Manuscript Details

Manuscript number	APAC_2019_1090_R1
Title	Method for in situ acoustic calibration of smartphone-based sound measurement applications
Article type	Research Paper

Abstract

Participatory measurement of environmental sound levels has gained interest in recent years. The calibration of measuring instruments is currently the main technical limitation. It is often the responsibility of the users and can be a potential source of error or add costs to the measurement protocol. In this article, a calibration protocol is proposed, based on the low variability of the average noise emission of individual vehicles. The advantage of this protocol for the user is that it does not require specific equipment, i.e. reference sound source or device, or special knowledge in acoustics. The method consists in measuring the noise level of a few vehicles as they pass through at different measuring points. The measured levels are compared to the levels expected by a numerical model, the difference serving as an offset for subsequent measurements. The robustness of the protocol is first tested over a large experimental campaign, and it turns out that measuring the passage of 15 vehicles at 3 different locations limits the error to -1.8 +/- 1.0 dB(A). Then, the protocol is tested in real conditions with a set of 8 smartphones. The comparison with a class 1 sound level meter on 6 control points shows an average error on all phones of -0.6 +/- 1.2 dB(A).

Keywords	Environnemental sound levels; low-cost calibration; participatory measurements
Taxonomy	Acoustic Modeling, Environmental Noise Control, Environmental Acoustics
Manuscript category	Europe and Rest of the World

Corresponding Author	Pierre Aumond
Corresponding Author's Institution	Ifsttar
Order of Authors	Pierre Aumond, Arnaud Can, Guillermo Rey Gozalo, Enrique Suarez, Nicolas Fortin
Suggested reviewers	Kurt Heutschi, Chuck Kardous, Deanna Tomerini, Eric Schreurs

Submission Files Included in this PDF

File Name [File Type]

CoverLetter.docx [Cover Letter]

Review.docx [Response to Reviewers]

Manuscript_v2.docx [Manuscript File]

figure1_step1.tif [Figure]

figure1_step2.tif [Figure]

figure1_step3.tif [Figure]

figure1_step4.tif [Figure]

Figure3.jpg [Figure]

declaration-of-competing-interests.docx [Conflict of Interest]

AuthorStatement.docx [Author Statement]

Manuscript.pdf [Supporting File]

Submission Files Not Included in this PDF

File Name [File Type]

figure2_a.png [Figure]

figure2_b.png [Figure]

figure2_c.png [Figure]

Figure3.PNG [Figure]

Figure4.png [Figure]

Figure5.PNG [Figure]

Figure6.PNG [Figure]

Figure7.PNG [Figure]

Available online at www.sciencedirect.com

Method for in situ acoustic calibration of smartphone-based sound measurement applications

Pierre Aumond[a]*, Arnaud Can[a], Guillermo Rey Gozalo[b,c], Enrique Suarez[d]

[a]IFSTTAR, CEREMA, UMRAE, F-44344 Bouguenais, France
[b]INTERRA, Departamento de Física Aplicada, Universidad de Extremadura, Avda. de la Universidad s/n, 10003 Cáceres, Spain
[c]Facultad de Ciencias de la Salud, Universidad Autónoma de Chile, 5 Poniente 1670, 3460000 Talca, Chile
[d]Instituto de Acústica, Universidad Austral de Chile, PO Box 567, Valdivia, Chile

Abstract

Participatory measurement of environmental sound levels has gained interest in recent years. The calibration of measuring instruments is currently the main technical limitation. It is often the responsibility of the users and can be a potential source of error or add costs to the measurement protocol. In this article, a calibration protocol is proposed, based on the low variability of the average noise emission of individual vehicles. The advantage of this protocol for the user is that it does not require specific equipment, *i.e.* reference sound source or device, or special knowledge in acoustics. The method consists in measuring the noise level of a few vehicles as they pass through at different measuring points. The measured levels are compared to the levels expected by a numerical model, the difference serving as an offset for subsequent measurements. The robustness of the protocol is first tested over a large experimental campaign, and it turns out that measuring the passage of 15 vehicles at 3 different locations limits the error to -1.8 +/- 1.0 dB(A). Then, the protocol is tested in real conditions with a set of 8 smartphones. The comparison with a class 1 sound level meter on 6 control points shows an average error on all phones of -0.6 +/- 1.2 dB(A).

Keywords: Environnemental sound levels; low-cost calibration; participatory measurements

1. Introduction

The characterization of sound environments has been enriched in the last years by an unprecedented expansion and diversification of measurement methods. The emergence of low-cost sensors, and more recently the possibility of making measurements via smartphones, are changing the way in which sound environments are characterized. Participatory measurement gained interest as a measurement protocol in which each citizen can perform geo-localized measurements via his smartphone, sent to a server where post-processing is performed [1]. The user thus becomes both a producer and a consumer of noise data: the data he communicates is used to calculate noise maps, and his smartphone sends him back information on his individual exposure.

Many smartphone applications have recently been developed to acquire acoustic data, such as Noise-Tube [2], [3], Ambiciti [4] and NoiseCapture [5], Hush city [6], Niosh SLM [7], etc. Thanks to the measurements

* pierre.aumond@ifsttar.fr

resulting from these applications, noise maps have been proposed, either by aggregating the measurements produced [2], [8] or by correcting modelled noise maps through data assimilation algorithms [9].

Researches have investigated in parallel the ability of smartphones to measure environmental noise. Kardous & Shaw have shown that while many applications provide erroneous results, some meet the criteria for environmental measurement [10], [11]. Ventura et al. [12] observed on a selection of mobile phones from the market that responses are linear for levels in the 45 to 75 dB(A) range. Aumond et al. [13] showed that used in similar conditions the instantaneous sound levels measured with mobile phones correlate very well ($r > 0.9$, $p < 0.05$) with sound levels measured with a class 1 reference sound level meter with a root mean square error smaller than 3 dB(A).

