Testing the accuracy of smartphones and sound level meter applications for measuring environmental noise

Enda Murphy* and Eoin A. King^

*School of Architecture, Planning & Environmental Policy, University College Dublin, Ireland

^Department of Mechanical Engineering, University of Hartford, USA

Abstract. This paper reports on experimental tests undertaken to assess the capability of noise monitoring applications to be utilised as an alternative low cost solution to traditional noise monitoring using a sound level meter. The methodology consisted of testing 100 smartphones in a reverberation room. Broadband white noise was utilised to test the ability of smartphones to measure noise at background, 50, 70 and 90 dB(A) and these measurements were compared with true noise levels acquired via a calibrated sound level meter. Tests were conducted on phones using the Android and iOS platforms. For each smartphone, tests were completed separately for leading noise monitoring apps culminating in 1472 tests. The results suggest that apps written for the iOS platform are superior to those running on the Android platform. They show that one of the apps tested – SLA Lite - is within +-1 dB of true noise levels across four different reference conditions. The results also show that there is a significant relationship between phone age and its ability to measure noise accurately. The research has implications for the future use of smartphones as low cost monitoring and assessment devices for environmental noise.

Keywords: Noise Measurement Applications; Environmental Noise; Crowd Sourced Noise Monitoring; Smartphones

Introduction and context

Smartphones have become a ‘must have’ for the majority of adult citizens in world’s developed nations. As of October 2014, 64 per cent of US adults own some form of smartphone [1]. To demonstrate the rapidity with which smartphones have infiltrated the US market, the corresponding figure for the spring of 2011 was 35 per cent [2]. Internationally, more recent research covering 32 countries estimates that 80 per cent of internet users own a smartphone. Of those, 54 per cent of phones utilise the Android operating system, 16 per cent operate the iOS and the remaining come from alternative operating systems such as Windows among others [3].

The development of smartphone technology and its impact on environmental noise studies has only recently begun to receive some attention in the academic literature. There are some studies which suggest that smartphones are capable of replacing traditional noise assessment devices such as sound level meters (SLMs) in the not too distant future. Kanjo [4] has outlined the possibility of developing a mobile phone platform for measuring noise in cities and highlights the potential of such avenues for the future. Similarly, D’Hondt et al [5] have demonstrated the possibility of smartphone-based noise apps to be utilised by ordinary citizens as a form of crowd sourced participatory noise assessment in cities. Studies such as these suggest that the future of noise assessment, whether it is in cities or elsewhere, will likely be tied closely to developments in smartphone and other forms of innovative mobile technology that are easily and relatively affordably accessed by ordinary citizens, especially in developed nations. A key challenge for noise mapping studies, in particular, is determining the accuracy of any smartphone based approach and to shed light on the margin of error that might be associated with the substitution of smartphones for sound level meters in future real world settings.

The current paper is concerned with trends in the development of smartphones and associated applications for the measurement of environmental noise specifically. There are only a small number of studies which have investigated issues that are relevant to the current research. Perhaps the most relevant is a recent study by Kardous and Shaw [6]. They tested the accuracy of 10 iOS and 4 Android apps for measuring noise in occupational settings on 8 smartphones and one tablet. Their research found that the iOS noise app – SoundMeter, developed by Faber Acoustical – has the best agreement in A-weighted sound levels (-0.52) with reference values while three other apps for the iOS were within +- 2dBA of reference values. This led the authors to conclude that devices running the iOS, in particular, had significant scope to be used as assessment devices for occupational settings. What is also interesting is that their research found that devices running the Android operating system were inadequate for the same purpose because they were 'built by several different manufacturers and that there is a lack of conformity for using similar microphones and other audio components in their devices' [6, p.192]. The focus of previous work by Kardous and Shaw was on examining the accuracy of smartphone apps rather than the smartphones themselves. Although they did offer some insights about the relationship between phone model and measurement accuracy, the sample of phones they used for testing was somewhat limited in scope (3 iPhone models and 5 Android devices).

Similarly, the work of Nast et al [7] tested five apps but only one phone - the iPhone 4S - thereby essentially controlling for the phone model in their analysis of noise measurement applications. Thus, their work provides no insight into the role of the smartphone hardware in producing accurate noise measurements or otherwise. Moreover, their tests did not utilise pink noise and/or white noise thereby limiting the spectral variability of the testing conditions to specific octave band analysis. Nevertheless, their results showed that for all apps tested, the results varied widely from that measured using a Type 1 SLM. The authors concluded that, with the exception of the Sound Meter App by Faber Acoustical, 'SLM apps are best used for entertainment purposes, as they are not accurate as SLMs…' [7, 253-254]. Indeed, their work pointed to large errors and nonlinearities at high sound levels, drawing into question the utility of apps for occupational purposes.

