XX IMEKO World Congress
Metrology for Green Growth
September 9–14, 2012, Busan, Republic of Korea

Statistical Thinking and Metrology: Problems and Decisions

Yu. Adler, V. Shper

Prof. of MSIS&A, Dr., Russia, adler.37@mail.ru

Abstract: It is shown that the current version of standards ISO 5725-1…6 includes insufficiently the main ideas of statistical thinking into its content. Moreover the necessary prerequisite of any measurement – stability of measurement process – looks like some kind of foreign body within the main frame of these standards. All these points are especially notable in the examples presented in the ISO 5725-6. Analysis shows that this part of 5725 standards does not respond to the basic requirement of any standard: to be an operational procedure for the best way of getting necessary results.

Keywords: measurement stability, control charts, Standards 5725, Statistical Process Control (SPC), Measurement System Analysis (MSA).

1. INTRODUCTION

Metrology and statistics are going side by side from the very beginning. They need each other. Metrology helps to get the original data about the features and characteristics of different objects. And statistics is engaged in data treatment and in the estimation of uncertainty inherent to any measurements. This union of metrology and statistics is a vital condition of transforming process data into information useful for interpretation and decision making.

Both metrology and statistics are developing. And in order to respond to the challenges of 21st century they need to take from each other its most useful achievements. We think that among them the notion of statistical thinking (ST) has the first priority. This notion was developed by Walter Shewhart at the first half of 20th century [1,2] and then it was generalized and promoted by famous quality guru Edwards Deming [3, 4].

The main idea of ST is that the behavior of any system may reliably be characterized then and only then when this system is statistically stable (or the system is under control). Otherwise any results including the results of any measurements are unreliable and cannot be reasonably interpreted. More that 80 years ago W.Shewhart gave us a very simple practical way to determine if a process is stable or not. To this end he suggested to construct an appropriate for this process picture (called now Shewhart Control Chart - SCC) and look at it. If all points of our chart lie within its limits (called control limits) the process is stable (or controlled or predictable) by definition. In other words SCC is an operational definition[1] of any process stability [3,4]. If there are points lying beyond the control limits of our chart the process is unstable (or uncontrolled or unpredictable) by definition. It is worth noting that control limits for all SCC are being calculated by the well-known simple formulas presented in any SPC book [5-8]. But there is a problem of extraordinary importance here. The thing is that the control limits for SCC can not be calculated mechanically by making simple arithmetical computations. Construction, conduction and interpretation of any SCC is something like communication with the process given: SCC is the voice of the process and only a person who knows this process thoroughly is able to understand what story the process wants to tell him. This problem is highly undervalued in SPC literature and as far as we know is being elucidated in very few books [5, 9, 10] and papers [11-13].

Now let us return to the problem of an interaction between metrology and SPC and look at the international standards ISO 5725 [14-19].

If one compares their contents with the core of ST [1-5] or with the technical specification ISO/TS 16949 [20] and its supplements [6, 21] he will see that many facets of ST are included into metrology standards. First of all we mean the appearance of such ST notions as process stability and SCC as a tool to define whether the process is stable or not. We think this is a huge leap into the right direction. But unfortunately there is a number of essential deficiencies in the current version of standards 5725, however [22]. It seems that ST has not yet become the inseparable part of metrological thinking. And as we wish this merging became the reality we are going to discuss here some examples of the lack of ST in ISO 5725. We'll limit our discussion with the standard ISO 5725-6 [19] only because the examples given there show the problems we are concerned with most clearly.

2. ANALYSIS AND RESULTS.

The first problem we'd like to raise is the problem of process stability in general. Let us look at the ISO 5725-6 where the examples of practical procedures are given with accompanying comments.

Example 1: Stability check of the repeatability standard deviation of a routine analysis (6.2.2 [19]). Here there are data from table 5 of [19] (table 1 is abridged for brevity).

[1] An operational definition is such a definition that is clear, and can be practically and unambiguously realized [3,4].

