forum acusticum 2026
8-12 September · Graz, Austria

ADVANCING RUBBER HARDNESS MEASUREMENT PROCEDURES FOR CPX REFERENCE TYRES: INSIGHTS FROM ISO/TS 11819-3 REVISION

Anneleen Bergiers[1], Andreas Fuchs[2], Jonathan Nowak[2]

[1] *R&D Sustainable Road Construction, Belgian Road Research Centre (BRRC), Belgium*
[2] *AIT Austrian Institute of Technology GmbH, Austria*

Abstract

Within the framework of working group ISO/TC 43/SC 1/WG33, which is dealing with measuring methods for comparing traffic noise on different road surfaces, the revision of ISO/TS 11819-3 is ongoing since 2025. This technical specification concerns reference tyres used in CPX (Close-Proximity) noise measurements. One of the remarks received during the systematic review of this technical specification was to improve the procedure to measure rubber hardness of the reference tyres, which is used as correction factor for the measured sound pressure levels. No distinction is currently made between digital and analogue durometers, although some users report different results depending on the device used. A measurement plan for a round robin test was set up where various institutes were requested to repeat measurements on their own P1 tyre by varying only one of the following parameters: durometer type, time setting after contact (for digital durometers), operator (including their level of expertise), measurement day. Participants were asked to measure more points around the tyre and to report all their results, including values that would be discarded according to the current procedure. A substantial dataset from seven institutes was obtained and analysed. Based on this analysis, a proposal for improving the measurement procedure has been put forward. It was determined that the operator's level of expertise and the device settings are significant parameters. The results will be used as input for the revision of the Technical Specification with the aim of reducing variability and increasing reproducibility of the CPX reference tyre hardness measurements.

Corresponding author: *a.bergiers@brrc.be*

Keywords: *CPX, rubber hardness, reference tyres, measurement reproducibility*

1. Introduction

The CPX method, described in ISO 11819-2 [1], is used for evaluating the acoustical performance of road surfaces. Tyre/road noise is measured by driving on the road with a measuring trailer or a self-powered vehicle using dedicated reference tyres, which are described in ISO/TS 11819-3 [2]. Rubber hardness of the tyres is monitored because it increases over time and it affects the CPX noise levels. Therefore, a correction of 0.12 and 0.2 dB/Shore A is applied for correcting the CPX noise levels to a reference tyre rubber hardness of 66 Shore A, for P1 tyre and H1 tyre respectively.

Annex A of ISO/TS 11819-3 describes the current procedure for measuring rubber hardness. A type A durometer is used to measure four equi-spaced points around the outer rib and around the inner centre rib of the tyre tread. Four consecutive readings at one point shall not differ by more than two units. The overall hardness is obtained from these eight positions.

During the systematic review of ISO/TS 11819-3 within working group ISO/TC 43/SC 1/WG33, several users reported issues. No distinction is currently made between analogue and digital durometers while users report different results depending on the device used and digital devices allow various time settings.

Previous work done by Vieira [3] suggests a re-examination of total variance related to the shore A hardness measurement reported in ISO/TS 11819-3. The objective of this study is to evaluate the reproducibility of the current procedure and to quantify the influence of operator expertise, durometer type, device settings and measurement position as input for the ongoing revision of ISO/TS 11819-3.

Improving the reproducibility of the hardness measurement procedure is essential for ensuring comparability of road surface noise measurements between institutes.

2. Measurement campaign

A round robin test (RRT) was organised. This chapter presents the measurement protocol, the received dataset and devices used.

2.1 Measurement protocol

A measurement template was developed to harmonize data collection across participating institutes.
Measurements were performed at ten predefined positions around the tyre's circumference, rather than the four positions defined in ISO/TS 11819-3. At each position, measurements were taken on both the inner and outer rib. For each position, up to eight individual measurements could be reported. Operators were instructed to report all measurements, including invalid measurements. The valid measurements were to be clearly marked. In addition, calibration blocks were measured. The template also recorded self-assessed operator expertise, device type, measurement settings, date of measurement, ambient temperature, and other relevant contextual information.