Within the foregoing context, the current paper builds on previous work which has sought to analyse the suitability of smartphones for use as a substitute for traditional SLMs. Whereas related studies has tended to place focus on the smartphone apps themselves, this research focusses not only on testing the leading apps on two leading platforms – iOS and Android - but we also test a much wider range of smartphones than has been tested in similar studies to date. In this regard, we are seeking to identify statistically significant differences in the ability of different smartphone models to measure noise accurately or otherwise using the same app while also assessing the suitability of the apps themselves and the platform being utilised to host them. The research also examines the relationship between smartphone age and measurement accuracy.

Methods

A representative sample of the most popular smartphones on the University of Hartford campus was acquired by asking students to volunteer their device for testing. In total 100 smartphones were tested; 65 were on the iOS platform while the other 35 were Android-based. A list of the phone manufacturers and individual models tested is presented in Table 1. For each iOS-based phone, four leading apps were tested while three apps were tested for each Android phone. This discrepancy was due to one app being taken down from the Google Play store after a small number of tests had been completed and because of this it was removed from the testing agenda. For an app to be included in the testing it had to satisfy certain criteria. These included: (1) being able to report A-weighted sound levels; (2) being able to report the sound level as a numeric value and (3) being either free or cost less than $5.00 While some apps allow for manual calibration of the in-built microphone prior to measurement, this was not completed for our experimental tests in order to simulate a typical real world situation. This conforms to the approach taken for similar testing studies [6, 7]. Table 2 provides a full list of the apps tested for our study – 7 in total - for the iOS and Android phones, the developer and version. All of the apps tested met our selection criteria and all were commercial apps. No tests were conducted on Windows-based devices given the dominance of iOS and Android phones of the smartphone market.

Table 1. Phones models tested and their frequency

Brand	Number
iPhone (4, 4s, 5, 5s, 5c, 6, +)	65
Galaxy (Note 2, Note 3, s3, s3 slim, s3 mini, s4, s4 active, s5,	24
Google (Nexus 5)	2
HTC (One, One Mini 2, M8)	4
LG (VS870, g2)	2
Motorola (Droid 2, Droid MAXX, Moto X 2nd gen.)	3
Total	100

Table 2. Smartphone apps selected for testing

Name	Developer (Price)	Web Link
Sound Level Analyzer Lite (iOS) version 1.3	Toon,LLC (€4.99)	https://itunes.apple.com/us/app/sound-level-analyzer/id886109671?mt=8
SPLnFFT (iOS) version 1.1	Fabien Lefebvre (HK€28)	https://itunes.apple.com/hk/app/splnfft-noise-meter/id355396114?mt=8
Decibel Meter Pro (iOS) version 2.05	Performance Audio (€0.99)	https://itunes.apple.com/ie/app/decibel-meter-pro/id382776256?mt=8
UE SPL (iOS) version 2.1.1	Logitech Inc. (Free)	https://itunes.apple.com/us/app/ue-spl/id332300068?mt=8
Sound Meter (Android) version 1.6	Smart Tools co. (Free)	https://play.google.com/store/apps/details?id=kr.sira.sound&hl=en
Noise Meter (Android) version 2.1	JINASYS (Free)	https://play.google.com/store/apps/details?id=com.pjw.noisemeter&hl=en
Decibel Pro (Android) version 1.4.22	BSB Mobile Solutions Tools (€4.99)	https://play.google.com/store/apps/details?id=bz.bsb.decibel.pro

