Ensuring the validity of results of hydrophone calibration by a mutual comparison calibration method

E G Oliveira, A V Alvarenga, R P B Costa-Felix

Laboratory of Ultrasound, Inmetro. Av. Nossa Sra das Graças, 50 – Xerém, Duque de Caxias, RJ, Brazil, ZIP 25.250-020.

rpfelix@inmetro.gov.br

Abstract. Hydrophone calibration is one of the most important basic metrology activities in ultrasound. Hydrophones are used in many different calibrations and testing methods in industry, for instance, the evaluation of medical electrical equipment basic safety and essential performance. Ensuring the validity of results of metrological services is a major concern for any laboratory. Primary calibration of measuring instrument or standard is a widely understood way to assure good results, despite it is not enough, it is not the unique way, neither it is easy to get in some circumstances. For instance, hydrophone calibration is a service restricted to a few laboratories worldwide, which demands a logistic effort and may be quite expensive. An alternative method to guarantee the traceability by in-house check is of interest. This paper presents a method to ensure the traceability of hydrophones after mutual calibrations. Eight previously calibrated hydrophones from different manufactures and characteristics (membrane and needle) were mutually calibrated, by comparison, all of them acting as standards and calibration items. Calibrations were performed from 1 MHz to 10 MHz and the results were evaluated by pair-difference chi-squared statistics. Uncertainties typically laid between 6% to 13% and sensitivities were between 33 to 318 nV Pa\(^{-1}\). After statistics analyses, 4 hydrophones were considered reliable to be used as standards, as their sensitivity keep on stable even after several years since the last calibration. The method can be used by the industry, accredited laboratories, or national metrology institutes to perform functional checks for virtually any measuring instrument or standard, being the limit the number of available devices to properly apply the statistical analysis.

Keywords. Metrology, ultrasound, hydrophone, mutual comparison calibration, pair-difference chi-squared.

1. Introduction

Hydrophones are used worldwide to perform measurements for an assortment of applications. In the industrial field, ultrasonic methods are commonly used to evaluate the internal structure of metallic pieces, detection of flaws in welds, concrete strength, the thickness of pressure vessel walls, among many others [1]. To investigate the capability of ultrasonic transducers to operate properly for those applications, hydrophones can be used to evaluate their output. For medical applications, hydrophones
are used to check the output field of many medical electrical devices, like those used for physiotherapy or diagnostics [2][3][4].

In all cases, hydrophone technical viability shall be checked periodically. This paper presents a method developed to check the reliability of a set of hydrophones to work as standards for secondary calibrations of sensitivity. The development was based on pair-difference chi-squared statistics.

1.1. The validity of measurement results and calibration status
A major concern for laboratories attending to the technical requirements of ISO/IEC 17025 is the assurance of the validity of measurement results, both for calibration or testing [5]. Metrological traceability (clause 6.5) is necessary but it is not enough to assure the quality of metrological services. Periodical functional check of the measuring system is another technical requirement that shall be fulfilled, as defined in the clause 7.7.1, to a laboratory guarantee that their outcomes are good enough for the intended use. Furthermore, calibrations are not always simple to get, as they may be off handing, expensive, or could depend on complex logistics. For instance, according to the Bureau International des Poids et Mesures (BIPM [6]), only three National Metrology Institutes carry out hydrophone calibrations in the MHz range: Physikalisch-Technische Bundesanstalt (PTB, Germany); Institute of Physical-Technical and Radiotechnical Measurements, Rosstandart (VNIIFTRI, Russian Federation); and the National Physical Laboratory (NPL, United Kingdom).

Calibration periodicity shall be defined within any laboratory Measurement Management System, preferable after the definition of metrological requirements for the measuring system [7]. It is important to check periodically the valid calibration status. As defined in clause 7.1.1 of [7], “recalibration of the measuring equipment is not necessary if the equipment is already in a valid calibration status”. The concern is how to confirm the metrological status of a measuring device. A proper technical method, based on robust statistics and rigid protocol, must be applied.