Quality characteristic is "Nickel content of private reference material".

Table 1. Extracts from table 5 of [19]

Sub-group No.	Observed values x1	Observed values x2	Range w = |x1-x2|
1	47,379	47,333	0,046
2	47,261	47,148	0,113
3	47,270	47,195	0,075
4	47,370	47,287	0,083
5	47,288	47,284	0,004
…	…	…	…
25	47,279	47,268	0,011
26	47,178	47,200	0,022
27	47,211	47,193	0,018
28	47,195	47,216	0,021
29	47,274	47,252	0,022
30	47,300	47,212	0,088

Then using these data a SCC shown in fig.1 was constructed[2]. This is a simple R-chart where we denoted a central line for the range calculated by us (see details below) as CLR; a central line for the range calculated in standard as CLR (Std); an upper control limit for range, calculated by us as UCLR; and an upper control limit for range calculated in standard as UCLR (Std).

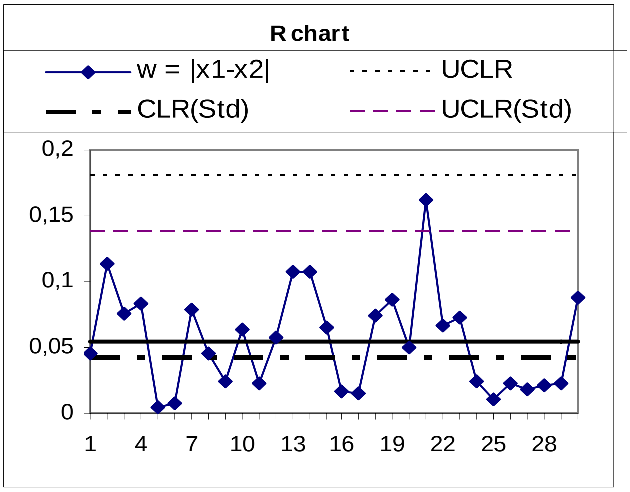

Fig.1

Both central line and upper control limit in standard were calculated using the data from the previous quarter of the year. The conclusion made in clause 6.2.2 [19]: "the test results are not stable because there is one point above the action limit...".

[2] Everywhere below we use the following notations: CL – the central line, CLR – the central line for range chart, UCL/LCL – the upper/lower control limit, UCLR – the upper control limit for range chart

Therefore…, but there is no any *therefore* in this standard.

Why?

As it is permanently being outlined in all literature about ST [1-13] the action limits are called "action" because the finding of any points beyond them should be followed by some actions, e.g. search for special causes of variation, their elimination, gathering new data, its analysis, etc. As W. Scherkenbach titled his paper "The Only Reason to Collect Data is to Take Action" [12].

Moreover if the process is unstable then any measurements are doubtful by definition,

but

There is no answer through all standards of ISO 5725 on the simple question: What should we do when the process is unstable?

This makes the goal stated in 1.1c of ISO 5725-6 – to estimate stability of measurement results and implement "quality control" into mesurement laboratories - absolutely useless.

Besides if the process is not stable why we use the previously found values for chart construction? Our chart shows both the limits from previous data (calculated in [19]) and the limits calculated by us from the current data. One can see that our limits tell us another story: the range chart shows stability if we pay no attention to six successive points below the central line (CL). But if we draw the chart for averages (see fig.2) we'll see that actually the process is not stable (the 9th point is above the UCL). But not due to range. It is not stable because of the process tuning.