For our experimental set up, we used broadband white noise in a 125 m^3 ISO 3741 [8] compliant reverberation room. This source was generated through Brüel & Kjær's Pulse Measurement System, version 18.1 and was played through a Type 4292-L OmniPower dodecahedron loudspeaker located in the centre of the room. The output voltage was adjusted in Pulse to produce a uniform sound field at 50 dB(A), 70 dB(A), and then 90 dB(A). These values were initially confirmed using both a rotating microphone boom fitted with a diffuse field microphone as well as a calibrated Brüel & Kjær Type 2250 SLM. Background noise was measured on each test day and was found to be 27 dB(A) in the reverberation room. Testing was conducted over 10 separate days. The diffuse sound field generated in the reverberation room meant that the precise location and size of the smartphone in the room did not influence the results of the study in any way. However, during measurements, phones were handheld at shoulder height by the same two individuals for the entire series of testing[1]; all phone covers were removed prior to testing to avoid any interference with the microphone. We collected a single measurement for each app at each test level (background, 50, 70 and 90 dB(A)). As an experimental precaution, the room was tested immediately before and after each testing schedule to ensure that the room acoustics remained consistent across testing schedules. The adoption of a handheld approach for testing differs from previous studies which utilised a tripod [6, 7]. The reason for this is that we were keen to attempt to simulate how phones would actually be utilised by the general public in a laboratory setting. Each phone was tested at background, 50 dB(A), 70 dB(A), and 90 dB(A) levels. Unlike other studies which tested smartphones for weighted and unweighted sound levels, we focussed only on the ability of phones to measure A-weighted sound level measurements given our interest in the capability of the devices for measuring environmental noise.

While recruiting students with smartphones, they were also asked a series of questions about their phone prior to testing taking place. Questions were asked about the precise make and model of the phone[2], the operating system[3] and the age of the phone in one of five categories[4]. This allowed for additional analysis of the test data with respect to these specific variables.

For data analysis, we performed ANOVA and t-tests to assess the difference in mean values associated with each platform (iOS/Android), across apps and phone models. In addition, descriptive statistics were utilised to determine operating system, app and phone performance while standard boxplot analysis was used to assess the variability in measurement scores across apps and phone models. In order to isolate the impact of certain variables on measurement outcome, sequential regression analysis was also undertaken. Sequential regression is utilised to determine the impact of independent variables on smartphones measurement differential from reference and allows the user to enter variables or sets of variables into the regression equation after other variables have been controlled for as a separate block. This allows the researcher to determine if such variables are contributing significantly to the prediction of the measurement outcome.

[1] For all tests, there was one individual testing in the reverberation room and one located outside operating the Pulse system for all tests.
[2] If students did not know the exact model, it was identified prior to testing.
[3] iOS or Android.
[4] These were: <6 months, 6-12 months, 1-1.5 years, 1.5-2 years, 2+ years.

Results

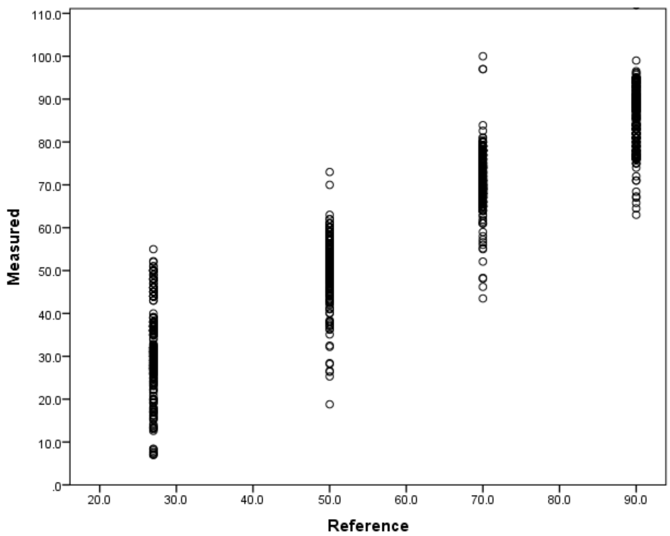
Table 3 shows descriptive statistics of the mean difference between measured values using smartphones and the pre-specified reference values. It can be seen that at the 50 and 70 dB(A) reference conditions the mean difference in app measurement from reference conditions is 2.09 and 1.33 respectively while at the other reference conditions the measurement results are more variable. Indeed, the results also show that the apps are less efficient at measuring at background and high noise levels; the applications over measure the true noise level by 5.33 dB(A) at background and underestimate it by 3.57 dB(A) at 90 dB(A). However, at noise levels between background and 90 dB(A) they do an adequate job of measuring to within an acceptable degree of error which is typically +-2 dB(A). The fact that the measurement apps do a poorer job of accurately measuring at high noise levels is a concern given that environmental noise at higher levels is the key area of concern from a public health perspective.