2.2 Dataset description

Data were submitted by seven institutes. Six institutes measured one tyre, while one institute measured ten tyres. In total, 29 operators contributed to the measurements, with some operators performing two repeated runs. Three institutes used two devices, whereas four institutes used only one device. Several institutes reported measurements obtained under multiple device settings using digital devices. The device settings affected the holding time of the durometer during measurement, with reported values ranging from 0s, corresponding to peak-value measurements, up to 3s. Only four operators classified themselves as experts. Reference block measurements were provided by five institutes.
Despite the use of a common measurement template, differences in interpretation of the standard and reporting instructions were observed. Institutes reported either four, eight, or ten measurements, and markings of selected measurements varied between operators and institutes. In some cases, four consecutive measurements within 2 Shore A were marked; in others, four non-consecutive measurements within two units were marked. Some datasets contained no markings. These differences indicate that further clarification of the measurement protocol and reporting requirements is necessary to improve consistency and comparability across institutes.
To obtain a representative and comparable dataset, only measurements which are valid according to the current version of the standard procedure, i.e. the first four consecutive measurements that differed by no more than 2 Shore A units, were considered.

2.3 Devices

All instruments used in this investigation are Type A durometers meeting the requirements of ISO 868 [4], see Table 1.

Table 1. Devices and settings used.

Device	Analogue	Digital	Setting [s]
ASKER-DD5 (lab)		x	0, 3
Bareiss HPEIII		x	1, 2, 3
BAQ SDA		x	1, 2, 3
DeFelsko PosiTector SHD-A		x	1, 2
Mitutoyo HH-336-01		x	0
Mitutoyo HH-336-11		x	0
ASKER Type A	x		
Checkline AD-A100-A	x		
RS PRO 307	x		
Schmidt PHPSA-M	x		
Zwick Shore A	x		

3. Analysis

This chapter presents the analysis of the reported RRT measurement data. The objective was to evaluate measurement consistency, identify relevant sources of variation, and assess the comparability of results across participating institutes, operator expertise, device types, device settings, etc. No temperature correction was applied to the measured hardness values, since each institute conducted the measurements on its own respective tires and the recorded temperatures remained highly stable across the individual laboratory measurements.
Due to the heterogeneity of the reported data, including differences in measurement procedures, reporting formats, and validity markings, all conclusions should be interpreted with caution.
In the following, the variability of the tyre hardness measurements is analysed at two levels: variation between the ten circumferential measurement positions around the tyre and variation within each individual measurement position. The latter was assessed using four valid values per position, where validity was determined according to the standard.
As always four measurements exist for each individual measurement position, these variations are calculated via an ANOVA table, where the circumferential measurement positions around the tyre are the grouping variable. The variations are calculated for each combination j of tyre, operator, run, measurement device and holding time as well as inner and outer rib position separately, as there is a significant difference between measurements at the inner and outer rib (see section 3.6).
The variation of the 4 $(= n)$ measurements within each individual measurement position is expressed by the standard deviation inside the measurement position

$$s_{\mathrm{r}j} = \sqrt{\frac{\sum_{i=1}^{p}\sum_{k=1}^{n}(y_{ijk} - \overline{y}_{ij})^2}{p(n-1)}} \tag{1}$$

with $p = 10$ for the ten circumferential measurement positions and the mean for each of these positions

$$\bar{y}_{ij} = \frac{1}{n}\sum_{k=1}^{n} y_{ijk}\,. \qquad (2)$$

Here, y_{ijk} denotes a single tyre hardness measurement. The standard deviation inside the measurement positions can be seen as an estimate of the repeatability of the measurements for this specific set of measurements *j*.