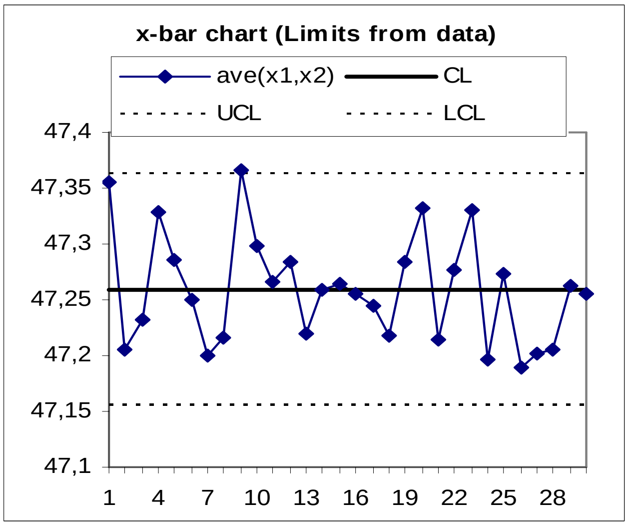

Fig.2

In the *example 2* (abridged data are shown in the table 2 below) the standard suggests the same procedure: the limits are being calculated on the base of the previous data, again only R-chart is being considered, and the conclusion is being made that the process is stable (see fig.3) because all points in fig.3 lie lower than the upper control limit for range (UCLR). But even a cursory glance at fig.3 shows that not all is good with these data. Firstly we'd like to draw a reader attention to the number of different values of range lying below the upper control limit on the range chart. This number is called the number of distinctive categories (*ndc*) and it should be more than 5 in order to consider the measurement system acceptable [9, 21]. Here we have that *ndc* just equals to 5 so the system is on the edge of acceptability. It is worth noting that the notion of the number of distinctive categories is not being discussed in ISO 5725 at all.

Table 2. Extracts from table 6 of [19]

Subgroup No.	Observed values x1	Observed values x2	Range w = |x1-x2|
1	0,56	0,56	0
2	0,48	0,5	0,02
3	0,57	0,58	0,01
4	0,6	0,58	0,02
5	0,58	0,58	0
…	…	…	…
25	0,46	0,45	0,01
26	0,6	0,58	0,02
27	0,59	0,56	0,03
28	0,54	0,56	0,02
29	0,47	0,49	0,02
30	0,59	0,58	0,01
31	0,49	0,52	0,03

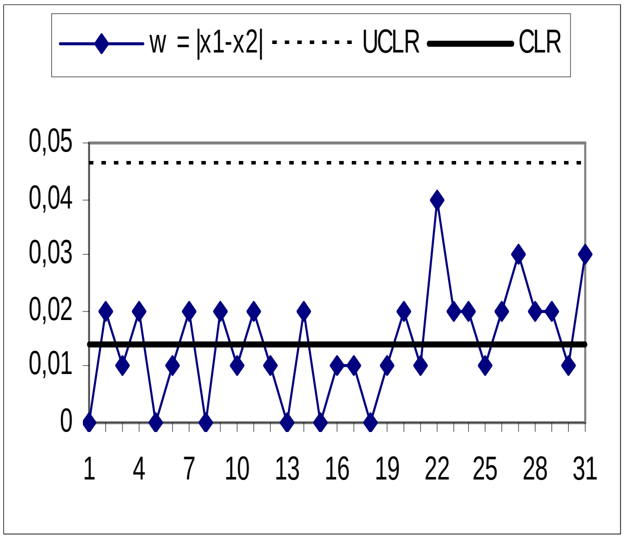

Fig.3

But if one depicts the chart for averages using the same data (see fig.4) he will face quite different situation. Again we calculated the limits by using the current data but in this case the difference between them and limits from the previous data is rather small. We think that the process shown in fig.4 can't be named stable. The data consist of two groups: 9 points lying below 0,5 and 22 points lying above 0,5 with almost regular interchange between them. In fact this is the issue we mentioned in introduction: before counting anything the process owner should depict the run chart for his process and look at it very attentively in order to decide: whether the data are taken from one population or not [11]. Or in other words: whether we can combine all data into one chart or not? If not then he has to decide how many charts he needs (i.e. how he will divide the data into different strata).