The calibration phase therefore remains the main technical lock of the measurement protocol, as it can be costly and time-consuming, although essential. A cross calibration procedure has been proposed in Can et al. [14], in which measurements made by a smartphone are compared with those made by the fleet of devices, in order to identify and filter operator/device pairs giving imprecise measurements, and to propose corrections for precise but biased measurements. The method requires nevertheless that a high density of measurements has already been collected on the server. In Picaut et al. [15], an individual calibration procedure is developed: the principle is based on the use of a reference smartphone, previously calibrated, communicating automatically with other smartphones that one wishes to calibrate, by means of an acoustic communication protocol.

Although valuable, these two methods do not respond to the specific case of scattered individual participatory measures, carried out individually. To fill this gap, some suggested using everyday objects (using a box full of coins or tearing a sheet of paper) [16], [17]. The standard deviations announced for these protocols may seem acceptable, on the order of 2 dB, but no document has yet been published rigorously presenting the repeatability of the experiment. Another source of noise in our daily lives that is well known and modelled because it has been studied for many years is road traffic. Thus, this article proposes a low-cost in-situ calibration based on traffic measurements that any user can follow without a previously calibrated reference (a sound level meter or another smartphone). The method is based on the hypothesis, which is tested in the article, that although individual vehicle sound power levels can be highly variable, their average varies little as long as: (i) traffic conditions are controlled (constant speed, conventional pavement), (ii) measurements are performed near the sound source and thus propagation effects are limited. The method therefore consists in measuring continuously sound levels on the side of several roads, letting each time several vehicles pass by. The measured sound levels are compared to the levels expected by the CNOSSOS-EU model, the difference being an estimate of the offset to apply for subsequent measurements which is consistent with the results of Schreurs et al. who showed that the median values of traffic noise maxima measured in situ on Brisbane roads are comparable to those predicted by the European IMAGINE/HARMONOISE models [18].

The method is tested in this article on an experimental campaign of 83 usable sound measurements (out of 302) of 15 minutes in the city of Talca, Chile.

The calibration protocol and the experimental campaign is first detailed in Section 2. Section 3 presents the different uncertainty items associated with this protocol. Section 3.1 presents the difference between the sound power level estimated by a numerical model and its estimated value from measurements and using the calibration protocol. Section 3.2 presents the uncertainty related to the input parameters of the protocol left to the user. Section 3.3 shows the analysis of the standard deviation of the measured estimates depending on the number of pass-by and location to consider keeping this uncertainty item below a targeted value. Section 3.4 presents an estimate of the overall uncertainty associated with this method. Section 4 presents the protocol assessment under real conditions. Discussion is presented in Section 5.

2. Method

2.1. Calibration protocol

Results from the literature show that the individual emissions of vehicles measured at a site are highly variable [19]. However, the proposed protocol relies on the hypothesis that, although the sound power levels of individual vehicles may vary widely, the average sound power levels of vehicles passing by a receiver during a measurement period vary little as long as :

- traffic conditions are controlled (constant speed, mainstream asphalt concrete or concrete pavement),
- measurements are made close to the sound source and the effects of propagation are therefore limited.

Thus, it becomes possible to compare the measurements made at the roadside with the expected results, provided by a numerical model. If the number of performed measurements is sufficient, the measurements will converge in theory towards the results provided by the model. In this article, the CNOSSOS-EU model is used, for light vehicles and assuming the reference road pavement [20].

The calibration protocol consists of four steps described below, and illustrated in Fig. 1.

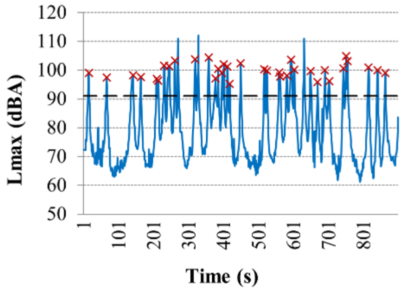
- Step 1: The user performs continuous measurements of roadside sound levels, with a 1-second time resolution ($L_{Fmax,1s}$) and a height of 1.5-m. The user is asked: (i) to wait the passage of a sufficient number of vehicles, (ii) to inform on the distance to the center of the road as well as the estimated mean speed of the vehicles, and (iii) to repeat the procedure at several locations. These numbers of vehicles and locations will be determined in section 3.3.
- Step 2: An algorithm detects the noise peaks in the signal that correspond to the individual pass-by of vehicles. The difficulty here stands in designing a noise-peak selection algorithm that captures vehicle pass-byes, knowing that the absolute noise levels measured might be biased. The algorithm used for this article is *findpeaks* from the R package *pracma v.1.9.9*. It allows to detect peaks in the measured time series $L_{Fmax,1s}$. A set of criteria is set up to keep only the peaks related to the passage of cars. Firstly, only peaks at least 3 seconds apart are retained, to avoid, for example, the measurement of a very close passage of two cars (probably from two different traffic lanes). Secondly, the sound level during the two seconds preceding the peak must be increasing and decreasing during the two seconds following it. Finally, only peaks that correspond to a sound level higher than L_{95} + 15 dB are retained because it is considered that for the protocol to be valid, the signal-to-noise ratio must be greater than this value. Fig. 1b gives several examples of the peaks detection algorithm.
- Step 3: The sound power level L_W that corresponds to each peak detected at the Step 2 is calculated based on the measured L_{max} and the following equation:

$$L_W = L_{max} + 20\log(r) + 10\log 2\pi \qquad (1)$$

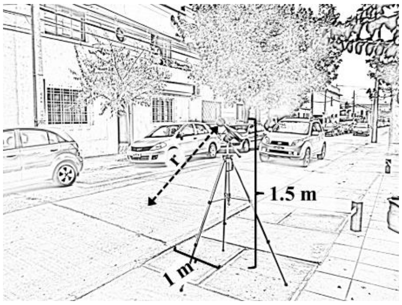
where r is the distance between the source and the microphone. It accounts for the divergence of the sound energy from the source (center of the roadway) to the receiver (microphone), as described in previous studies [19], [21].

- Step 4: The bias between the median of the L_W values and the $\widehat{L_w}$ given by the numerical model for light vehicles and reference conditions is calculated for each of the n sampled site. Then, the average of the n calculated bias is calculated. This value corresponds to the offset that will be stored. The interest of taking the median in there is that it statistically removes heavy or abnormally noisy vehicles, in order to enable the comparison with the numerical model for light vehicles. The advantage of calculating the offset in two steps, calculating an average of the bias calculated for each site, is that it avoids giving too much weight to a site if it has been measured during the passage of more vehicles than the other locations.