The standard deviation between the circumferential measurement positions is calculated from the between group variation:

$$s_{\mathrm{b}j} = \sqrt{\frac{s_{\mathrm{d}j}^2 - s_{\mathrm{r}j}^2}{n}} \qquad (3)$$

with

$$s_{\mathrm{d}j} = \frac{1}{p-1}\sum_{i=1}^{p} n\left(\bar{y}_{ij} - \bar{\bar{y}}_j\right)^2$$

$$\bar{\bar{y}}_j = \frac{\sum_{i=1}^{p}\bar{y}_{ij}}{p}. \qquad (4)$$

The standard deviation between the circumferential measurement positions is an estimate of the variation of the values around the tyre and can be caused by a non-homogenous hardness.

Subsequently, the calculated standard deviations inside and between circumferential measurement positions are plotted as histograms. Further, a kernel density estimator is used to show the probability density function.

3.1 Level of expertise

No guidelines nor descriptions of the several levels of expertise were given to participants. Operators were only asked to do a self-assessment choosing between “beginner”, “advanced” and “expert” for the degree of experience with rubber hardness measurements in general and for the degree of experience with the specific durometer used for the test. This subjective choice of category may have an impact on the uncertainties of the related analyses. For these limited data not much difference was found between the analyses based on general experience or based on specific durometer experience. Therefore the following is based on the general experience.

Figure 1 shows the standard deviations inside measurement positions for the categories beginner, advanced and expert for all validated datasets (including analogue and digital devices with various settings).

Operators who defined themselves as advanced have smaller standard deviations inside measurement positions.

Experts seem to have higher standard deviations but this is because experts reported numerous measurements with 0s peak-value setting, which show higher variability.

3.2 Device types

Results from three institutes measuring with both analogue and digital durometer indicate that digital devices tend to measure lower hardness values.

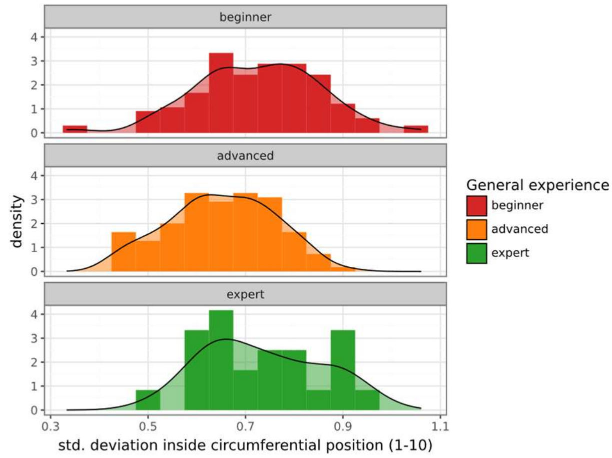

Figure 1. Standard deviations inside measurement positions for different levels of experience with hardness measurements.

Based on the data which comply with the validity criterion defined in the standard, analogue devices show systematically higher standard deviations inside a measurement position than digital devices for all levels of expertise and settings, see Figure 2.

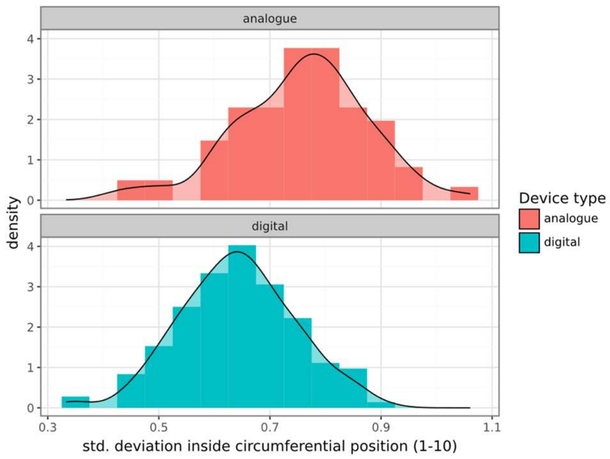

Figure 2. Standard deviations inside measurement positions for analogue and digital devices.

Analogue durometers are reported in integer values only, which is resulting in lower measurement resolution compared to the digital devices, which register results in decimals. Therefore, an interval of almost 3 Shore A may be assumed as valid for analogue devices and not for digital devices, according to the current procedure in the TS (e.g. 65.5 – 68.4 Shore A results in “within 2 Shore A” for an analogue device with integer reading). This contributes to a lower standard deviation of digital devices and may explain the skewedness of the distribution.