Fig.5 reveals the whole story by using *x-mR* chart: the process is not stable and demonstrates the presence of special causes of variation both on *x*-chart and *mR*-chart. And again the lack of homogeneity is quite obvious therefore the chart discussed in this example is senseless.

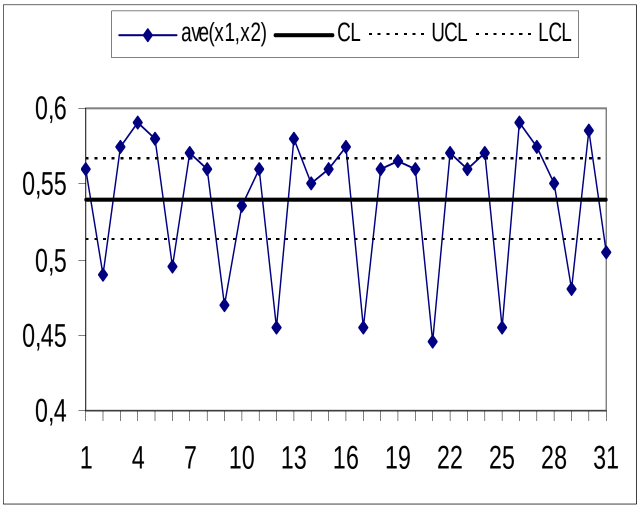

Fig.4

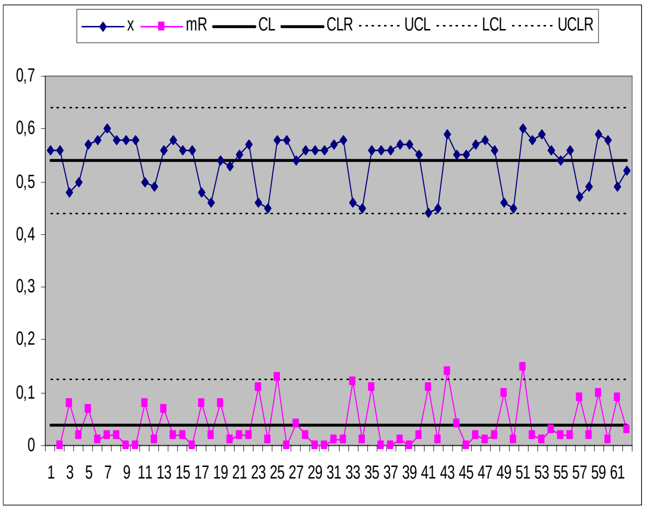

Fig.5

Here we stop discussing examples from [19] because the new ones do not add anything principally new.

The problem we see here is not the problem of unlucky choice of some examples as someone could say. This is the problem of gaining new knowledge and experience and it lies much deeper than it seems at first glance.

The thing is that ST is not a tool or technique – this is the way of thinking, and this new way of thinking can't be implemented by a number of examples discussed in a standard. Moreover this is not the problem of only metrologists. E. Deming in his Foreword to the new edition of Shewhart's book of 1939 wrote: "Another half-century may pass before the full spectrum of Dr. Shewhart's contribution had been revealed in liberal education, science, and industry" [2, Foreword]. Philip Stein in the special publication of ASQ [23] 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 will not change this situation. On the contrary it can get much worser because many laboratories will be using the recommendations of 5725-6 mechanically; they will be doing senseless calc's, writing reports about stability of their processes; and all this heap of work will be producing waste.

3. SUMMARY

It seems to us that statistical thinking has made only its first step into metrology area: the problem of process stability had been designated and the key words – Shewhart control charts, statistical stability of a process, control limits, etc. – appeared in the text of ISO 5725. But from practical point of view this is absolutely insufficient.

In current version of standards ISO 5725 there are no clear answers on many questions, namely:

- 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;
- How to construct cumulative control chart.

Besides the examples should be carefully analyzed from different points of view and should be worded operationally.

Another important issue, which we had not touched upon here, is the compatibility of terms and methods in ISO 5725 and ISO/TS 16949 manual on measurement system analysis [21].