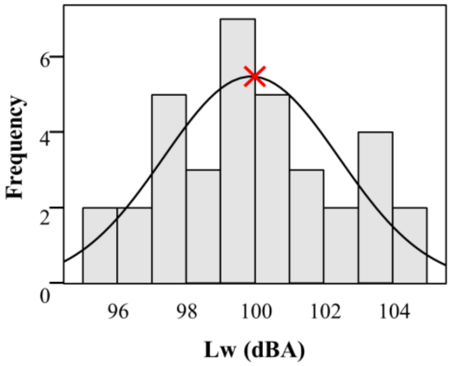
(a) Step 1 – Perform measurement with app or soundmeter



(b) Step 2 – Detect peak levels



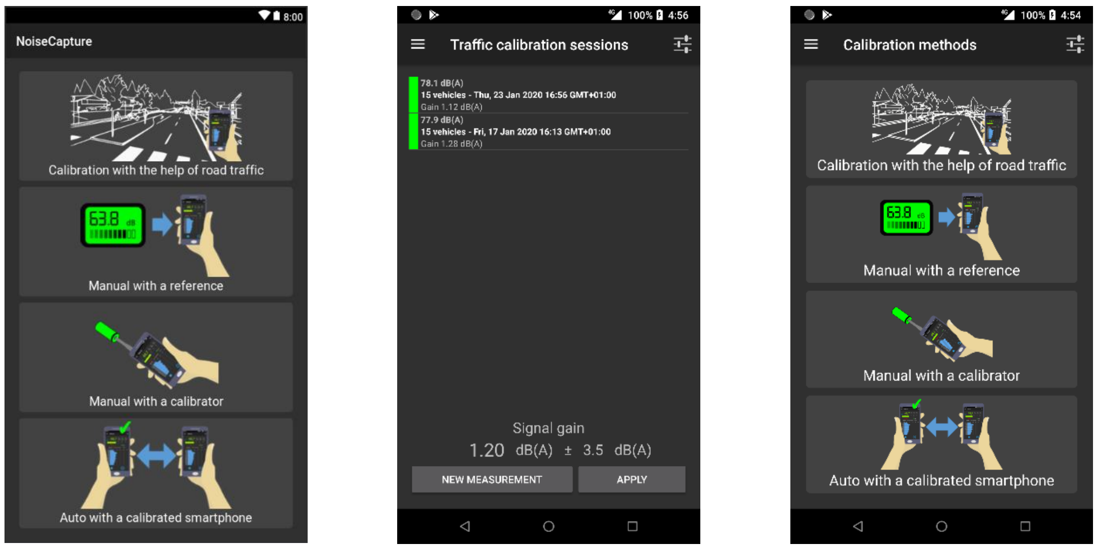
(c) Step 3 – Calculate the sound power level from the peak level and measurement conditions



(d) Step 4 – Get the median sound power level Lw

Fig. 1 The four steps of the calibration protocol.

Note that the user is only concerned by Step 1, the three following ones consisting on post-treatment that need to be implemented on the sensor. The calibration protocol as presented in this paper has been implemented in the NoiseCapture mobile application [5]. Figure 2 shows the user interface for this step within this application. First the user chooses the calibration mode (Figure 2.a), then decides to add a measurement series (Figure 2.b). The user then has to report the distance to the road and the estimated speed of the vehicles and wait for 15 vehicle pass-byes (Figure 2.c). Each correct pass-by is automatically detected in real time in the application which can inform the user.



(a) (b) (c)

Figure 2 User interface of the calibration protocol in the NoiseCapture mobile application. (a) Selection of the calibration record. (b) Adding a set of pass-by measurement. (c) While the measurements are in progress, the user can inform the distance to the road centerline, the estimated speed of the vehicle flow and see how many pass-by were validated.

2.2. Experimental campaign

A total of 302 sound measurements of 15 minutes each were carried out in different locations in the city of Talca (Chile) between 2015 and 2016 [22]. Measurements were carried out on working days. A type-I sound level meter (2250 Brüel & Kjaer) was used with tripod and windshield and it was placed at a height of 1.5 m in free-field conditions following the ISO 1996-2 guidelines [23], [24]. The sound-level meter was located 1 m from the curb. Calibration of the equipment was checked at the beginning and end of each measurement using a 4231 Brüel & Kjær calibrator. The noise parameters L_{eq} and L_{Fmax} for each 1/3 octave band were recorded at 1-second intervals. Flow rates and the average vehicle speed over the 15 minutes were registered together with the sound measurements. Relevant urban features (street dimensions, road surface type, state of the road surface) were also noted. Road traffic was the main source of noise during the sound measurements. When other sound sources or anomalous noise events (e.g., horns and sirens) were detected, the measurements were deleted. In addition, the sampling points that encountered the following conditions were selected:

- measured roads with no or very slight slope;
- measured flow free and with constant speed. The selected average speed were between 57 and 63 km/h (The urban speed limit in Chile was 60 km/h);
- measured roads with standard pavement. Pavements in very poor condition or with specific acoustic performance (as cobbled roads) were discarded;
- measured roads with one or two lanes.

As a result of this filtering process, a total of 83 sound measurements of 15-min taken at different sampling points were used for the analysis.

3. Results

3.1. Bias on the estimation of the sound power level between a numerical model and measurements

Fig. 2 presents the estimated sound power levels at each of the 83 locations, using the protocol presented in section 3 over the 15-min measurements (boxplots). The results are reported by octave band and for global values in dB and dBA, and are compared to the sound power levels estimated with the CNOSSOS model, cars driving at 60 km/h on reference pavement (red curve).