To explore the data variability, a scatterplot comparing measured values with pre-specified reference conditions – 27 dB(A)background, 50 dB(A), 70 dB(A), 90 dB(A) – was completed and is shown in Figure 1. It demonstrates the extent of variation in measured versus reference values across the full range of measurements. The high degree of variation between measured and reference scores suggests that the reliability of smartphones for measuring environmental noise depends to a significant degree on having a relatively large number of sample data points rather than a few isolated measurements.

Table 3. Descriptive statistics showing smartphones testing results by reference condition (dB(A))

Reference (dB(A)	**N**	**Mean difference from reference**	**S.D.**	**S.E.**	**Range**
Background (27)	368	5.33	9.64	0.50	48.0
50	368	2.09	6.50	0.34	54.2
70	368	1.33	6.27	0.33	56.5
90	368	-3.57	6.99	0.36	51.0

Figure 1. Scatterplot of reference versus measured noise values using smartphones



Turning specifically to an analysis of the relationship between measurement accuracy and the phone platform being utilised, an independent samples t-test was performed to examine whether a significant difference existed between the mean measured values across the iOS and Android platforms at each of the four reference conditions.[5] The results are presented in Table 4. They demonstrate a significant difference in mean scores for the two platforms for all but one reference condition – 70 dB(A). With the exception of the 70 dB(A) reference condition, the results show that Android devices have a mean value which is closer to the true noise level for all other reference conditions. However, these results come with a caveat because they also demonstrate that Android devices are associated with higher standard deviation values relative to the iOS indicating poorer reliability in terms of measurement consistency.

Table 4. Relationship between measurement accuracy and the phone platform[6]

	Platform	N	Mean	Mean Difference	Std. Deviation	Std. Error Mean	t	P-value
B'grd (27 dBA)	iOS	263	35.436	11.35	7.57	0.46	10.93	0.00*
	Android	105	24.219		9.50	0.92		
50 dBA	iOS	263	53.407	5.31	3.70	0.22	5.45	0.00*
	Android	105	48.291		9.71	0.94		
70 dBA	iOS	263	71.253	0.67	4.67	0.28	0.71	0.48
	Android	105	70.856		9.13	0.89		
90 dBA	iOS	263	87.994	5.50	5.84	0.36	7.29	0.00*
	Android	105	82.491		8.02	0.78		

A further interesting issue to investigate is the relationship between the phone manufacturer and measured noise values. Table 5 shows the mean difference from reference values by phone brand. The results show that the best performing phone brand is HTC with only a 0.33 difference from the true noise level. Samsung is the next best followed closely by Apple. Table 5 also shows a break out of the results for each manufacturer by reference condition. It shows that at the background reference condition the HTC phone performs best (the mean difference from reference is 2.06 dB(A)) while at the 50 dB(A) reference condition the Google phone performs best. The test results also show that at the 70 and 90 dB(A) reference conditions the iPhone (1.52 dB(A)) and the HTC phone (-1.61 dB(A)) respectively perform best. Thus, the HTC phone performs best at two of the four reference conditions.

The wide variation in results for Android phones is interesting because it demonstrates that the phone brand being utilised for measurement has a significant bearing on its ability to measure noise accurately when the same app is being used. This implies that considerable variation exists in the quality of the hardware components among Android phones. More specifically, it points to a high degree of variation in the quality of MEMS microphone components used in different devices. However, the results of a sequential regression to examine the effect of make on the ability of a smartphone to measure noise accurately was not statistically significant (p=0.68) when other factors were controlled for, such as phone age, platform and the app being utilised. Thus, further testing is needed to examine more extensively the relationship between phone brand and ability to measure noise accurately.

Turning our attention to specific phone apps, the results found that the best app was on the iOS platform (SLA Lite) with the second best app associated with the Android platform (Sound Meter). Overall, the testing regime showed that iOS apps over measured

[5] For all t-tests, the data was first test for homoscedasticity using Levene's test and the appropriate p-value was taken depending on the result.

