XXI IMEKO World Congress “Measurement in Research and Industry”
August 30 – September 4, 2015, Prague, Czech Republic

THE METROLOGY STANDARDS AND STATISTICAL THINKING: WHAT TO DO?

Adler Yuri [1], *Shper Vladimir* [2]

[1] Prof. of Moscow Institute of Steel & Alloys, Moscow, Russia, adler.37@mail.ru
[2] Assoc. Prof. of Moscow Institute of Steel & Alloys, Moscow, Russia, vlad.shper@gmail.com

Abstract – This paper widens our suggestions made in Busan [1]. Here we suggest much more general plan of changes which – we think - are necessary for metrology to march with the times and not to become obsolete. Our suggestions refer to general basis of metrological standards and to their link with the main notions of statistical thinking introduced by such pillars of statistics as R. Fisher, W. Shewhart, E. Deming and G. Box.

Keywords: statistical thinking, standard ISO 5725, urgent changes

1. INTRODUCTION

We think that it is high time to start a broad discussion about essential widening of an area covered by metrological standards issued under the ISO patronage.

Such widening should go through four main directions:

- Inclusion of nonmetric scales into common practice of metrology;
- Refusal from the practice of using the normal distribution as the main model of error variations;
- Adoption of the notion of process stability as the main prerequisite for any measurements;
- Implementation of DOE into the most part of practical methods used for measurement error assessment.

Below we briefly consider these four paths which are part and parcel to metrology transformation in accordance with the demands of 21st century. Additionally we’d like to emphasize that movement in these directions will require the metrologists to become “team player” because practical realization of such a plan implies the joint efforts of different specialists. As a result the fifth direction of necessary changes will be the skill of teamwork.

At the end, the conclusion and a short list of references are given.

2. THE MAIN DIRECTIONS OF CHANGES

2.1. The scope of metrology: metric and nonmetric scales

Traditionally metrology limits its scope by measuring so-called physical values like time, weight, length, etc. All the more so relates to metrological standards, which limit their field with measurements of mainly materials on continuous scale [2]. But the world is quickly entering the epoch of the 3rd industrial revolution which leads to the great changes of what and from what materials will be produced in the nearest future. Besides the part of the population busy in the production of “things” is expected to fall sharply and the most part of people will be working in the areas of service. This means that the most part of data that should be measured will be gathered in the areas of not only discrete scales but scales without metrics we are used to operate. In other words we are sure that metrological standards should include procedures with all four types of scales: nominal, ordinal, interval and ratio. The theoretical basis for this is being discussed in statistical literature (see, for example, [3, 4] and the references there) so metrologists together with statisticians should start moving in this direction.

Perhaps theoretically more important reason for such change is linked with that simple fact that any measurement means the intervention into the system measured and afterwards this system inevitably changes. This is the problem that was in the center of discussion between such outstanding personalities as A. Einstein and N. Bohr in the middle of the last century [5]. And in our case this intervention is linked with a human being, and Man’s behavior is unpredictable. That’s why we need to be able to make measurements in all conceivable dimensions.

2.2. Refusal from normality assumption

Widely used normal model sometimes surely works but to hope it is a reality in the world of infinite variety is senseless. Of course one may hope that an operator working during a short interval of time with the same equipment will get normally distributed results while making replicates on the same thing but to hope to get such results in different laboratories by different operators seems to us utopian phantasy. Besides there are many situations when normality assumption knowingly does not work – for example, limited from one side features of materials such as damage threshold, or features that can’t be negative like time, or well-known highly skewed features like wages in almost all organizations, and so on. And surely the number of such cases will be increasing as a result of new material appearance (with new and often highly unexpected properties – for example, graphene, etc.).

But nonparametric estimates which may be used instead of the normal distribution lead to larger intervals of uncertainty. So the problem we are facing is the problem of the choice: in what case our losses will be bigger? There are many alternatives to parametric models – robust and Bayes estimates, bootstrap simulation and so on, and it is obvious that the universal answer does not exist – so it is necessary to widen the applications of different models and discuss and analyse the results obtained (see [3, 4, 6] and references there). The choice of alternative is principally a team work: metrologists, statisticians, engineers, managers, professionals in specific fields – all should work together. That means that all participants of such teams should be taught to work collectively and should have shared terminology.

2.3. Measurement process stability

This problem was discussed by us in [1], but we can't help saying about it again. The current version of standards ISO 5725 had already included the notion of "process stability" into the text but made it completely unsatisfactorily.

The problem we see here is not the problem of unhappy choice of some examples as we tried to demonstrate in [1]. The real problem lies much deeper than it seems at first glance. This is a problem of the lack of statistical thinking among most of metrologists. It is worth noting that the words "statistical thinking" we use here do not mean the knowledge of standard statistical tools (like analysis of means, confidence intervals, point and interval estimates and so on).

The notion of statistical thinking we are talking about is the understanding of the ideas of W. Shewhart and E. Deming about common and assignable causes of variations, the skill to analyze the stability of any process by using Shewhart control charts, and the understanding of what actions should follow after an analysis of process stability had been made [7-9].