Although the measurements were carried out in Chile, on a different vehicle fleet from the one on which the relationships of the CNOSSOS emission model were built, the average sound power values estimated based on

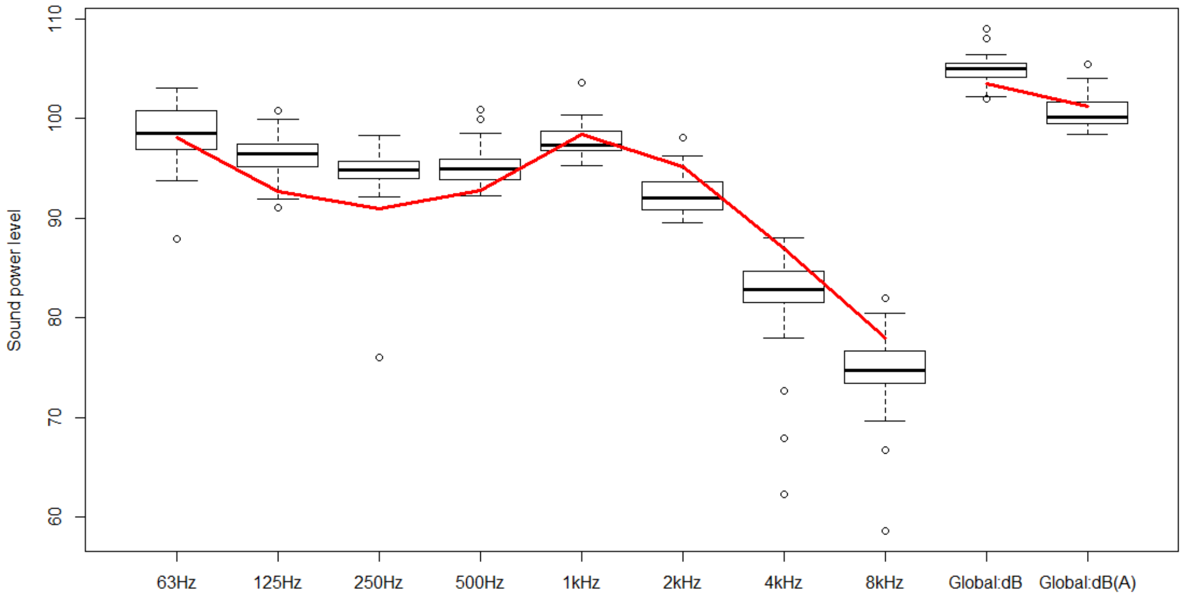


Fig. 3. Estimated sound power levels calculated for each of the 83 selected 15-min measurements using the protocol (boxplots) compared to the estimated sound power levels using the CNOSSOS model (red line)

measurements are close to the ones estimated with the CNOSSOS model. The average bias in the global sound power levels is low (0.5 dBA or -1.8 dB), and the interquartile range is moderately low (2.1 dBA or 1.4 dB). Concerning frequency bands, the average bias and its interquartile ranges are lowest for 500 Hz and 1000 Hz octave bands. Sound power levels in the octave bands from 125 Hz to 250 Hz are slightly overestimated, whereas sound power levels in the octave bands from 2kHz to 8kHz are underestimated. This might be due to differences in the vehicle car fleet or road pavement between Chile and the reference conditions used in CNOSSOS but also to the close proximity of the sound source to the measuring point ($r < 5$m). For example, effects related to the vertical directivity of the source for some frequency bands may not be negligible. In addition, the interquartile range is higher for the octave bands 63 Hz, 4 kHz and 8 kHz. This suggests that the calibration protocol is more robust for global and mid-frequency values. However, since the response of the microphones embedded within smartphones is flat between 250 and 4000 kHz [25], and in view of the results obtained, we suggest applying the protocol at the global level (in dB or dBA), applying the same equivalent offset for each octave band.

Finally, the estimated global sound power levels show a relatively low interquartile range between the measured locations (2.1 dBA or 1.4 dB). Paragraph 3.3 shows that offering the user to reproduce the measurements at different locations and using the resulting average noise level can further reduce this range.

3.2. Uncertainty related to the input parameters of the numerical model

The calibration protocol proposed relies on the comparison between an estimated value of the mean sound power level relying on measurements and its numerical estimation. In this study, the numerical model is composed of CNOSSOS-EU emission part linked to a simple law for acoustic propagation (Eq. 1). For the application of the protocol, three input parameters have to be estimated by the user, distance from the road, height of the measuring point and average vehicle speed. User estimation of these parameters leads to uncertainty in the estimation of the calibration offset. A study of the sensitivity of the model to these parameters is presented in this section.

The height h and the distance to the road d are reference values mentioned in the protocol, which are recommended to be set at h = 1.5 m and d = 1m. A distance to the roadside d = 1m corresponds to a distance $d_{total} = d + 1.5\text{m} = 2.5\text{m}$ suggesting that the width of the road is 3m. Eq.2 allows to set the source-receiver distance r in the Eq. 1.

$$r = \sqrt{d_{total} + (h - 0.05)} \qquad (2)$$

with 0.05 m corresponding to the source height as suggested in the CNOSSOS model.

The user can change these input values if needed. These values can also shift around these references of a few tens of centimeters due to a wrong appreciation of the user, what can lead to uncertainty in the estimate.

The vehicles mean speed can vary within the range of validity of the CNOSSOS emission model, namely from 20 to 130 km/h and must be estimated by the user, who can alternately use the limit speed on the road segment. The literature reports that the estimation of vehicle speeds by pedestrians is subject to error. Strauss et al. show that the standard deviation in the percentage of error on the task of estimating the vehicle speed is about 20% [26].

This section evaluates the influence of errors in the estimation of these parameters, based on Eq. 1 and the emission model used in the reference conditions. Fig. 4 shows the sensitivity to each input parameter. Errors in the height of the measurement point have a low impact on the numerical estimation of the mean sound power level, which stands below 0.5 dB(A) in the tested range (Fig. 3.a). Errors in the distance to the road are slightly more influential, and can exceed 1 dB(A) (Fig. 3.b); however it can be assumed that this parameter is estimated with little error by the user. Errors in the estimated vehicle speed are more influential. For a reference speed of 50 km/h, an underestimation (resp. overestimation) of 10 km/h of the estimated vehicle speed leads to an underestimation of 2.3 dB(A) (resp. overestimation of 2 dB(A)) of L_W. The estimated offset would be biased accordingly.