[6] Asterisks denotes significant at the 0.05 alpha level.

true noise levels by an average of 2.93 dB(A) (N=1052) while apps on the Android platform under measured noise levels an average of 2.79 dB(A) (N=420). While this suggests that apps on the Android platform were slightly more successful at measuring true noise levels, the high standard deviation value associated with Android apps (SD=9.58 dB(A)) highlights the greater degree of variability associated with measurement apps on that platform; in short, apps on the iOS (SD=6.81 dB(A)) were more consistent and less erratic in terms of their measurement values. Thus, while the results show that Android devices have mean values closer to true noise levels at most reference conditions, the best performing and most consistent apps in terms of measurement reliability are on the iOS platform

Table 5. Mean difference from reference conditions and phone manufacturer

		N	Mean	Standard Deviation	Range
iPhone	Background (27)	263	8.57	7.57	30.50
	50	263	3.61	3.70	20.00
	70	263	1.52	4.68	26.00
	90	263	-2.01	5.84	29.50
	Total	1052	2.92	6.80	51.00
Galaxy	Background (27)	72	-4.49	8.38	32.20
	50	72	0.96	7.30	36.80
	70	72	2.42	7.02	39.00
	90	72	-7.32	7.65	49.50
	Total	288	-2.10	8.40	55.50
Google	Background (27)	6	-3.15	5.34	11.80
	50	6	-16.85	13.80	39.20
	70	6	-7.45	14.00	34.80
	90	6	-13.05	8.69	25.40
	Total	24	-10.12	11.62	39.50
HTC	Background (27)	12	2.06	9.42	35.10
	50	12	-4.14	7.78	27.50
	70	12	5.00	6.75	22.40
	90	12	-1.61	6.95	21.90
	Total	48	0.33	8.32	42.70
LG	Background (27)	6	16.15	5.96	15.50
	50	6	6.43	5.62	14.30
	70	6	6.27	5.12	14.50
	90	6	-4.08	4.65	11.20
	Total	24	6.19	8.85	33.90
Motorola	Background (27)	9	-7.94	5.36	12.80
	50	9	-15.14	5.78	14.60
	70	9	-15.27	5.75	14.80
	90	9	-15.48	5.90	15.20
	Total	36	13.45	6.33	23.20

A detailed breakdown of the differential between measurement values for individual apps and reference conditions for all tests is provided in Table 6. With regard to the performance of specific apps, the best performer in this regard was SLA Lite. Across the four reference values, the app had an average under measurement of only -0.37 dB(A) and was consistently within 1 dB(A) of the true noise level at each reference condition which compares very favourably with SLMs. Moreover, the standard deviation associated with measurements using SLA Lite was small (1.41) highlighting the consistency of the app in terms of its measurement accuracy. Despite the ability of the app to measure accurately, one of the main drawbacks is its inability to log data over a specified time period; however, upgrading to SLA for a fee of €4.99 enables data logging.

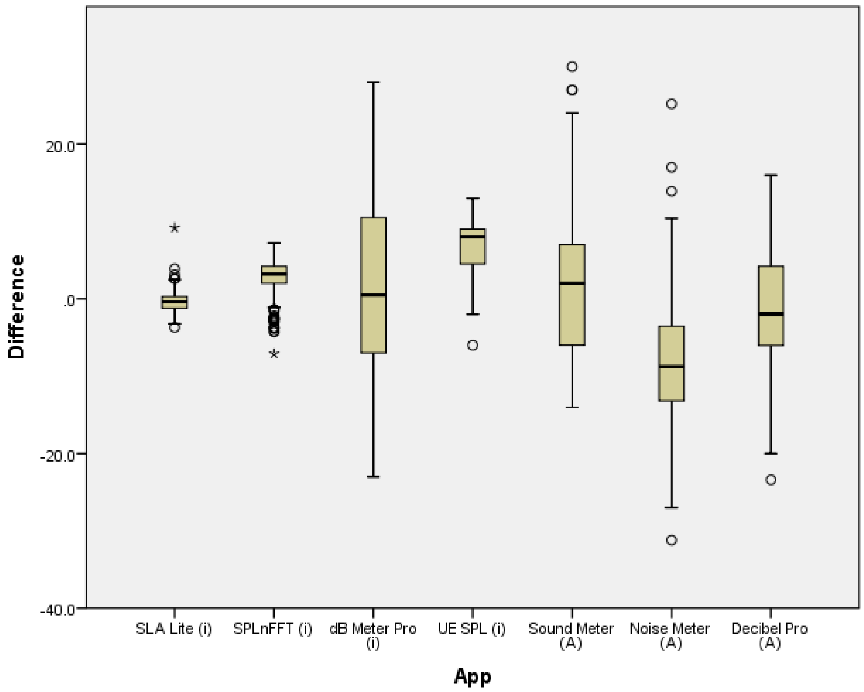
Table 6. Performance of individual apps compared to reference conditions

