What you should know about the accuracy and repeatability of portable hardness testing probes

Article highlights

- Accuracy for portable hardness testing devices without a definition is an "inaccurate term" and shall always be replaced by the coefficient of variation and measurement deviation.
- Coefficient of variation (also known as repeatability or precision) and measurement deviation (also known as accuracy) shall be used together to better describe the device parameters as demanded by the most rigorous standards.
- ASTM A1038 and ASTM A956 standards control only measurement deviation and do not monitor repeatability.
- DIN 50159 and GB/T 34205 (UCI devices) and DIN EN ISO 16859 are the most rigorous standards that ensure both: accuracy and repeatability.
- The best practice is to use portable hardness testing probes calibrated against precision-demanding standards to make sure that the measurements are not only accurate but repeatable (precise).
- Customers, who choose devices calibrated against repeatability (precision) and not only against accuracy can offer much higher product and service quality to internal and external customers, being ahead of everyone else using only accuracy-based standards

After reading this 3-minute-long article you will understand the difference between accuracy and precision, why they are essential and how to make the best choice when choosing your equipment.

Introduction

Many producers provide technical datasheets and talk about the portable testing device's accuracy and how accurate the probes are, but this creates even greater confusion among the users when it comes to the technique itself.

So why is the probe's accuracy an inaccurate statement? The probe accuracy refers to “how accurate the technology and its components” are but does not define what multi-point accuracy the probe can deliver.

For UCI, for example, even more critical is that UCI method is typically executed with a handheld device and the operator's experience or handling contributes to the absolute values.

For the portable hardness testing method, the two parameters that directly deal with the probe's performance are of much greater significance for the user: **measurement deviation (also referred to by standards as accuracy)** and **coefficient of variation (repeatability or precision)**. Both of them are used to calibrate devices complying to most rigorous standards (DIN & GB/T for UCI and ISO for the Leeb method).

Measurement deviation (accuracy) and coefficient of variation (repeatability)

How are these two parameters described and what do they mean?

According to all major standards defining portable hardness testing: ASTM A956, ASTM A1038, DIN 50159, DIN 50157, ASTM A1038, GB/T 34205 and ASTM E3246 the **measurement deviation** (accuracy) is defined as follows:

$$E = \frac{\bar{H} - H}{H} \cdot 100\ \%$$

(E – measurement deviation, – Average value of n measurements, H – reference value i.e. test block)

In other words: it describes how the average value deviates from the reference value on % scale and is also strongly correlated with the quality of the reference and the calibration.

The coefficient of variation of UCI device (repeatability) is defined in **DIN 50159 and GB/T 34205** and **DIN EN ISO 16859** and describes the relative difference between the highest and the lowest hardness value concerning the average:

$$r = \frac{H_{max} - H_{min}}{\bar{H}} \cdot 100\ \%$$

(r-repeatability, H_{min} and H_{max} – the lowest and the highest hardness values respectively, - average value)

In other words: it describes how far the measurement values are scattered from each other. Repeatability is mainly dependent on the quality of the instrument and is sometimes used interchangeably as the precision of the device.

Why is repeatability important?

To illustrate better the meaning of accuracy and repeatability we use a simple target.

Imagine, that you are shooting against a target with a specific amount of bullets and your final score is computed as the average of those shots. So, depending if you are hitting close to the centre of the target or hitting random spots of the target, your accuracy and repeatability will differ.