3.3 Setting of digital devices

Based on measurements from three institutes using digital durometers with various time settings, increasing the time setting yields lower hardness results.

In Figure 3 standard deviations between measurement positions 0-10 for various settings of digital devices are shown for all validated datasets. Standard deviation decreases with increasing number of seconds (at least up to 2s). The 3s results are based on data from only three institutes.

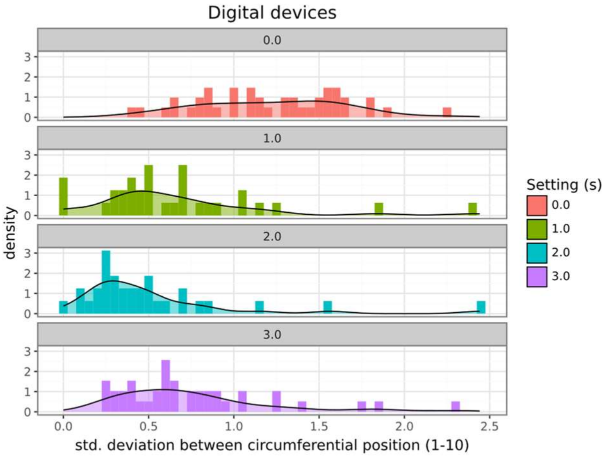

Figure 3. Standard deviations between measurement positions 1-10 for various settings of digital devices (0, 1, 2 and 3 s).

A setting of 2s seems to be a good compromise between limiting the needed measurement time and maximising the precision.
Mean values of valid measurements for different positions around the tyre can differ in hardness but should not depend on device setting (or operator / device type). If this standard deviation depends on setting or other parameters, this means that there is a variation between the measurement points which cannot be explained by actual difference in hardness around tyre.
Figure 4 shows the standard deviations inside measurement positions for various settings of digital devices. The result of 3s is only based on the data of one institute so they should be handled with care. The results confirm that a setting of 2s is a good choice.

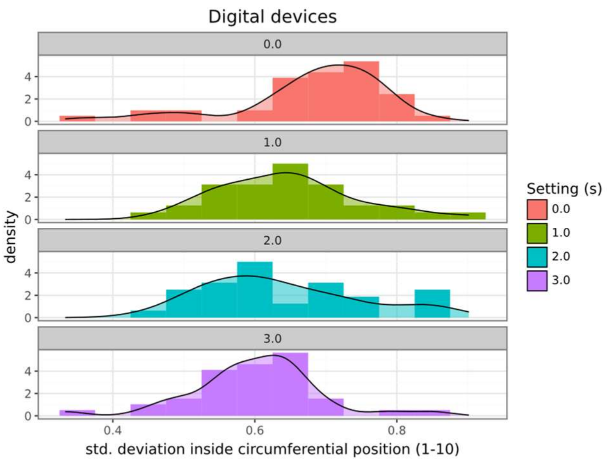

Figure 4. Standard deviations inside measurement positions for various settings of digital devices.

3.4 Number of used measurement positions around the tyre

Again, only measurements that pass the standard validity criterion are considered. Only data from institutes that measured 10 measurement points around the tyre are used.
The mean rubber hardness value of all 10 measurement positions is calculated and used as reference, separately for each setting, experience and device type. Then, for all possible combinations of using only 4 to 9 measurement positions, the mean values are calculated. Again, separately for each setting, experience and device type. The difference between this mean value and the reference mean value is then calculated.
The results are shown in Figure 5. The number of used measurement positions is plotted on the X-axis. The share of mean values within certain intervals Δ is plotted on the Y-axis (60 to 100 %). The columns (from left to right) show the results for analogue device, digital device with settings 0s, 1s, 2s and 3s. Each row represents a different interval (from top to bottom): $\pm$ 0.5 Shore A; $\pm$0.75 Shore A and $\pm$ 1 Shore A which can be an allowable deviation from the reference value.

