of freedom in the Student's t-distribution table are well below the 99.5% probability level. Thus, it is clearly demonstrated that the difference between the mean values is not significant, so it can be stated statistically that the results obtained from the capture of the specimens do not come from two different statistical populations.

In the end, we could not demonstrate that the modified test parameters would cause a significant change in the results. This way we can exclude this module from the list of significant uncertainties. The measurement instructions may include a tolerance limit of $\pm 0,1$ g for each component of the test solutions. [16]

Uncertainty caused by the duration of pre-treatment processes

Of the pre-treatment parameters, the duration of soaking and thermal and high-pressure treatment in the oven was also chosen because of the lack of tolerance limits underlying the above consideration and the existence of relevant practical experiences. Therefore, the question for the next experiment was whether there can be any change experienced compared to the original, 30 minutes long pre-treated control group and a soaking duration of 45 minutes with an experimental setting based on practical considerations. Also, we suspected that pressure and thermal treatment extended to 4.5 hours instead of 4 hours can cause fluctuations in the results. The average of 2x12 samples per test type exposed to aqueous, alkaline, and acid solutions showed variations of 0.02 to 0.06 degrees, not always in the same direction. The two-sample statistical tests described in the chapter before did not show any change that could not be called random, so we excluded these two process elements from the significant factors of uncertainty.

Uncertainty caused by visual assessment

The importance of evaluating visual assessment is demonstrated by the lack of precision data of the measuring system, and the key role it plays in all colour fastness tests. The resulting Gage R&R value can also be used for later detection steps of uncertainty. R&R analysis quantifies the repeatability and reproducibility of measurement systems. Repeatability refers to the degree of variation within a measurement sequence, i.e. it gives the variation of measurement results over a short period of time,



leaving all elements of a given measurement system unaltered. Reproducibility defines the difference between a series of measurements that are generated differently by one particular circumstance (such as the person performing the measuring). [5]

With Gage R&R study, we have examined and precisely quantified the accuracy of visual assessment that is subjective, meaning the accuracy of the measurement tools are completely depending on the operators' eyes, professional experience and judgment. To do this, we classified the results of 12-12 different standard pre-treatment processes of 6 samples of raw material, 3 times per operator, excluding the effect of memory. Taking into account the results measured by the three operators, the repeatability of the measuring system (mean of the ranges measured by one operator) is 0.593 grades, its reproducibility (the scales of the averages measured by all the operators) is 0.275 and the R&R value equals 0.654 grades. As a result of the study, the latter value will be used as a factor of uncertainty for the test process component of the visual evaluation.

1.8 Estimation and summary of variance components

The resulting standard uncertainty can be calculated from the following equation, using the standard deviations of the previously selected and calculated factors, expressed in scale (grades).

This can be given by the following formula for colour fastness to perspiration test results:

$$U = \sqrt{U_{\text{sampling}}^2 + U_{\text{visual assessment}}^2} \quad (3)$$

Accordingly, due to the high degree of similarity between the processes of colour fastness to water and perspiration and the factors examined in this paper, their combined measurement uncertainty (for the time being) will be as follows (calculated with a 95% confidence interval):

$$U = \sqrt{0,266^2 + 0,654^2} = 0,706 \text{ grades.}$$

Based on the examination of the modules of process of colour fastness to rubbing and their uncertainty components, it can be stated that its two significant sources of uncertainty are also sampling and visual evaluation. The next step in further development is to determine the sampling uncertainty of the test and to assign a separate uncertainty value for this test based on the indicators of the R&R study.

4. SUMMARY

During the research, we aimed at the complete monitoring system of measurement uncertainties and the systematic management of changes in the measurement tests of a textile testing laboratory by creating a spreadsheet database. To this end, we summarize information about the factors of uncertainty, the risk calculated for particular test methods, the current value of uncertainty, and related improvements. Following the systematization, we present the procedure for determining the "U" value for a selected test, which was already a test of this development process. In the future, we plan to validate the method before its final implementation and then adapt it for certification testing.



5. REFERENCES

- [1] EA Laboratory Committee: JCGM 100:2008 – EVALUATION OF MEASUREMENT DATA – GUIDE TO THE EXPRESSION OF UNCERTAINTY IN MEASUREMENT, 2003
- [2] JCGM 100:2008 – EVALUATION OF MEASUREMENT DATA – GUIDE TO THE EXPRESSION OF UNCERTAINTY IN MEASUREMENT
- [3] ISO/IEC 17025:2017 – GENERAL REQUIREMENTS FOR THE COMPETENCE OF TESTING AND CALIBRATION LABORATORIES, 2017
- [4] ISO 105-E04:2013 – TEXTILES – TESTS FOR COLOUR FASTNESS – COLOUR FASTNESS TO PERSPIRATION, 2013
- [5] AIAG: MEASUREMENT SYSTEMS ANALYSIS (MSA), 4TH EDITION, 2010

6. CORRESPONDING AUTHOR:

Tibor GREGÁSZ

Institute of Media Technology and Light Industry, Rejtő Sándor Faculty of Light Industry and Environmental Engineering, Óbuda University

H-1034, Budapest, HUNGARY

E-mail: gregasz.tibor@rkk.uni-obuda.hu