But from the system point of view the main question is: what should we do to change this situation in the right direction?

Our answer is both very simple and very difficult in realization. It is necessary to teach and train all metrologists how to use ST and apply SPC in their everyday life. So as a first step the special teaching course with conditional name "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 team of experts in SPC and metrology methodology.

And simultaneously of course it is absolutely necessary to issue a new version of ISO 5725 standard in order to make it useful for purposes announced in it and for the real friendship between statistics and metrology. To this end this new version should be supplied with very detailed supplement about using SPC methods in metrology issues.

4. REFERENCES

[1] **Shewhart W.A.** Economic Control of Quality of Manufactured Product. – 1980, ASQ, Milwaukee.

[2] **Shewhart W.A.** Statistical Method from the Viewpoint of Quality Control. – N.Y., Dover Publications, Inc., 1939 (reprint 1986). – 163 p.

[3] **Deming W.E.** Out of the Crisis. – 1986, Cambridge, Massachusetts Institute of Technology.

[4] **Deming W.E**. The New Economics. For Industry, Government, Education Second Edition, 1995. – Cambridge, Ma: Massachusetts Institute of Technology, 1993. – 247 p.

[5] **Wheeler D., Chambers D.** Understanding Statistical Process Control, 2nd Ed. – 1992, SPC Press, Inc., Knoxville.

[6] Statistical Process Control Reference Manual (SPC). – Chrysler, Ford, GM. - 1995.

[7] **Montgomery D.** Introduction to Statistical Quality Control, 6th Ed. – 2009, John Wiley & Sons, Inc.

[8] **Duncan A.J.** Quality Control and Industrial Statistics, 5th Ed. – 1986, Irwin, Homewood, Il.

[9] **Wheeler D.J.** Advanced Topics in Statistical Process Control. – 1995, SPC Press, Inc., Knoxville.

[10] **Balestracci D.** Data Sanity: A Quantum Leap to Unprecedented Results. – 2009, Medical Group Management Association.

[11] **Wheeler D.** The Four Questions of Data Analysis. – www.qualitydigest.com/inside/quality-insider-column/four-questions-data-analysis.html

[12] **Scherkenbach W**. The Only reason to Collect Data is to Take Action! – www.spcpress.com/pdf/only_reason_to_collect.pdf

[13] **Neave H.** A Study of Statistical Process Control. – www.spcpress.com/ink_pdfs/study%20of%20SPC.pdf

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

[15] International Standard ISO 5725-2. Accuracy (trueness and precision) of measurement and results. Part 2: Basic method for the determination of repeatability reproducibility of a standard measurement method.

[16] International Standard ISO 5725-3. Accuracy (trueness and precision) of measurement and results. Part 3: Intermediate measures of the precision of a standard measurement method.

[17] International Standard ISO 5725-4. Accuracy (trueness and precision) of measurement and results. Part 4: Basic method for the determination of the trueness of a standard measurement method.

[18] International Standard ISO 5725-5. Accuracy (trueness and precision) of measurement and results. Part 5: Alternative methods for the determination of the precision of a standard measurement method.

[19] International Standard ISO 5725-6. Accuracy (trueness and precision) of measurement and results. Part 6: Use in practice of accuracy values.
[20] Technical Specification ISO/TS 16949. Quality management systems - Particular requirements for the application of ISO 9001:2000 for automotive production and relevant service part organizations.
[21] Measurement System Analysis. Reference Manual. Third Edition. – Chrysler, Ford, GM. - 2002.
[22] **Adler Yu., Shper V.** Statistical Thinking and Metrology: Enemies or Friends?! – In: Abstracts. –The 17th International Conference of the Israel Society for Quality. – Jerusalem, Israel, November 18-20, 2008. – P. 58.
[23] **Stein P.** Statistical Issues in Measurement. – Special Publication. ASQ Statistics Division. – Fall, 2002.