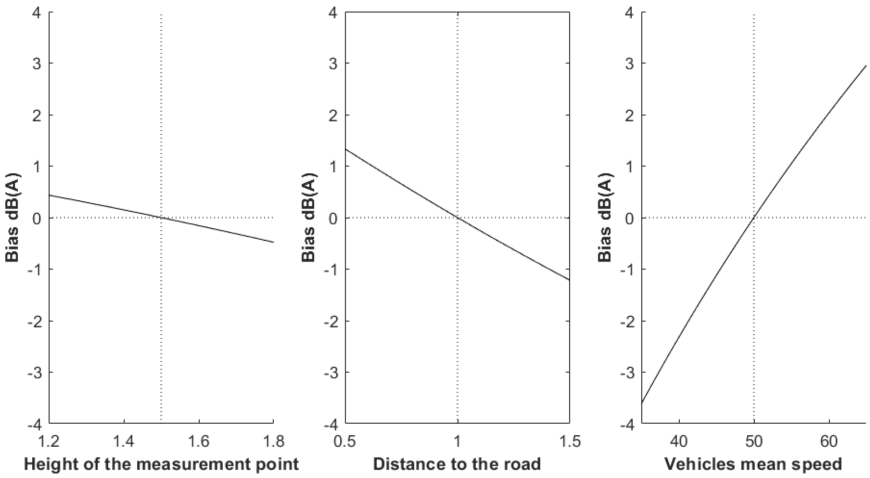


Fig. 4. Influence of the height, distance to the road, and vehicle mean speed on the numerical estimation of the average sound power level estimated at a site. The offset determined following the calibration protocol would be biased accordingly.

In addition, the protocol leaves the responsibility to the user to sample locations with classical road pavements, the errors on which might also affect the L_W value calculated following the CNOSSOS emission model and thus the calibration offset. The user could possibly be asked to choose a corresponding pavement from the database available for his country [27].

3.3. Influence of the number of locations and vehicles per site on the variability of the average L_W estimates

To estimate the average sound power level from measurements with the lowest possible uncertainty, it is necessary to measure a representative number of vehicles, ideally at different locations, in order to obtain a reliable convergence towards the actual average sound power levels of the vehicles fleet. Under an individual calibration protocol context, the proposed method must also be short enough not to discourage the user. Therefore, a sensitivity analysis is conducted to determine the influence of the number of considered locations and vehicles per site on the estimated $\overline{L_W}$ variability.

For each couple "number of locations / number of pass-by per site", a loop of 100 repetitions in which passages are taken randomly within the complete dataset is achieved, on which statistics are calculated. For instance, if the number of locations is set at 7 and the number of vehicle passages is set at 15, for each of the 100 repetitions 7 locations and 15 vehicle passages are randomly selected for each location. Thus, 100 average sound emission levels $\overline{L_W}$ values are estimated according to the procedure presented in Section 2.1.

In rare cases, the method may deviate widely from the standard due to sensitivity to parameters related to peak detection. As these cases are rare, they may be considered outliers. In this case, values that deviate from the mean by two or more times the standard deviation are eliminated from the statistical analysis. A user implementing the protocol would probably also detect these outliers by observing an exaggerated calibration in one measurement relative to the others (> 15 dB). In this case, he would eliminate it himself.

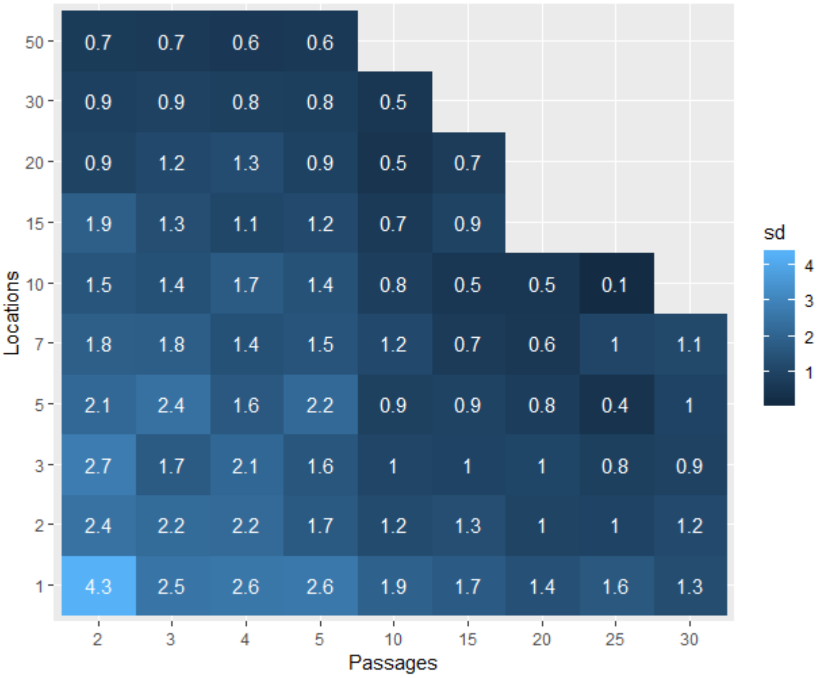


Fig. 6 Standard deviation of the estimated average sound power level calculated over the 100 repetitions as a function of the number of locations and the number of passages per locations.

Fig. 6 shows the standard deviation of the estimated average sound power level $\overline{L_W}$ over the 100 repetitions as a function of the number of locations and the number of passages per site. The results suggest that it is possible to limit the standard deviation under a low value with an acceptable number of vehicle passage and number of locations. For instance, 3 locations and 15 vehicles per location result in a standard deviation of 1.0 dB(A).

A multiple linear regression is calculated and proposed in Table 1 to estimate the standard deviation of the estimated average sound power level from these two variables. The explained variance is 81%. On this basis, reverse modeling can allow the user to choose the best compromise to achieve a targeted uncertainty. For example, using Table 1, for a target uncertainty of less than 1.25 dB, the user can measure 15 vehicle passages at 3 different locations.

Table 1. Estimation the standard deviation from the logarithm of the considered number of pass-by and locations.