		N	Mean	Standard Deviation	Range
SLA Lite (i)	Background (27)	65	0.57	1.11	6.40
	50	65	-0.76	1.21	6.20
	70	65	-0.55	1.68	12.00
	90	65	-0.75	1.14	4.90
	Total	260	-0.37	1.41	12.90
SPLnFFT (i)	Background (27)	66	3.97	1.20	6.90
	50	66	2.90	1.68	12.70
	70	66	2.31	2.55	10.90
	90	66	1.79	3.02	10.80
	Total	264	2.74	2.36	14.30
dB Meter Pro (i)	Background (27)	66	19.92	2.95	20.00
	50	66	4.23	2.97	20.00
	70	66	-3.38	2.80	19.00
	90	66	-10.94	3.10	17.00
	Total	264	2.45	11.81	51.00
UE SPL (i)	Background (27)	66	9.70	1.47	9.00
	50	66	8.02	1.56	10.00
	70	66	7.68	1.77	10.00
	90	66	1.89	2.19	11.00
	Total	264	6.82	3.43	19.00
Sound Meter (A)	Background (27)	35	3.60	6.00	31.00
	50	35	3.11	8.77	36.00
	70	35	4.80	9.34	44.00
	90	35	-3.77	9.40	37.00
	Total	140	1.93	9.04	44.00
Noise Meter (A)	Background (27)	35	-6.73	9.77	45.20
	50	35	-7.49	8.87	41.20
	70	35	-5.09	7.82	40.40
	90	35	-13.65	5.64	30.30
	Total	140	-8.24	8.71	56.40
Decibel Pro (A)	Background (27)	35	-5.21	8.99	36.00
	50	35	-0.75	8.58	34.40
	70	35	2.86	7.10	30.50
	90	35	-5.11	4.22	19.30
	Total	140	-2.05	8.11	39.40

Turning to the Android platform, the most accurate app was Sound Meter which under measured noise by 1.93 dB(A), under the typically acceptable error threshold of +-2 dB(A). However, across all reference levels it can be seen that the average differential from the true noise level is between 3-4 dB(A). It can be seen also that despite the mean values for Android apps holding up well when compared to true noise levels, the standard deviation values associated with most Android apps are typically a lot higher than those associated with iOS apps. This suggests a lack of measurement consistency for Android apps when compared to corresponding apps for the iOS.

