ANNALS of Faculty Engineering Hunedoara – International Journal of Engineering
ISSN: 1584 - 2665

THE QUALITY OF VICKERS HARDNESS TESTER CALIBRATION

Jozef PETRÍK, Pavol PALFY

Technical University of Košice, Faculty of Metallurgy,
Department of Integrated Management, Košice, SLOVAKIA

ABSTRACT:
The objective of submitted work is to analyze the influence of the hardness non-uniformity of the used CRM and appraisers doing calibration on its outcome, i.e. conformance or non-conformance of the Vickers hardness tester (HV30) by uncertainty analysis. The results were validated by measurement systems analysis (MSA), unpaired t-test, ANOVA and Z-score. Repeated calibration carried out by two appraisers leads to the identification of tester non-conformance. The statistically significant effect of the CRM non-uniformity and appraisers on the tester non-conformance was proved. According to MSA the studied process is non-capable.
KEYWORDS:
calibration, hardness HV30, uncertainty, MSA

1. INTRODUCTIONS

The Vickers test is the standard method for measuring the hardness of metals, particularly those with extremely hard surfaces: the surface is subjected to a standard pressure for a standard length of time by means of a pyramid-shaped diamond with vertex angle 136° [1, pp. 319]. Like in any test of mechanical properties, there is obvious requirement for reliability of measurement results, which is unthinkable without sufficient quality of measurement process. Metrological confirmation (calibration and verification) shall be designed and implemented to ensure that the metrological characteristics of the measuring equipment satisfy the metrological requirements for the measurement process [2].

Calibration is checking of a measuring instrument against an accurate CRM (certified reference material, standard) to determine any deviation and correct for errors [3]. A calibration laboratory, or a testing laboratory performing its own calibrations, shall have and shall apply a procedure to estimate the uncertainty of measurement for all calibrations and types of calibrations. When estimating the uncertainty of measurement, all uncertainty components which are of importance in the given situation, shall be taken into account using appropriate methods of analysis [4].

The objective of this article is to analyze the influence of possible CRM's hardness non-uniformity and influence of appraisers on conformance or non-conformance of tester by analysis of calibration uncertainty. The results were validated by measurement systems analysis (MSA), unpaired t-test, analysis of variance (ANOVA) and Z-score.

2. EXPERIMENTAL

The calibration was carried out by two approximately equally skilled appraisers (A, B). The measurement points were along the diameter (rim to rim) of the CRM in equidistant intervals. Appraiser A performed a calibration (5 indentations) followed by appraiser B. The indentations of both appraisers were evenly distributed around the center of the filed of view complying with the standard's requirement for the minimal spacing between the adjacent indentations [8]. The force application time was 10 seconds. The values of average hardness $\overline{H}$ and standard deviation s_H are in tab. 1 and 2.

Calibrated tester HPO 250 was made by VEB Werkstoffprüfmaschinen „Fritz Heckert" (East Germany) in 1982. The magnification of measuring device is 140×. The hardness tester is not a legal measuring equipment (Slovak regulation No. 210/2000 Z.z.), therefore its metrological confirmation is limited to calibration [5]. The test force/load F = 294.2 N (30 kg). The ambient temperature was 21.0°C, relative humidity 56.5%. According the direct calibration (VI/08) the test force deviation is -0.2% and the measurement device deviation is +0.1% (diagonal 0.4 mm). The certified reference material (CRM) in form of hardness reference block with specified hardness H_c = 472.4 HV30 and standard uncertainty u_{CRM} = ± 4.724 HV30 was used as a standard.

Table 1. Results of calibration, appraiser A

Calibration No.	$\overline{H}$	s_H	r_{rel}	E_{rel}	u_H	u_{ms}	u_{HTM}	U_{HTM}	U_{rel}
1	498	3.16	0.73	1.55	1.625	0.705	5.04	10.08	3.63
2	484	4.46	1.17	2.45	2.296	0.715	5.30	10.60	4.58
3	480	1.55	0.81	1.55	1.565	0.705	5.02	10.04	3.62
4	482	3.58	0.96	1.97	1.843	0.710	5.12	10.23	4.05
5	481	5.75	1.54	1.73	2.957	0.707	5.61	11.23	4.04
6	477	3.41	0.95	1.07	1.754	0.700	5.08	10.17	3.19
7	475	2.80	0.73	0.45	1.440	0.694	4.98	9.97	2.55
8	482	7.77	0.84	2.01	3.997	0.710	6.23	12.45	4.55
9	476	1.34	0.36	0.66	0.687	0.696	4.82	9.64	2.68
10	471	3.37	0.80	0.30	1.731	0.686	5.07	10.15	2.46
Normality test	p: 0.150576								