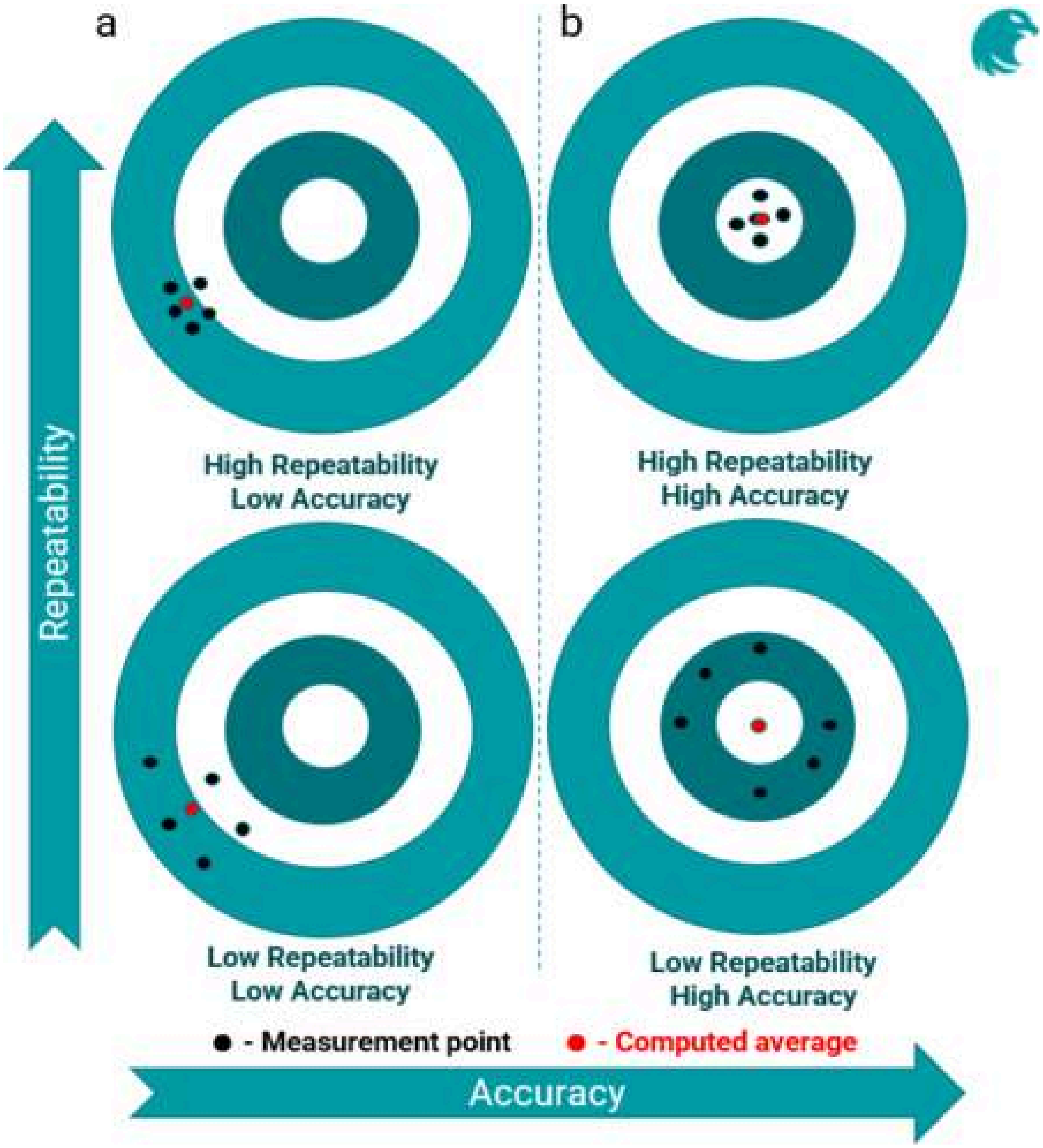
If your final average is close to the centre of the target, your shoots are accurate and if they at the same time are close to each other they are also precise. On the contrary, if you are hitting only the border of the target, your average value will differ a lot from the true value represented by the center - your shooting is inaccurate, and if your single shoot values are away from each other, then they are also imprecise.

Note, that you can also be inaccurate but precise. - a situation, in which your shoot constantly the border of the target, but your values are very close to each other. In this case one needs to expect a systematic error.

The same analogy can be transferred to portable hardness testing or in fact any measurement. Typically several measurements are executed to compute the average - the true value compared against the reference test block.

In the below example, the four possible measurement outcomes are compared and classified into two columns **a** - low accuracy and **b** high accuracy. For both cases **a** and **b** the red point indicates the computer average **which is identical** (for a column and b column respectively).

A high accuracy but low repeatability indicates that measurements must be carried in a higher population of measurement points to compute the average value, as the single data points are widely spread.



This is a problem for many hardness testing applications, for example, Heat Affected Zones (HAZ) where a weld profile is inspected by collecting a hardness profile of a weld, that consists of single measurements. In this particular case, the readings may be distorted to the extent that the border between the affected and unaffected zone is not easy to spot or blurred. Furthermore, calibrations of the devices are done in laboratories with high precision in a very controlled environment, minimizing the user's influence on the measurement - users use those in the field on non-ideal surfaces and not always perpendicularly to the tested surface, which is crucial. And hence accurate but not repeatable devices add unnecessarily deviation into the quality and reliability of the data.