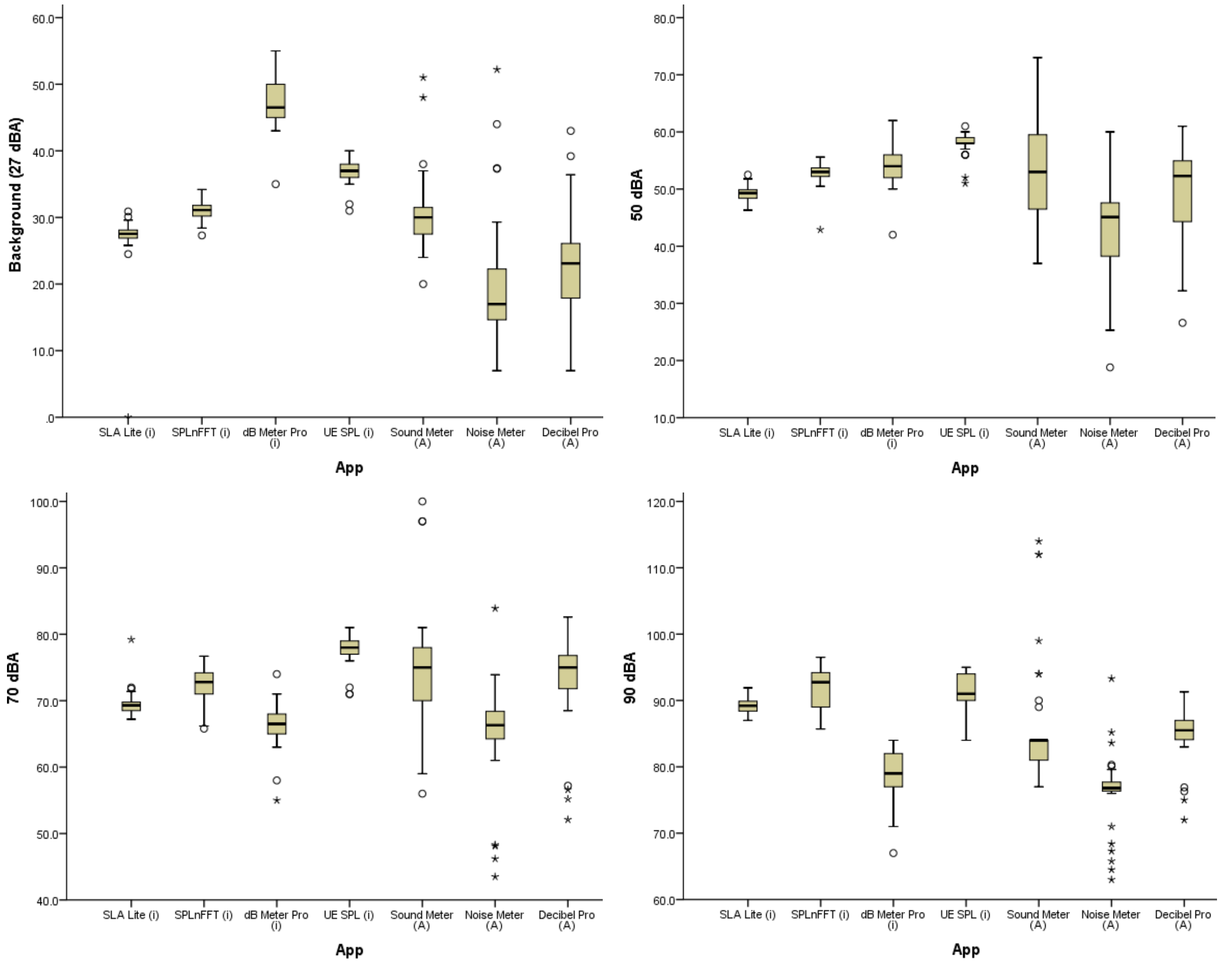
The boxplot in Figure 2 shows a visual breakdown of the distribution of the difference between reference and measured data by noise measurement application while Figure 3 shows a similar visual breakdown but for each specific reference condition – background, 50, 70, and 90 dB(A). It can be seen that, with the exception of dB Meter Pro, the applications with the lowest degree of variability are all on the iOS platform with those on the Android platform associated with more varied data distributions. Indeed, apps such as SLA Lite and SPLnFFT, in particular, have data ranges which are considerably narrower than other apps indicating that those apps are more consistent in terms of their ability to measure environmental noise accurately. The more detailed breakdown by specific reference condition shows that the highest degree of variabilty lies at the background refence condition and also shows that Andrpid apps are associated with a higher degree of variability at all reference conditions. Moreover, at the 90 dB(A) condition there exists a signifcant number of data outliers[7] which suggests a more erratic pattern of measureement at higher noise levels when compared with the 50 and 70 dB(A) conditions.

Figure 2. Boxplot showing data distribution of difference between reference and measured values by smartphone application



[7] Outliers are indicated by asterisks and circles.

Figure 3. Boxplots showing measurement results by smartphone application at various reference conditions



In order to examine the relationship between app and measurement accuracy more concretely, a sequential regression was undertaken to examine the effect of noise measurement application on the ability of a smartphone to measure noise accurately. The results of the regression were statistically significant ($p=0.00$) when other factors were controlled for such as phone brand, platform and the age of the smartphone indicating a statistically significant relationship between the app being used and measurement accuracy highlighting the importance of choosing the correct app for environmental noise measurement.

Table 7 shows the results of smartphone testing by the age of phone broken down into five categories. They show that the newest phones are also the phones that measure noise closer to true noise levels; phones that are less than six months old have an average differential from reference noise levels of only 0.15 dB(A) when compared with smartphones that are more than two years old (2.76 dB(A)). Moreover, the results show that all phones perform poorly at measuring background noise irrespective of age. At both 50 dB(A) and 70 dB(A) phones that are less than six months old perform best while at 90 dB(A) phones that are more than two years old perform best. This suggests either that newer phones are equipped with better microphone technology when compared with older phones or that the performance of the microphone in smartphone devices deteriorates with age. While this trend is hardly surprising, what is counter-intuitive is the fact that the standard deviation values tends to decline with smartphone age. This implies that while younger phones are typically better on average at measuring true noise levels they are also less consistent in doing so when compared with older smartphones which are associated with measurements that have a tighter distribution around the mean. Indeed, the result of one-way ANOVA confirmed a significant difference between the mean values of smartphones by age category (p=0.00). In order to examine the issue further, a sequential regression was undertaken to examine the effect of phone age on the ability of a smartphone to measure noise accurately. The results were statistically significant (p=0.01) when other factors were controlled for, such as phone brand, platform and the app being utilised confirming a statistically significant relationship between phone age and ability to measure noise accurately.