Table 2. Results of calibration, appraiser B

Calibration No.	$\overline{H}$	s_H	r_{rel}	E_{rel}	u_H	u_{ms}	u_{HTM}	U_{HTM}	U_{rel}
1	484	2.23	0.52	2.51	1.145	0.715	4.91	9.82	4.58
2	490	7.57	1.78	3.62	3.896	0.727	6.16	12.33	6.01
3	485	1.55	0.37	2.75	0.799	0.718	4.84	9.68	4.67
4	484	1.48	0.37	2.54	0.761	0.715	4.83	9.67	4.47
5	484	5.70	1.62	2.42	2.932	0.714	5.60	11.20	4.68
6	481	4.87	1.17	1.82	2.506	0.708	5.39	10.78	4.03
7	484	3.70	0.88	2.42	1.902	0.714	5.14	10.28	4.49
8	481	3.65	0.96	1.91	1.876	0.709	5.13	10.26	4.00
9	479	0.92	0.22	1.46	0.472	0.704	4.80	9.59	3.44
10	478	2.24	0.59	0.30	1.153	0.702	4.91	9.82	3.29
Normality test	p: 0.017882								

The first step of analysis is to estimate whether the discrimination (effective resolution) d*- the value (HV) of the smallest scale division (graduation) is sufficient. A general rule of thumb is that the discrimination ought to be at least one tenth the process variation (standard deviation s_H in tab. 1 and 2). Because d* = 1.44 HV, used tester does not satisfy this condition [6, p. 44, 74].

Grubbs' test (with significance level α= 0.05) detected one outlier (appraiser B, calibration No. 2). The statistical outliers indicate that the process is out of statistical control.

The normality was estimated by Freeware Process Capability Calculator (3.0.0) software, using Anderson – Darling test. The file of appraiser A has normal distribution, but normality of appraiser's B file was not confirmed. The MSA assume normal probability distribution. If normality of the file is not confirmed, the measurement system error is overestimated [6, p. 48].

3. THE CALCULATION OF CALIBRATION UNCERTAINTY

The repeatability of tester is the mean value, d_5 is the maximum and d_1 is minimum value of indentations diagonals. The repeatability r_{rel} of satisfactory tester is less than 2% for used CRM.

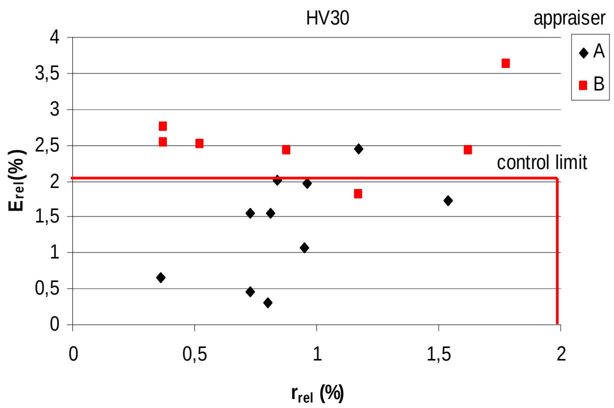

Fig. 1. The values of r_{rel} and E_{rel} for both appraisers

$$r_{rel} = 100 \times \frac{d_5 - d_1}{\overline{d}} \% \quad (1)$$

The error at specific conditions of calibration $\overline{H}$ is the average hardness of CRM, measured at calibration

$$E = \overline{H} - H_c \quad (2)$$

Relative maximum error

$$E_{rel} = 100 \times \frac{\overline{H} - H_c}{H_c} \% \quad (3)$$

The relative maximum error E_{rel} of satisfactory tester is less than 2 % for used CRM [5]. The values of r_{rel} , s_H and E_{rel} are on fig. 1 and tab. 1,2.

The present calculations supposed, that the result of calibration equals to its ideal value. But, the uncertainty is a parameter associated with the result of a measurement, characterizing the dispersion of the values that could reasonably be attributed to the measured value [7]. This fact must be regarded. The uncertainty of indirect calibration:

$$u_{HTM} = \sqrt{u_{CRM}^2 + u_{CRM-D}^2 + u_H^2 + u_{ms}^2} \quad (4)$$

© copyright FACULTY of ENGINEERING - HUNEDOARA, ROMANIA

The standard uncertainty of used CRM u_{CRM} = 4.724 HV30. The uncertainty resulted drift of CRM u_{CRM-D} was ignored (used CRM was calibrated only once, XI/2005). Standard uncertainty of hardness tester s_H is standard deviation of the results of calibration, Student's factor t = 1.15 for n = 5 (number of trials in one calibration) and significance level α = 0.317 [5]. Another source of uncertainty is measuring device. The sensitivity of indentations measuring device δ_{ms} = 0.0005 mm.