$R^2_{adj.} = 81.4$ % | $n = 71$	Estimates	t value
Intercept	3.31	28.18 (p<0.001)
log10(number of locations)	-1.13	-14.91 (p<0.001)
log10(number of pass-by)	-1.30	-13.85 (p<0.001)

Fig. 5 Standard deviation of the estimated sound power level as a function of (a) the number of passages for each number of measured location (dots) and (b) the number of locations for each number of pass-by (dots). Lines are average value.

3.4. Estimation of the overall uncertainty related to the calibration protocol

Three principal items of uncertainty linked to the proposed protocol are:

- The uncertainty u_m linked to the bias of using the emission and propagation model under different conditions than the reference ones, evaluated within the section 3.1. This bias was less than 2 dB(A) in this case study, although it presents the unfavorable case of using the European model CNOSSOS on a Chilean experiment. This bias should be relatively small if local models are used;
- The uncertainty u_p related to the input parameters of the calibration protocol that the user must specify. In view of the results presented in section 3.2, this uncertainty should be in most cases smaller than 2.5 dB;
- The uncertainty u_s related to the sampling strategy, that is the number of locations and pass-byes considered by the user when applying the protocol, evaluated in section 3.3. Applying the protocol with 3 locations and 15 pass-byes per location seems an acceptable compromise, which limits the uncertainty to less than 1.25 dB.

Assuming that all these uncertainties are Gaussian, the overall uncertainty U can be obtained according to Equation 2:

$$U = \sqrt{(u_m)^2 + \left(u_p\right)^2 + (u_s)^2} \qquad (2)$$

The overall uncertainty associated to this calibration protocol is therefore smaller than $\sqrt{(2)^2 + (2.5)^2 + (1.25)^2}$ = 3.4 dB. Moreover, considering that u_m and u_p are fixed to 2 and 2.5 dB respectively alleviates the sampling strategy. Indeed, applying the protocol in a constraining case with 10 locations and 20 pass-byes will yield an uncertainty u_s of 0.5 dB and consequently to an overall uncertainty slightly reduced to 3.2 dB. On the other extreme, the protocol is sensitive to a degradation of the sampling strategy: an unfavorable case with 2 locations and 5 pass-byes would yield an uncertainty u_s of 2.2 dB and consequently to an overall uncertainty of 3.8 dB.

4. Protocol assessment under real conditions

4.1. Method

8 smartphones were selected to assess the calibration protocol in real conditions. 8 participants, owners of the phones, used the NoiseCapture application in which the protocol was implemented. All phones were different (brand and model) and use the android operating system. For the calibration procedure, participants went to 3 locations and waited while 15 vehicles passed by (automatically detected in real time by the app) following the instruction given by the app. As shown in Figure 1, after filling in the measuring distance and evaluating the speed of passage of the vehicles, a calibration value was provided to them.

In a second step, 6 measurements of 3 minutes each were taken. These measurements concern 6 urban sound environments (park, boulevard, quiet street, busy street, courtyard and ring road). These sound environments were measured simultaneously between a Class 1 sound level meter and the 8 phones. The quality of the protocol is assessed on the basis of this set of measures.

The L_{50} indicator is used for the comparison because it is less sensitive than the L_{eq} to measurement conditions as short-term noise close to a specific smartphone.

4.2. Results

Figure 7 shows the sound levels measured by the sound level meter and phones before and after applying the calibration protocol.

For seven of the eight phones, the average error decreases from 4.2 dB(A) before calibration to -0.6 dB(A) after calibration. Analysis of the results also shows that there is a large average standard deviation between the phones prior to calibration 4.8 dB(A) which is strongly reduced after calibration 1.2 dB(A). The eighth phone is considered an outlier and is discarded from the analysis. A saturation of the phone for the highest sound levels may have induce this unexpected behavior. However, the calibration protocol still reduces its bias from 10.4 dB(A) to 4.2 dB(A).

The standard deviation of the bias between sound level meters and smartphones averages 0.7 dB(A). This corresponds to the part of the total standard deviation due to the measurement conditions (*e.g.* directivity, distance between the sound level meter and the phones, etc.). The resulting standard deviation, only due to the proposed calibration protocol is about 1 dB(A).

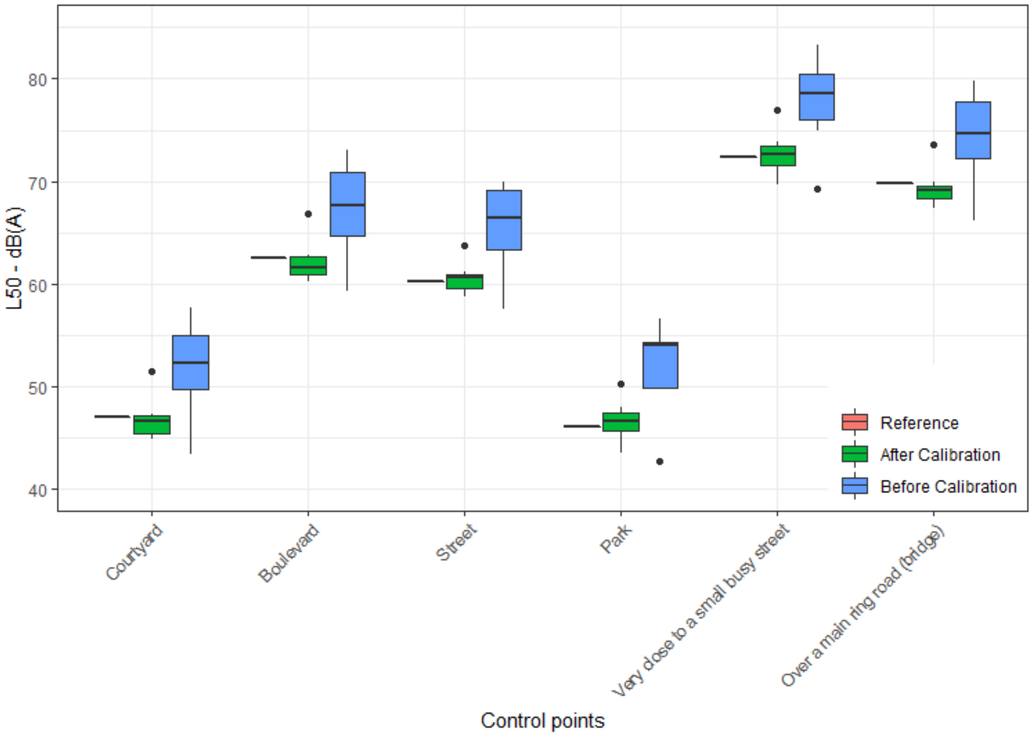


Figure 7 Sound levels measured by the sound level meter and phones before and after applying the calibration protocol.