1.2. Hydrophone calibration
Hydrophones are devices used to transform mechanical waves into electrical signals. The complexity of the transduction process or the hydrophone construction is beyond the scope of this paper, but one could find details about that somewhere else [8][9]. As a metrological device, hydrophones shall be periodically calibrated and functionally checked to allow proper signal transduction, from dynamic pressure (measured in Pa) and effective voltage (in V measured in RMS). The technical international standard IEC 62127-2 [10] discloses many different calibration methods, including primary and secondary approaches (please refer to the International Vocabulary on Metrology – VIM – to access the definitions and use of terms in metrology [11]).

Clause 12 of IEC 61127-2 defines generally the comparison calibration method using a standard hydrophone, which is a traditional calibration by substitution. A standard is a hydrophone that has its characteristics traced to a primary standard or it is a primary standard itself. The main characteristics of a hydrophone are the sensitivity (magnitude and phase), directivity, and linearity. Within this paper, the quantity of interest is the magnitude sensitivity for propagation waves with normal incidence to the active element of the hydrophone for amplitudes within the linear response range of the hydrophone.

The method consists of generating an ultrasound wave in a propagation medium (typical water at room temperature) with the aid of an auxiliary output transducer (emitter), a signal generator, and, optionally, a power amplifier. The ultrasonic wave captured by the hydrophone (receiver) is converted to an electrical signal which is recorded, being the effective voltage the quantity to be stored. The generated waveform may have dissimilar impulse responses, for which the most usual is the burst of sines in a defined frequency of interest. The length of the burst, i.e., the number of cycles depends mainly on the distance between the transducer and the reflecting surfaces, targets or scatters that may have in the field (tank wall or water surface, for instance), as well as the direct path between the emitter and the receiver. In a second step, the standard hydrophone is replaced by the item under calibration, which active element must be placed in the same spot that the standard hydrophone was previously positioned. The distance is determined by the time of flight of the waveform from the emitter to the receiver, measured in units
of time, and the directivity is determined by rotation and tilt of the emitter and the receiver. The objective is to maximize the input signal, which defines the best alignment between the transducer and the hydrophone. The procedure shall be repeated a couple of times to assess the random component of the uncertainty under repeatability conditions (type A uncertainty). Please refer to VIM [11] and Guide to the expression of uncertainty in measurement (GUM – [12]) for further details.

The general equation for the comparison calibration of hydrophone sensitivity is [10]:

\[ M_{HUT}(f) = \frac{V_{HUT}(f)}{V_{Std}(f)} M_{Std}(f) \]  

In this research, the sensitivities \( M_{HUT} \) and \( M_{Std} \) for the hydrophone under test (HUT) and the standard (Std) are scalars representing the magnitude as a function of frequency \( f \), whilst the RMS voltage for each frequency is represented by \( V_{HUT} \) and \( V_{Std} \).

1.3. Normalized error and pair-difference of the measurement result

When a pair of calibrations are done with devices that are supposed to have the same (or similar) response to the stimuli they had, statistical analyses are mandatory in a proper metrological approach. For international comparisons, like those organized by BIPM or Regional Metrology Organizations (RMO), the most accepted statistics is the normalized error. Rationales about the relevance and applicability of the normalized error to compare the results of a pair of calibrations can be found in [13][14][15]. Normalized error, typically represented by \( E_n \), accounts for the normalized difference between a pair of metrologically assessed values of a quantity or a parameter. The normalization is done upon quantification of the combined dispersion of the results for each of the measurement results. Mathematically, the normalized error \( E_n \) is expressed by (2).

\[ E_n = \frac{1}{k} \frac{|x_1 - x_2|}{\sqrt{u^2(x_1 - x_2)}} \]  

That formulation is a slightly different arrangement from the definition presented in [14] and [15], mixing both but without misrepresenting the original meaning of either. In that formulation, \( x_1 \) and \( x_2 \) are the results of two realization of a quantity of interest (for instance, calibrations of a measuring device). The value of \( k \) depends on the effective degrees of freedom \( v_{eff} \) of the combined uncertainty \( u(x_1 - x_2) \). In the case that \( x_1 \) and \( x_2 \) can be considered independent realizations of the quantity under calibration, or if \( u_{x1} \) and \( u_{x2} \), respectively the combined uncertainties computed for \( x_1 \) and \( x_2 \), have any correlated uncertainties sources disregarded, equations (3) and (4) apply [14], in which \( v_{x1} \) and \( v_{x2} \) are their degrees of freedom, respectively.

\[ ku(x_1 - x_2) = [(ku_{x1})^2 + (ku_{x2})^2]^{1/2} \]  

\[ v_{eff} = v_{x1} + v_{x2} \]  

For a defined coverage probability (or level of confidence) of \( p = 0.95 \) and \( v_{eff} \rightarrow \infty \), \( k \rightarrow 1.96 \) or \( k \cong 2 \).