$$u_H = \frac{t \times s_H}{\sqrt{n}} \qquad (5)$$

$$u_{ms} = \frac{2\overline{H}}{d}\,\frac{\delta_{ms}}{2\sqrt{3}} \quad \text{[8, step 7 M2]} \qquad (6)$$

$$d = \sqrt{0.1891 \times \frac{F}{\overline{H}}} \qquad (7)$$

The error of calibration

$$\overline{b} = \overline{H} - H_c = E \qquad (8)$$

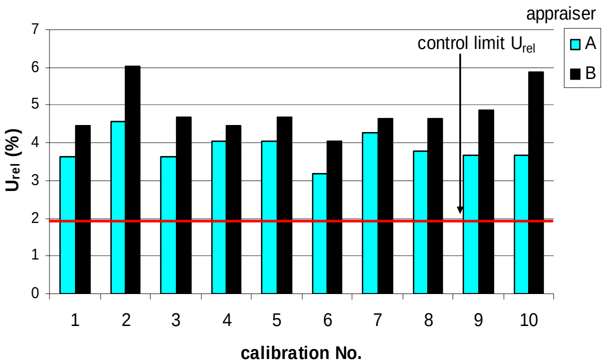

Fig. 2.: The values of U_{rel} for both appraisers

The maximum permissible error of the tester including the measurement expanded uncertainty U_{HTM}:

$$U_{HTM} = k \times u_{HTM} \qquad (9)$$

$$\Delta H_{HTM\max} = U_{HTM} + \left|\overline{b}\right| \qquad (10)$$

The coverage factor k = 2. Relative maximum permissible error of the tester (relative expanded uncertainty):

$$U_{rel} = \frac{\Delta H_{HTM\max}}{\overline{H}} \times 100\% \qquad (11)$$

The relative maximum permissible error U_{rel} of satisfactory tester is less than ± 2 % for used CRM. The values of U_{rel} for individual calibrations and appraisers are on fig. 2. It is possible that high value of uncertainty of calibration is a result of low capability [9] and low resolution of the tester.

The results of individual calibrations by individual appraisers were evaluated by t-test which compares the mean hardness of two groups (values of hardness measured by appraisers A and B on the same calibration's place). As can be seen in tab. 3, there were extremely statistically significant differences between results of appraisers A and B at calibrations No. 1, 9 and statistically significant at calibrations No. 3, 7 and 10.

The r_{rel} is satisfied for all calibrations. According to two factor analysis (ANOVA) without replication the influences of appraisers (p=0.718895) is not statistically significant and it of calibration's place - used place of CRM (p=0.010187) is statistically significant on r_{rel}.

As far as E_{rel}, the tester is satisfactory for half of calibrations. The influences of appraisers (p=0.000446) and calibration's place (p=0.006103) are both statistically significant.

The tester is not satisfactory for all calibrations and all appraisers with respect to its U_{rel}. The influence of appraiser (p=0.003039) and calibration's place (p=0.014634) on the uncertainty are both statistically significant. The two factor analysis with replication was used for evaluation of hardness values measured on the CRM. The influences of appraiser (p = 4.53 E-7) and calibration's place (p = 2.82 E-7) are both statistically significant [10, p. 129].

Table 3. The results of unpaired t-test

calibration no.	1	2	3	4	5	6	7	8	9	10
p value	0.0001	0.1980	0.0059	0.1593	0.3943	0.1383	0.0021	0.9028	0.0008	0.0036
difference	5	1	4	1	1	1	4	1	5	4

Note: 1. by conventional criteria, this difference is considered not to be statistically significant, 4. very statistically significant and 5 extremely statistically significant

4. THE CALCULATION OF PROCESS CAPABILITY BY MSA

The capability of calibration process can be evaluated by analysis of measurement system (MSA – Measurement Systems Analysis), which helps to conform with ISO/TS 16 949:2002 requirements. If the analyzed measurement system (equipment, appraisers, methods, software, environmental conditions...) is capable, it is likely that the measurement process, taking place in it, is capable as well.

The GRR, one of MSA methods is an experimental and mathematical method of determining measurement repeatability and reproducibility. The computation of capability indices was carried out according to [6, p. 102-120]. The software Palstat CAQ with confidence level α=0.01 (5.15σ) was used for calculation.

The measurement system ought to be under statistical control before capability is assessed, the range (R) control chart is used. The process is under the control, if all ranges are between control limits. This condition was not satisfied for appraiser A, calibration No. 6 (detected outlier!). The number of distinct categories ("ndc", based on Wheeler's discrimination ratio) is connected with

© copyright FACULTY of ENGINEERING - HUNEDOARA, ROMANIA

question of the resolution of measurement equipment. The “ndc” is greater than or equal to 5 for capable processes, the calculated “ndc” value 1.087 is unsatisfactory.