It is worth noting here that some important industries (for example, the most well-known is automobile) are now use the so-called Measurement System Analysis (MSA) as a tool for ensuring the accuracy of their measurements and this approach is entirely based on simple Shewhart charts and on the link between standard deviation and range. Unfortunately most of metrologists do not know these issues because they are not included into their education. Philip Stein in the special publication of ASQ [10] wrote: "Yet, statistical thinking about measurement results and measurement data is far too rare. As a laboratory assessor, I have visited more than 80 calibration labs – some of them more than once. Only two of them were using control charts, and only one of those was doing it correctly". We are sure that current version of ISO 5725 standards can't change this situation. On the contrary it worsens the state of affairs because many laboratories while using the recommendations of ISO 5725 do not understand what is the goal of stability analysis according to these standards. The thing is that in the current version of standards ISO 5725 there are no answers to such obvious questions as:

Why we need measurement process stability;

What to do when our measurement process is not stable;

What not to do when our measurement process is stable;

How to calculate control limits rightly in different circumstances;

When we can use data from the previous period of time and when not.

And what is way worse is that when people from analytical laboratory meet people from the shop floor to discuss some problem of measurements in their organization they suddenly found that they do not understand each other: they speak different languages. Unfortunately the terms of ISO 5725 standards and the guide for MSA are not harmonized yet.

We are sure that the path to improve this situation goes through teaching and training all metrologists in statistical thinking and using Shewhart control charts in their everyday work. So as a first step the special teaching course with a conditional name "Statistical Process Control (SPC) in metrology" should be developed. It should be included into all programs of teaching specialists in metrology. But before this the lecturers themselves should be trained in application of SPC in their professional areas. This course should be focused on those working in metrology that's why it should be developed by a joint team of experts in SPC and metrology methodology.

2.4. Metrology and DOE

The current version of metrological standards uses the methods of DOE. But we think that the full potential of DOE has not been used yet. Instead of different parts of the standard for various conditions it seems quite possible to work out the general plan of experiment and thus to convert six standards into two: the first with general introduction and the second with general plan and its partial realizations. In fact the main factors that contribute to the variability of measurement results usually are known and not too numerous (see, for example, 0.3 in [2]). These are: operator, measurement system (which may include several factors), measurement algorithm (which may be characterized by some factors), used for measuring materials, different labs, times, and so on. For all listed factors one can create full or fractional factorial design. It seems possible to have this plan being orthogonal so before study one may eliminate the factors not necessary at current moment. Sometimes the preliminary screening experiment may turn out to be necessary but such cases won't be too often. And here we again encounter the problem of teaching and training all metrologists some additional knowledge – this time it is more deep understanding of DOE and its applications.

Additionally it is worth noting that making a plan of experiment is a necessary phase of such stages as designing equipment for future measurements, planning the algorithm of analysis and so on, i.e. DOE should be used at all stages of life cycle of measurement system if we want to reach its optimization [11].

3. CONCLUSIONS

Thus the main idea of our presentation is to facilitate the discussion about radical transformation of metrological standards. The changes we consider necessary include the widening these standards to nonmetric measurements, refusal from the normality assumption as almost ubiquitous feature of measurements, teaching and training all metrologists in statistical thinking (in accordance with W.

Shewhart and E. Deming) as well as in DOE. And all these changes should be accompanied with much more spread of skills for teamwork among the very different professionals than it takes place nowadays and is described in the current version of ISO 5725. Besides – but this is the theme for another paper – the compatibility of terms and methods between ISO 5725 standards and ISO TC 16949 guide for Measurement System Analysis should be reached as soon as possible because more and more companies throughout the world are forced to implement these systems at their shop floor simultaneously.

REFERENCES

[1] Yu. Adler, V. Shper. "Statistical Thinking and Metrology: Problems and Decisions", *XX IMEKO World Congress*, pp. 1-5, Busan, Republic of Korea, Sept. 2012.

[2] International Standard ISO 5725-1. Accuracy (trueness and precision) of measurement and results. Part 1: General principles and definitions.

[3] S. Vardeman, M.S. Hamada, T. Burr, etc., "An Introduction to Statistical Issues and Methods in Metrology for Physical Sciences and Engineering", *J. of Quality Technology,* vol. 46, nº. 1, pp. 33-62, Jan. 2014.

[4] J. de Mast, W.N. van Wieringen, "Modelling and Evaluating Repeatability and Reproducibility of Ordinal Classifications", *Technometrics*, vol. 52, nº. 1, pp. 94-106, Febr., 2010.

[5] Bohr N. *Bohr - Einstein Dialogue*. - In: Kennedy, P. J. *Niels Bohr: A Centenary Volume.* - Cambridge, Massachusetts: Harvard University Press, 1949, pp. 121–140.

[6] Yu. Adler. The Statistics: Past, Present, and Future. (A New Paradigm in Statistics). – *13th Annual ENBIS Conference*, Ankara, Turkey, Sept. 2013.

[7] Shewhart W.A. *Statistical Method from the Viewpoint of Quality Control.* – N.Y., Dover Publications, Inc., 1939 (reprint 1986).

[8] Deming W.E. *The New Economics. For Industry, Government, Education*. – Cambridge, Ma: Massachusetts Institute of Technology, 1995.

[9] Wheeler D., Chambers D. *Understanding Statistical Process Control*. SPC Press, Inc., Knoxville, 1992.

[10] Stein P. *Statistical Issues in Measurement*. – Special Publication. ASQ Statistics Division. – Fall, 2002.

[11] Adler Yu. P., Markova E.V., Granovsky Yu. V. *The Design of Experiments to Find Optimal Conditions. A Programmed Introduction to the Design of Experiments*. – Transl. from Russian by G. Leib. – Moscow: MIR Publishers, 1975.