2. Material and method

The subject of this research is the mutual calibration of hydrophones under the comparison method. Eight hydrophones were used as standard and HUT, each one at its turn. That means that a total of 56 calibrations were done, each one of the 8 hydrophones working as standard for 7 HUT.

2.1. Measuring devices

In table 1, there is a summary of all hydrophones used in the experiment. Table 2 discloses the calibrated sensitivity of each device. All hydrophones were calibrated by an NMI except H4 and H6, which were
calibrated by their manufacturers. The last time they were externally calibrated was in 2007 (H4), 2010 (H2, H3, H5, and H8), 2011 (H1 and H7), and 2015 (H6). Internal calibrations were done periodically ever since but with a non-systematic control of periodicity. The measurements reported in this paper were done in 2017.

Table 1 – Hydrophones description

| Id | Model | Manufacturer       | Hydrophone s.n. | Characteristics               |
|----|-------|--------------------|----------------|------------------------------|
| H1 | UC1604-029 | Precision Acoustics | UC1604-029     | Membrane; 0.4 mm active element; 15 µm PVDF |
| H2 | UC1604-068 | Precision Acoustics | UC1604-068     | Membrane; 0.2 mm active element; 15 µm PVDF |
| H3 | UC1602-023 | Precision Acoustics | UC1602-023     | Membrane; 0.2 mm active element; 15 µm PVDF |
| H4 | 804-224   | Aertara (formerly Sonora Medical) | S4-224     | Membrane; 0.4 mm active element; 25 µm PVDF |
| H5 | S4-225    | Aertara (formerly Sonora Medical) | S4-225     | Membrane; 0.4 mm active element; 25 µm PVDF |
| H6 | NH0500 | Precision Acoustics | NH0500     | Needle; 0.5 mm active element; 9 µm PVDF |
| H7 | NH0200   | Precision Acoustics | NH0200     | Needle; 0.2 mm active element; 9 µm PVDF |
| H8 | NH0500 | Precision Acoustics | NH0500     | Needle; 0.5 mm active element; 9 µm PVDF |

Table 2 – Hydrophones calibrated sensitivities

| Id | Hydrophone | 1 MHz | Ucal-0.95 [%] | 3 MHz | Ucal-0.95 [%] | 5 MHz | Ucal-0.95 [%] | 7 MHz | Ucal-0.95 [%] | 10 MHz | Ucal-0.95 [%] |
|----|------------|-------|--------------|-------|--------------|-------|--------------|-------|--------------|--------|--------------|
| H1 | UC1604-029 | 71    | 6%           | 75    | 6%           | 76    | 6%           | 76    | 6%           | 74     | 7%           |
| H2 | UC1604-068 | 54    | 6%           | 56    | 6%           | 56    | 6%           | 56    | 6%           | 55     | 7%           |
| H3 | UC1602-023 | 33    | 7%           | 36    | 7%           | 37    | 7%           | 38    | 7%           | 37     | 7%           |
| H4 | S4-224     | 77    | 12%          | 84    | 12%          | 88    | 12%          | 93    | 12%          | 98     | 12%          |
| H5 | S4-225     | 63    | 6%           | 75    | 6%           | 80    | 6%           | 83    | 6%           | 85     | 7%           |
| H6 | 1085       | 208   | 13%          | 201   | 13%          | 178   | 13%          | 200   | 13%          | 206    | 13%          |
| H7 | 1085       | 33    | 7%           | 37    | 7%           | 51    | 6%           | 51    | 6%           | 47     | 7%           |
| H8 | 1085       | 318   | 8%           | 286   | 8%           | 245   | 8%           | 266   | 8%           | 267    | 7%           |

Table 3 discloses the characteristics of the auxiliary transducers used in the experiment. One must observe that the frequencies reported therein are the nominal as informed by the manufacturer. Tx2 and Tx4 were driven at 3 MHz and 7 MHz, respectively, instead of in their nominal frequency.