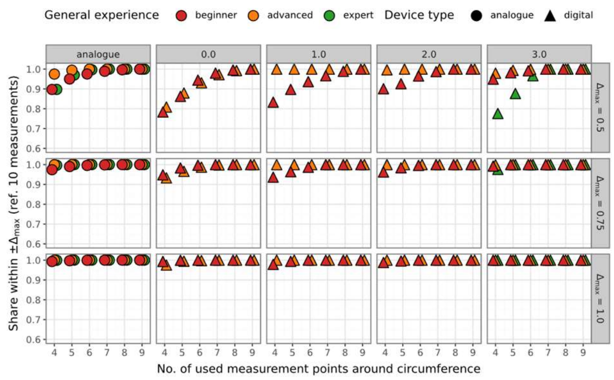

Figure 5. Share of combinations that stay within $\pm$0.5, $\pm$0.75 or $\pm$1 Shore A of the reference value obtained from 10 positions as a function of the number of measurement positions used.

With analogue devices used by beginners and expert operators, approximately 90 % of the possible combinations of 4 measurement positions yield a mean value which differs by maximum 0.5 Shore A from the reference value.
Peak-values 0s with digital devices yield a large spread. Research of Vieira [3] also recommended not to use maximum values. It is therefore advised not to use peak-values.
Less than 95 % of combinations of up to 6 measurement positions yield deviations less than 0.5 Shore A for digital devices with 1s setting operated by beginners.
In Figure 6, the difference between the mean value and the reference mean value of 10 measurement points is shown as a function of the number of measurement points for various settings of digital devices (0, 1, 2 and 3s) and general experience (beginner, advanced, expert).

With increasing experience, time settings and number of measurement positions the deviation between the mean value and the reference mean value is less.
For the 2s digital values fewer measurement values are necessary for the same accuracy in comparison to the other digital settings as well as to analogue devices.

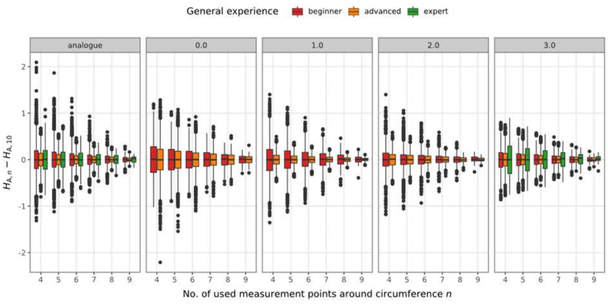

Figure 6. Deviation from the 10-position reference value as a function of the number of measurement positions used, grouped by digital device setting (0, 1, 2, 3s) and operator experience (beginner, advanced, expert).

For the 3s values, the results seem to be better for beginners than for experts. However, numerous 3s value data from experts come from a digital device that is not meant for handheld measurements but for measuring in a laboratory with a test stand, causing this larger spread.
Based on these data no definite conclusion can be drawn on how may repetitions for each measurement spot are necessary as well as on the need of obtaining “consecutive” values which differ by no more than 2 units. Additional measurements are foreseen.

3.5 Reference blocks

Reference block measurements were compared with tyre measurements to investigate whether a relationship exists between the respective measurement biases. If such a relationship can be established, reference block measurements can potentially be used to correct tyre hardness measurements.
The bias of each individual operator is calculated with respect to the mean of the corresponding institute for analogue and digital devices separately. For the latter all device settings are included. Figure 7 shows the results.

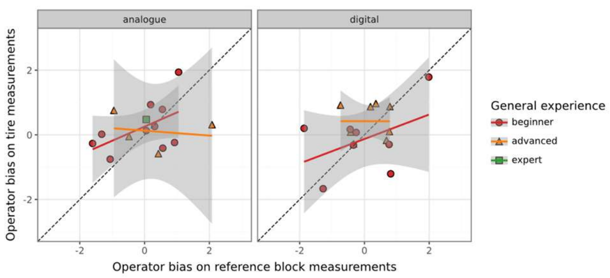

Figure 7. Operator bias on tyre measurements as a function of bias on reference block measurement. The dashed line shows the 1:1 relation.