5. Discussion

In this article, an *in situ* and low-cost calibration protocol is proposed, based on roadside measurements of vehicle passages at different locations. The measured levels are compared to the levels expected by a numerical model, the difference serving as an offset for subsequent measurements.

The uncertainties linked to the protocol are evaluated in section 3. It appears that the two main uncertainty items are the speed evaluated by the user and the fact of using the model under conditions different from those of reference, *i.e.* 2 dB and 2.5 dB respectively.

As concerning the use of the model under different conditions than the reference one, the consequence is a bias in the estimation, which is supposed to be constant under similar conditions: for instance a bias of +0.5 dB in the case of this Chilean experiment. Thus, the consequence would be that all the smartphones calibrated on this study area would have an excessive offset of 0.5 dB, which can be evaluated and corrected if fixed sensors are deployed in parallel.

Beyond the listed uncertainties, the method is dependent on the proposed algorithm for finding peaks. If it is too loose, then events that cannot correspond to vehicle pass byes will be selected, hence adding a bias to the method. The proposed algorithm to find peaks algorithm relies on the threshold L_{95+15} that was shown to remove the low levels peaks. Such a threshold remains however inefficient in the case of high background noise. Therefore, the user is asked to only sample relatively quiet roads.

The robustness of the model should be tested on a wider number of locations, taken in a large variety of countries. Here, the use of the European CNOSSOS model on a data set recorded in Chile leaded to a small and acceptable bias. The optimal approach would be to use locally recommended noise emission models for each territory, where they exist, or to adapt a model according to the vehicle fleet, the type of surfacing most commonly used, etc. A review on the existing noise emission models can be found for instance in [28].

Finally, Section 3.4 shows that the overall uncertainty associated with this calibration protocol remains lower than 3.2 dB. This uncertainty may seem important but these results have to be contrasted:

- The range of sound levels in environmental acoustics can vary from 40 to 95 dB(A). The error is therefore quite low front of the total variability.
- The biases of smartphones without prior calibration can be very important. Mallet, 2017 shows that numbers of phone models bias range between -20 dB(A) and 7 dB(A) [12];
- The measurement uncertainty of a properly calibrated smartphone can reach 2 to 3 dB [13], [29].
- The short experiment carried out under in situ conditions shows that this error is probably overestimated.

In addition the proposed protocol has the advantage of being simple and can be applied by any user without the need to have a calibrated reference sensor nearby. It is therefore particularly suitable for participatory measures of the noise environment. In view of the above considerations, this protocol seems to offer a valid and simple alternative to the calibration of phones.

Finally, the height of 1.50 m chosen in the experiment to hold the measuring tool is a little high (at eye level because it is more suitable for the sound level meter). For an application of the protocol with mobile phones, the height can be reduced to chest height without changing the protocol.

6. Conclusion

This article proposes a calibration protocol based on a noise level measurement of a few vehicles as they pass through different measuring points. The measured levels are compared with the levels expected by a numerical model, with the difference serving as an offset for subsequent measurements. The robustness of the protocol is tested over a large experimental campaign and it turns out that measuring the passage of 15 vehicles at 3 different locations limits the error to -1.8 +/- 1.0 dB(A). Then, the same protocol was applied in real conditions using 8 different phones. 6 control points were used to compare measurements made by a Class 1 sound level meter and the phones. The resulting error after the application of the calibration protocol is -0.6 dB(A) +- 1.2 dB(A).

The present protocol proposes a calibration that does not require special knowledge of the user and a limited time (a few minutes per measuring point). The participatory measurement of environmental noise levels has gained interest in recent years and the main technical limitation is currently the calibration of measuring instruments. Although the uncertainty associated with the protocol is not negligible, it could allow a significant increase in the quality of the measurements collected.

Acknowledgements

This research was funded by the Chilean National Commission for Scientific and Technological Research (CONICYT) through the project FONDECYT No. 1180547. This work was also supported by IFSTTAR (France), a Juan de la Cierva—Incorporación contract from the Spanish Ministry of Economy, Industry and Competitiveness (IJCI-2016-28923), and by the Consejería de Economía, Ciencia y Agenda Digital of Junta de Extremadura, through grants for attracting and returning research talent to R&D&I centres belonging to the Extremadura Science, Technology and Innovation System (TA18019), where University of Extremadura was the beneficiary entity.

References

[1] G. Guillaume *et al.*, "Noise mapping based on participative measurements," *Noise Mapping*, vol. 3, no. 1, pp. 140–156, Jan. 2016, doi: 10.1515/noise-2016-0011.

[2] E. D'Hondt, M. Stevens, and A. Jacobs, "Participatory noise mapping works! An evaluation of participatory sensing as an alternative to standard techniques for environmental monitoring," *Pervasive and Mobile Computing*, vol. 9, no. 5, pp. 681–694, Oct. 2013, doi: 10.1016/j.pmcj.2012.09.002.

[3] N. Maisonneuve, M. Stevens, M. E. Niessen, and L. Steels, "NoiseTube: Measuring and mapping noise pollution with mobile phones," *Environmental Science and Engineering (Subseries: Environmental Science)*, pp. 215–228, 2009, doi: 10.1007/978-3-540-88351-7-16.

[4] R. Ventura, V. Mallet, V. Issarny, P.-G. Raverdy, and F. Rebhi, "Estimation of urban noise with the assimilation of observations crowdsensed by the mobile application Ambiciti," presented at the Internoise 2017, Hong-Kong, 2017.

[5] J. Picaut, N. Fortin, E. Bocher, G. Petit, P. Aumond, and G. Guillaume, "An open-science crowdsourcing approach for producing community noise maps using smartphones," *Building and Environment*, vol. 148, pp. 20–33, Jan. 2019, doi: 10.1016/j.buildenv.2018.10.049.