Table 3 – Auxiliary transducers description

| Id | Model | Manufacturer                | Transducer s.n. | Characteristics             |
|----|-------|-----------------------------|----------------|-----------------------------|
| Tx1| A3025 | Olympus Panametrics-NDT     | 626858         | Nominal diameter 25.4mm; 1 MHz |
| Tx2| A3825 | Olympus Panametrics-NDT     | 540399         | Nominal diameter 12.7mm; 3.5 MHz |
| Tx3| A3095 | Olympus Panametrics-NDT     | 538675         | Nominal diameter 12.7mm; 5 MHz |
| Tx4| A3205 | Olympus Panametrics-NDT     | 569143         | Nominal diameter 12.7mm; 7.5 MHz |
| Tx5| A3115 | Olympus Panametrics-NDT     | 561040         | Nominal diameter 12.7mm; 10 MHz |

2.2. Calibration method

Calibration by substitution as defined in [10] was done in 4 repetitions for each frequency and each pair of hydrophones. Repetition consists of the voltage measurement of each hydrophone for each frequency at a time so that 8 voltage measurements per repetition were undertaken. The frequencies were 1 MHz, 3 MHz, 5 MHz, 7 MHz, and 10 MHz.

For each frequency, the hydrophone active element was positioned at the beginning of the far-field to maximize the ultrasonic wave reception. The alignment between the emitter (transducer) and receiver (hydrophone) was thoughtfully conducted to assure the better signal to noise ratio for each given situation. The driven electrical signal to the transducers was a 20 cycles burst, 20 Vpp generated by an
arbitrary waveform generator model 33250A (Agilent Technologies, USA). The received signals were
digitized with an oscilloscope model TDS 3012B (Tektronics, USA).

The sensitivity for each hydrophone was computed 7 times, each one using another hydrophone as
standard, using equation (1).

2.3. Statistical analyses
For each hydrophone, each one of its 7 comparison calibrations was compared to the external calibration
presented in table 2. The statistics is the normalized error as disclosed in equation (2) with the
simplification of equations (3) and (4). The expanded uncertainties \( U_{x_i} = k u_{x_i} \) where used as a
normalization for the error, being \( k = 2 \) to simplify the assessment as the effective degrees of freedom
were huge. For that pair-difference in which \( E_n \leq 1 \), the comparison calibration was considered
equivalent to the last external calibration.

2.4. Election of standard hydrophones
The normalized error for all pairs of calibrations was the criteria to select those hydrophones that could
be considered good standards. For all frequencies, if \( E_n > 1 \) for 4 or more times for any hydrophone, it
was not computed as standard and the final mutual calibration value would not use it. After the election
of the standard hydrophones mutually comparable, the final sensitivity was assessed as the uncertainty
weighted average calculated as following [13][15]:

\[
M_{w-j} = \frac{\sum_{i=1}^{N_S} \frac{M_i}{u^2_{M_i}}}{\sum_{i=1}^{N_S} \frac{1}{u^2_{M_i}}} \tag{5}
\]

The weighted average sensitivity \( M_{w-j} \) is computed for each hydrophone \( j \) \((1 \leq j \leq 8)\) as a function
of the sensitivities \( M_i \) of each elected standard hydrophone \((1 \leq i \leq N_S, \text{ in which } N_S \text{ is the number of elected standard hydrophones})\).

Finally, the normalized error for each hydrophone is computed comparing the mutual comparison
weighted average sensitivity and the last external calibration.

3. Results and Discussion
Results are divided into three parts:
- Normalized errors;
- Election of standard hydrophones;
- Final magnitude sensitivity.

3.1. Normalized errors
Table 4 is a general compilation of the normalized errors for all hydrophones and all frequencies. The
table for each frequency represents in their columns the hydrophone working as standard, which means
that the lines are the results for hydrophones working as HUT. The last line of each table is the total of
\( E_n \geq 1 \) for each standard acting hydrophone. Likewise, the last columns disclose the number of times a
HUT has its normalized error greater than unity. The bottom left table (“MAX ALL”) shows the
maximum \( E_n \) for each pair \( H_i \times H_j \) considering all frequencies under test.