Beginners show a similar tendency relative to the institute mean on both the reference blocks and the tire. No correlation is found for advanced users. Not enough data of experts are available. Additional data collection under controlled conditions is planned.

3.6 Inner/outer measurement positions

Hardness was measured on inner and outer ribs, as specified in the current procedure of the TS. The measurement results were mostly lower on the outer rib than on the inner rib. Figure 8 shows the results for the various institutes. This corresponds to the research of Vieira [3], where on average a difference in 0.56 Shore A was found. One institute, reporting hardness values above 75 Shore A, exhibited the opposite trend. This may indicate a different behaviour at higher hardness levels; however, a misunderstanding of the inner and outer rib definitions cannot be excluded.

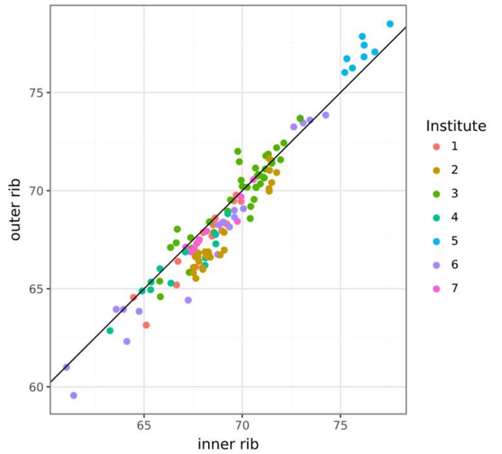

Figure 8. Shore A Hardness results on inner versus outer rib for various institutes.

4. Conclusion

A round robin test involving seven institutes was conducted to investigate the rubber hardness measurement procedure specified in ISO/TS 11819-3. Operator, durometer type and digital time setting influences on shore A hardness measurements of CPX reference tyres were studied. Digital durometers generally provide lower variability than analogue devices. Digital durometers tend to yield lower hardness results than analogue devices.
For digital devices a holding time of 2 s appears to offer a good compromise between measurement stability and practicality. Increasing time settings involves lower hardness results.
Operator experience matters. A smaller variability is noted for advanced users compared to beginners. It is therefore advised to include training and/or a defined minimum proficiency level in the Technical Specification.
More measurement points are needed for analogue devices and beginners to obtain a similar measurement uncertainty as digital devices and advanced operators.

No clear relationship was found between operator bias on reference blocks and tyre hardness measurements, indicating that reference block measurement may not be directly used to correct tyre hardness measurements. The analysis demonstrated that hardness values are generally lower on the outer rib than on the inner rib, confirming findings of another study.
Additional data collection performed by advanced users is planned including the use of more digital device types and measuring on the same tyre to supplement these data.
Results will be used as input for the ongoing revision of ISO/TS 11819-3.

5. Acknowledgments

We gratefully acknowledge the contributing institutes for providing the data used in this study: AIT (AT), BRRC (BE), M+P. (NL), New Zealand Transport Agency (NZ), Nissan Motor Co. (JP), Université Gustave Eiffel (FR) and Slovenian National Building and Civil Engineering Institute ZAG (SI).

References

[1] [International Organization for Standardization, ISO 11819-2 Acoustics – Measurement of the influence of road surfaces on traffic noise – Part 2: The close-proximity method, 2017.

[2] [International Organization for Standardization, ISO/TS 11819-3 Acoustics – Measurement of the influence of road surfaces on traffic noise – Part 3: Reference tyres, 2021.

[3] Vieira, T. et al. : "Evaluation of uncertainty on Shore hardness measurements of tyre treads and implications to tyre/road noise measurements with the Close Proximity method", vol. 162, no. 107882, ISSN 0263-2241, 2020.

[4] International Organization for Standardization, ISO 868 Plastics and ebonite – Determination of indentation hardness by means of a durometer (Shore hardness), 2003.