The area within the control limits of the X-bar control chart represents measurement sensitivity (“noise”), one half or more of the averages should fall outside the control limits. The condition of sensitivity was not satisfied, 70% of measurements were between control limits.

The %EV index represents the cumulative influence of measurement equipment, used measuring method and environmental conditions on the variability. It is a function of average range of trials of all appraisers. High value of this index, 61.37% is the result of aforementioned low resolution and high U_{rel} of the tester. %PV index is a function of range of average hardness of individual calibrations. Because his value of %PV is between 50% and 70% (61.06%) used tester is “inaccurate” [11, p. 29]. But, on the other hand, low value of %PV index confirms high homogeneity of used CRM. %AV index represents the influence of appraisers on variability, for example their competence, perceptions, skills disciplines and vigilance. It is a function of average values from individual appraisers. Relatively high value of %AV, 50.06%, means similar different quality of both appraisers work. Analyzed process is not capable, the value of %GRR is 79.20% (above 10% limit). This low capability is a consequence of aforementioned effects.

5. THE CALCULATION OF Z-SCORE

Z-score method is routinely applied in inter-laboratory comparisons. The value for individual calibration is:

$$z_i = \frac{x_i - \bar{x}}{s} \qquad (12)$$

x_i is the average hardness of one calibration, $\bar{x}$ the average hardness of all calibrations and „s“ is standard deviation of all calibrations. The results $|z_i| \leq 2$ are satisfactory and $|z_i| \geq 3$ are unsatisfactory [12, p. 217]. The differences between z-score results obtained for both appraisers and all calibrations are not significant (fig. 3). All results are satisfactory.

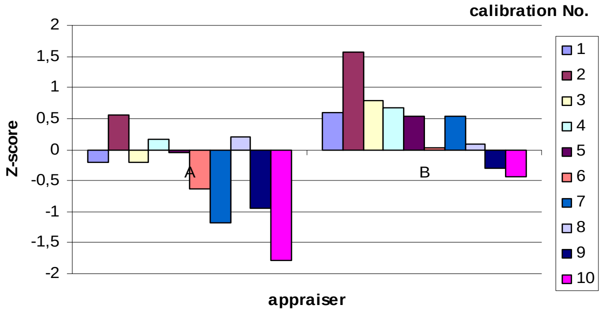

Fig. 3. Z-score

6. CONCLUSION

- The hardness tester is non-conforming for all repeated calibrations carried out by two appraisers on the same CRM.
- Statistically significant effect of the place of CRM used for calibration and appraisers on E_{rel} and U_{rel} was obtained.
- According to measurement systems analysis the studied process is non-capable, which corresponds with the first conclusion.
- With regard to the identified insufficient resolution it is recommended to use larger magnification for the calibration.

REFERENCES / BIBLIOGRAPHY

[1.] *Veles, P.: Mechanické vlastnosti a skúšanie kovov. SNTL Praha/Bratislava 1985.*
[2.] *ISO 10 012:2003 Measurement management systems - Requirements for measurement processes and measuring equipment.*
[3.] *[online]. [cited 03 July 2008]. <http://encarta.msn.com/dictionary_1861594147/calibration.html*
[4.] *ISO/IEC 17 025:2005 General requirements for the competence of testing and calibration laboratories.*
[5.] *STN EN ISO 6507-2:2005 Metallic materials. Vickers hardness test. Part 2: Verification and calibration of testing machines.*
[6.] *Measurement system s analysis (MSA). Reference manual. Third edition. 2003.*
[7.] *International Vocabulary of Basic and General Terms in Metrology VIM. [online]. [cited 14 March 2008]. Available from: <http://www.ntmdt.ru/download/vim.pdf>.*
[8.] *STN EN ISO 6507-1:2005 Metallic materials. Vickers hardness test. Part 1: Test method.*
[9.] *Tobolski, E.: Uncertainty in Hardness testing. (part 2). Advanced materials & processes, Vol. 161, No. 5 (2003), pp. 25.*
[10.] *Chajdiak, J.: Štatistika v exceli. Statis Bratislava 2002. ISBN 80-85659-27. [11] Meloun, M. – Militký, J.: Kompendium statistického zpracování dat. 1. vydanie. Praha: Academia, 2002. ISBN 80-200-1008-4.*
[11.] *Palenčár, R. et al.: Štatistické metódy v metrologických a skúšobných laboratóriách. Bratislava. 2001, ISBN 80-968449-3-8.*

ACKNOWLEDGEMENTS

This work was supported by the Slovak Grant Agency for Science VEGA 1/4141/07.

© copyright FACULTY of ENGINEERING - HUNEDOARA, ROMANIA