What are the limits used by manufacturers and why do those values depend on hardness and test force?

Let the following table be a guideline concerning the maximum tolerable measurement deviation and repeatability. Please note that these values are used for the device's calibration by the manufacturers and not as a basis for the daily verification conducted by the end-user.

Scale / Range	Max. measurement deviation (E) in % DIN 50159, ASTM A1038, and GB/T 34205								Repeatability (R) / %			
	DIN & GB/T	ASTM	DIN & GB/T	ASTM	DIN & GB/T	ASTM	DIN & GB/T	ASTM	DIN & GB/T	ASTM	DIN & GB/T	ASTM
	<250 HV		250 -500 HV		500 – 800 HV		>800 HV		≤ 250 HV		> 250 HV	
HV 0.1	5	6	6	7	7	8	8	9	8	Not required	6	Not required
HV 0.3	5	6	6	7	7	8	8	9	8		6	
HV 0.8	4	6	4	7	5	8	6	9	8		6	
HV 1	4	5	4	5	5	7	6	7	8		6	
HV 5	4	5	4	5	4	7	4	7	5		5	
HV 10	4	5	4	5	4	7	4	7	5		5	

Table 1. The summary of maximum tolerable errors for measurement deviation and repeatability from DIN 50157-2, ASTM A1038 and GB/T 34205

Probe type	Max. measurement deviation (E) and repeatability values in DIN EN ISO 16859-2 and ASTM A956 depending on the probe type and hardness level					
	ISO	ASTM	ISO	ASTM	ISO	ASTM
D, DC	<500 HLx		500-700 HLx		>700 HLx	
DL, S	<700 HLx		700-850 HLx		>850 HLx	
C, E	<600 HLx		600-750 HLx		>750 HLx	
G	<450 HLx		450-600 HLx		>600 HLx	
Max. mmeasurement deviation (E)	4%	± 6 HLx	3%	± 6 HLx	2%	± 6 HLx
Max. rrepeatability (R) / %	2.5%	Not required	2%	Not required	1.5%	Not required

Table 2. Summary of maximum tolerable errors for measurement deviation and repeatability DIN EN ISO 16859-2 and ASTM A956. Note: x represents D, DC, DL, S, C, E, G for the respective probe.

What is the best practice?

This article shows how accuracy and repeatability are computed and how important repeatability is for the end users. It is important to highlight is that ASTM standard does not demand repeatability during the calibration process (see table 1 and table 2), hence the users can not avoid purchasing accurate but not repeatable instrument devices.

It is always advisable to use devices that are also controlled against repeatability, which is demanded by German DIN 50159, Chinese GB/T 34205 and international DIN EN ISO 16859 standards. By using devices that are compliant with all three standards, the end users make sure that their equipment is best-in-class not only by means of accuracy but also in repeatability and foremost reliability of harvested data.

Note: This document shows only a fraction of the information described in ASTM A956, ASTM A1038, DIN 50159, DIN 50157, GB/T 34205 and ASTM E3246. Screening Eagle Technologies has done everything in its power to translate the sections of the DIN 50159, DIN 50157 and GB/T 34205-2017 standards accurately. For authorized translations or more information, interested readers are encouraged to read the full version of standards DIN, ASTM A1038, GB/T 34205 and ISO available at www.beuth.de, www.astm.orgorwww.spc.org.cn and www.iso.org respectively.

References:

Metallische Werkstoffe – Härteprüfung nach dem UCI-Verfahren – Teil 2: Prüfung und Kalibrierung der Härteprüfgeräte, DIN 50159-2:2015-01, 2015

Standard Test Method for Portable Hardness Testing by the Ultrasonic Contact Impedance Method, ASTM A1038-19, 2019

Metallic materials – Hardness testing – Ultrasonic contact impedance method, GB/T 34205-2017, 2017

Metallic materials — Leeb hardness test — Part 1: Test method, DIN EN ISO16859-1

Metallic materials — Leeb hardness test — Part 2: Verification and calibration of the testing devices, DIN EN ISO16859-2

Portable Hardness Testing. Theory practice, Applications, guidelines. Burnat, D., Raj, L., Frank, S., Ott, T. Schwerzenbach, Screening Eagle Technologies AG, 2022.