One can observe that when acting as standards, \( H_2, H_4, H_6, \) and \( H_8 \) were not equivalent to their
external calibration more than 3 times each. Observe also that \( H_1, H_3, \) and \( H_5 \) were not “good standards”
only for \( H_2 \) and \( H_8 \), while \( H_7 \) has \( E_n \geq 1 \) or \( H_2, H_8, \) and \( H_5 \).
Table 4 – Normalized errors

| Hydrophone | 1 MHz | 2 MHz | 3 MHz |
|------------|-------|-------|-------|
| H1         | 0,00  | 3,26  | 0,33  |
| H2         | 0,46  | 0,11  | 0,31  |
| H3         | 0,31  | 0,15  | 0,31  |
| H4         | 0,46  | 0,11  | 0,31  |
| H5         | 0,37  | 0,15  | 0,31  |
| H6         | 0,46  | 0,11  | 0,31  |
| H7         | 0,37  | 0,15  | 0,31  |
| H8         | 0,37  | 0,15  | 0,31  |

3.2. Election of standard hydrophones

Analyzing Table 4 in detail, according to the selection criteria presented previously, it could be considered a proper choice of standard hydrophones H1, H3, H5, and H7. One can argue that H7 has a normalized error greater than unity when it was used as the standard to calibrate H5. It is true indeed, but the opposite did not happen (H5 as standard calibrating H7), and that fact happened just twice (1 MHz and 3 MHz). Comparing to other hydrophones, the performance could be considered partially satisfactory. Because of that, both H5 and H7 were elected as standard hydrophones.

Observe that H2 and H8 had a bad performance for all frequencies as standards, what is quite symptomatic about their behavior on long term stability. On the other hand, H4 and H6 were just quite good, as they were not good acting as standards in four cases. However, a close look in table 2 will highlight the fact that those hydrophones had the greatest uncertainties in their last external calibration. Because of that, their normalized error is somehow understated.

Table 5 summarizes the mutual comparison calibration for all devices and frequencies. Observe that the expanded uncertainty reported in table 5 for the weighted mean was assessed based on [13][15]. Therefore, the combined weighted uncertainties are lower than the originally calibrated uncertainty. Despite it is quite acceptable for key comparisons, the use of such values of uncertainties within the scope of the method proposed in this paper should be done with reasonable caution and under the discretion of the user. The hydrophones were greyed out in table 5 if they were not considered reliable enough to be elected as standards.
3.3. Final magnitude sensitivity

Table 6 summarizes the findings of this research. The weighted sensitivity mean $S_w$ and its weighted expanded uncertainty $U_w$ were the comparison values to define the normalized error $E_n$ concerning the external calibration. Only the elected standard hydrophones were used to assess $S_w$. Observe that the normalized error was great than unity only for H2 and H8, confirming the suspicion that they were not reliable since the mutual calibration pair-difference was assessed (see table 4).

Table 6 – Sensitivity magnitude by mutual calibration and normalized errors

| Id   | Hydrophone | Original | Mutual | Original | Mutual | Original | Mutual |
|------|------------|----------|--------|----------|--------|----------|--------|
|      |            | $S_w$    | $U_w$  | $E_n$    | $S_w$  | $U_w$    | $E_n$  |
|      |            | [nV Pa$^{-1}$] [%] | [nV Pa$^{-1}$] [%] | [nV Pa$^{-1}$] [%] |
| H1   | UC1604-029 | 70.2 4.8% 0.1 | 75.4 4.7% 0.1 | 75.2 4.7% 0.0 | 77.1 5.0% 0.1 |
| H2   | UC1604-068 | 44.2 4.8% 0.2 | 47.4 4.7% 0.1 | 46.8 4.7% 0.2 | 45.3 5.0% 0.2 |
| H3   | UC1602-023 | 33.3 4.8% 0.3 | 37.7 4.7% 0.2 | 37.8 4.7% 0.1 | 36.9 5.0% 0.0 |
| H4   | S4-224     | 75.8 4.8% 0.0 | 95.3 4.7% 0.6 | 98.1 4.7% 0.4 | 98.5 5.0% 0.0 |
| H5   | S4-225     | 66.6 4.8% 0.7 | 76.3 4.7% 0.2 | 80.8 4.7% 0.1 | 84.0 4.7% 0.2 | 85.5 5.0% 0.1 |
| H6   | 1085       | 238.1 4.8% 1.2 | 269.9 4.7% 0.6 | 222.2 4.7% 1.0 | 235.5 4.7% 1.3 | 235.5 5.0% 1.5 |

4. Conclusion

Calibration of measuring devices is considered the most reliable way to assure the quality of the metrological services that those calibrated devices could offer. It is partially true. Regardless of the issues with environmental conditions and personnel, the measuring system contribution with the final results should be scrutinized even further. A proposal to accomplish with that premise is the periodical check of the measuring system.