[6] "From Crowdsourced Data to Open Source Planning: the Implementatio...: Ingenta Connect." [Online]. Available: https://www.ingentaconnect.com/contentone/ince/incecp/2018/00000258/00000004/art00083. [Accessed: 13-Jan-2020].

[7] "The potential use of a NIOSH sound level meter smart device appli...: Ingenta Connect." [Online]. Available: https://www.ingentaconnect.com/contentone/ince/ncej/2019/00000067/00000001/art00003. [Accessed: 13-Jan-2020].

[8] E. Murphy and E. A. King, "Smartphone-based noise mapping: Integrating sound level meter app data into the strategic noise mapping process," *Science of The Total Environment*, vol. 562, pp. 852–859, 2016, doi: 10.1016/j.scitotenv.2016.04.076.

[9] R. Ventura, V. Mallet, and V. Issarny, "Assimilation of mobile phone measurements for noise mapping of a neighborhood," *The Journal of the Acoustical Society of America*, vol. 144, no. 3, pp. 1279–1292, Sep. 2018, doi: 10.1121/1.5052173.

[10] C. A. Kardous and P. B. Shaw, "Evaluation of smartphone sound measurement applications," *The Journal of the Acoustical Society of America*, vol. 135, no. 4, pp. EL186–EL192, Apr. 2014, doi: 10.1121/1.4865269.

[11] C. A. Kardous and P. B. Shaw, "Evaluation of smartphone sound measurement applications (apps) using external microphones—A follow-up study," *The Journal of the Acoustical Society of America*, vol. 140, no. 4, pp. EL327–EL333, Oct. 2016, doi: 10.1121/1.4964639.

[12] R. Ventura, V. Mallet, V. Issarny, P.-G. Raverdy, and F. Rebhi, "Evaluation and calibration of mobile phones for noise monitoring application," *The Journal of the Acoustical Society of America*, vol. 142, no. 5, pp. 3084–3093, Nov. 2017, doi: 10.1121/1.5009448.

[13] P. Aumond *et al.*, “A study of the accuracy of mobile technology for measuring urban noise pollution in large scale participatory sensing campaigns,” *Applied Acoustics*, vol. 117, pp. 219–226, Feb. 2017, doi: 10.1016/j.apacoust.2016.07.011.

[14] A. Can, G. Guillaume, and J. Picaut, “Cross-calibration of participatory sensor networks for environmental noise mapping,” *Applied Acoustics*, vol. 110, pp. 99–109, Sep. 2016, doi: 10.1016/j.apacoust.2016.03.013.

[15] J. Picaut, N. Fortin, and P. Aumond, “Calibrage Acoustique de Smartphones pour la Mesure du Bruit dans l’Environnement,” in *Actes du Congrès Français d’Acoustique*, Le Havre, France, 2018.

[16] *3-Punkt-Kalibrierung von Apps zur Schallpegelmessung*. Available: https://www.youtube.com/watch?v=_5Ste5q4Vv0 [Accessed: 13-Jan-2020].

[17] M. Ziegler, “Bedienungsanleitung. Spaichinger. Schallanalysator. Version 2.2. Freeware,” Gymnasium Spaichingen, September2019.

[18] E. Schreurs, L. Brown, and D. Tomerini, “Maximum pass-by noise levels from vehicles in real road traffic streams: comparison to modeled levels and measurement protocol issues,” presented at the Inter-Noise 2011, 2011, vol. 5, pp. 4496–4501.

[19] A. L. Brown and D. Tomerini, “Distribution of the noise level maxima from the pass-by of vehicles in urban road traffic streams,” *Road and Transport Research*, vol. 20, no. 3, pp. 41–54, 2011.

[20] *Commission Directive (EU) 2015/996 of 19 May 2015 establishing common noise assessment methods according to Directive 2002/49/EC of the European Parliament and of the Council (Text with EEA relevance)*, vol. OJ L. 2015.

[21] J.-F. Hamet, F. Besnard, S. Doisy, J. Lelong, and E. Le Duc, “New vehicle noise emission for French traffic noise prediction,” *Applied Acoustics*, vol. 71, no. 9, pp. 861–869, Sep. 2010, doi: 10.1016/j.apacoust.2010.05.003.

[22] G. Rey Gozalo and J. M. Barrigón Morillas, “Analysis of Sampling Methodologies for Noise Pollution Assessment and the Impact on the Population,” *Int J Environ Res Public Health*, vol. 13, no. 5, 11 2016, doi: 10.3390/ijerph13050490.

[23] J. M. Barrigón Morillas, D. Montes González, and G. Rey Gozalo, “A review of the measurement procedure of the ISO 1996 standard. Relationship with the European Noise Directive,” *Science of The Total Environment*, vol. 565, pp. 595–606, Sep. 2016, doi: 10.1016/j.scitotenv.2016.04.207.

[24] International Organization for Standardization, “ISO 1996-2:2007 (Acoustics -- Description, measurement and assessment of environmental noise -- Part 2: Determination of environmental noise level),” *Standard*, pp. 40–40, 2007.

[25] K. Heutschi and K. Eggenschwiler, “Schallpegelmess-Apps,” Empa, 5214.018547, 2018.

[26] M. Strauss, J. Carnahan, and R. Ruhl, “The Accuracy of Pedestrians in Estimating the Speed of a Moving Vehicle,” SAE International, Warrendale, PA, SAE Technical Paper 2013-01–0785, Apr. 2013.

[27] S. Shilton and F. A. Ledee, “Conversion of existing road source data to use CNOSSOS-EU,” p. 6, 2015.

[28] A. Can and P. Aumond, “Estimation of road traffic noise emissions: The influence of speed and acceleration,” *Transportation Research Part D: Transport and Environment*, vol. 58, pp. 155–171, Jan. 2018, doi: 10.1016/j.trd.2017.12.002.

[29] W. Zamora, C. T. Calafate, J.-C. Cano, and P. Manzoni, “Accurate Ambient Noise Assessment Using Smartphones,” *Sensors (Basel)*, vol. 17, no. 4, p. 917, Apr. 2017, doi: 10.3390/s17040917.