The results presented in this paper were based on a mutual comparison calibration of hydrophone magnitude sensitivity. Eight hydrophones were calibrated pair to pair, and the pair-differences were statistically compared. Evaluating the final results, 4 hydrophones were considered reliable enough to keep on their valid calibration status as standards, and as so they can be maintained as references to be used in calibrations furthermore.
References

[1] Alvarenga, A.V., Silva, C.E.R., Costa-Félix, R.P.B. Monte Carlo uncertainty assessment of ultrasonic beam parameters from immersion transducers used to non-destructive testing. Ultrasonics 69, pp. 144-151. 2016.

[2] Souza, R.M., Santos, T.Q., Oliveira, D.P., (...), Alvarenga, A.V., Costa-Felix, R.P.B. Standard operating procedure to prepare agar phantoms. Journal of Physics: Conference Series 733(1), 012044. 2016.

[3] Abrunhosa, V.M., Soares, C.P., Batista Possidonio, A.C., (...), Costa, M.L., Mermelstein, C. Induction of Skeletal Muscle Differentiation InVitro by Therapeutic Ultrasound. Ultrasound in Medicine and Biology 40(3), pp. 504-512. 2014.

[4] Alvarenga, A.V., Costa-Félix, R.P.B. Uncertainty assessment of effective radiating area and beam non-uniformity ratio of ultrasound transducers determined according to IEC 61689:2007. Metrologia 46(3), pp. 367-374. 2009.

[5] International Organization for Standardization; International Electrotechnical Commission. ISO/IEC 17025 – General requirements for the competence of testing and calibration laboratories. Ed. 3.0. Genève: ISO/IEC. 30p. 2017.

[6] Bureau International des Poids et Measures. BIPM Calibration and Measurement Capabilities (CMS) from the Key Comparison Database (KCDB). https://kcdb.bipm.org/AppendixC/. Accessed on Aug 12th, 2019.

[7] International Organization for Standardization. ISO 10012 – Measurement management systems – Requirements for measurement processes and measuring equipment. Ed. 1.0. Genève: ISO/IEC. 26p. 2003.

[8] Bacon, D R. Characteristics of a PVDF membrane hydrophone for use in the range 1-100 MHz. IEEE Transactions on Sonics and Ultrasonics, v. 29(1), pp 18-25. 1982. DOI: 10.1109/T-SU.1982.31298.

[9] Costa-Felix, R P B; Machado, J C. Hydrophone calibration methods. Research on Biomedical Engineering (formerly Brazilian Journal of Biomedical Engineering), v. 23(3), pp 277-292. 2007.

[10] International Electrotechnical Commission. IEC 62127-2 – Ultrasonics – Hydrophones – Part 2: Calibration for ultrasonic fields up to 40 MHz. Ed. 1.2. Genève: ISO/IEC. 30p. 2017.

[11] Bureau International des Poids et Measures; Joint Committee for Guides in Metrology. BIPM JCGM 200:2012 – International vocabulary of metrology – Basic and general concepts and associated terms (VIM). 3rd Ed. Sèvres: BIPM. 108p. 2012.

[12] Bureau International des Poids et Measures; Joint Committee for Guides in Metrology. BIPM JCGM 100:2008 – Evaluation of measurement data – Guide to the expression of uncertainty in measurement (GUM). 1st Ed. Sèvres: BIPM. 132p. 2008.

[13] Douglas, R J; Steele, A G. Pair-difference chi-squared statistics for Key Comparisons. Metrologia v. 43, pp 89-97. 2006. doi:10.1088/0026-1394/43/1/013

[14] Steele, A G; Douglas, R J. Extending $E_k$ for measurement science. Metrologia v. 43, S235-S243. 2006. doi:10.1088/0026-1394/43/4/S10

[15] Cox, M G. The evaluation of key comparison data: determining the largest consistent subset. Metrologia v. 44, pp 187-200. 2007.