Table 7. Descriptive statistics showing phone testing results by age of smartphone

		N	Mean	Standard Deviation	Range
< 6 Mths	Background (27)	145	4.18	10.16	43.80
	50	145	0.81	7.67	54.20
	70	145	0.32	6.71	53.80
	90	145	-4.69	7.13	46.20
	Total	580	0.15	8.61	61.20
6-12 Mths	Background (27)	76	5.19	9.12	43.00
	50	76	1.81	7.11	44.70
	70	76	1.62	7.71	53.50
	90	76	-3.32	8.09	51.00
	Total	304	1.32	8.55	54.00
1-1.5 Yrs	Background (27)	39	6.15	9.74	39.60
	50	39	3.05	4.31	15.10
	70	39	2.38	4.91	19.90
	90	39	-2.58	5.87	19.70
	Total	156	2.24	7.21	39.60
1.5-2 Yrs	Background (27)	76	5.76	9.60	42.00
	50	76	3.90	4.65	20.00
	70	76	2.58	4.84	25.40
	90	76	-2.93	6.64	32.00
	Total	304	2.32	7.44	53.50
2+ years	Background (27)	32	8.85	7.77	27.50
	50	32	3.12	3.40	10.20
	70	32	1.01	4.04	13.00
	90	32	-1.90	4.90	16.70
	Total	128	2.76	6.55	37.00

Discussion and Conclusion

The use of smartphones for measuring environmental noise, while currently in its infancy, has significant potential in the future to act as a form of crowd sourced noise monitoring. The use of everyday technology such as a smartphone to measure environmental noise has the potential to improve the monitoring of the sound environment in cities and the countryside alike but it potentially has the added advantage of engaging and indeed empowering citizens to contribute to monitoring the environment in which they live and work. Moreover, if smartphone-based noise measurement apps prove to be useful in the future, they could play an important role for mapping environmental noise in cities and assisting with the implementation of the EU Environmental Noise Directive and associated action planning [see 9, 10, 11]. The examination of the errors associated with smartphone-based noise apps is a useful first step in this regard.

Compared with previous studies that have tested the accuracy of smartphones for measuring noise [6, 7], this study includes a much more extensive range of testing. First, it tests 100 phones of various makes and models comprising 1472 tests. Smartphones from seven manufacturers were tested comprising 18 different Android phone models and 7 different iOS models. By virtue of the testing range, smartphones across a variety of age cohorts are included in the analysis thereby reflecting to a greater extent the population of smartphones in use among the general public. Second, we tested a range of leading smartphone apps across the iOS and Android platforms. While other studies have also completed similar testing, the testing of apps has not been completed across such a volume and variation of phone makes and models as is included in this study.

The accuracy of noise measurement apps varied widely relative to pre-specified reference levels. Overall, there is little doubt that iOS apps performed better than Android-based apps. While some Android apps performed better than those for the iOS in terms of mean differential from reference values (e.g. Sound Meter), the high degree of measurement variability associated with such apps renders their reliability questionable. What we can say is that if a large number of sample measurements are being taken then Android apps such as Sound Meter and Decibel Pro will tend to converge on a noise measurement level that is roughly within +-2 dB(A) of the true noise levels. However, in the absence of a large number of sample measurements, iOS apps such as SLA Lite and SPLnFFT should be utilised due to their ability to measure with less variability around the mean noise level. In fact, SLA Lite was the only app accurate to within +-2 dB(A) across all of the reference conditions – background, 50, 70 and 90 dB(A) - even though other apps such as SPLnFFT (iOS) and Sound Meter (Android) performed relatively well. Thus, as things stand currently we can conclude that noise apps are not quite ready to replace traditional SLMs but our results suggest the likelihood that as software and hardware technology improves there is ample scope for noise apps to perform an important role in crowd sourced environmental noise monitoring in the near future. The accuracy of the SLA Lite app clearly demonstrates that a combination of good hardware and software achieves noise monitoring results that are very accurate provided an adequate number of sample measurements are taken.

Another issue is the fact that three out of the seven apps tested reported average sound levels below reference levels. This is somewhat worrying because apps that incorrectly report noise levels below the true level are more problematic from a public health perspective. Adaptation of the precautionary principle informs us that it is better for an app to report noise levels slightly above the true level because at least then regulators have information which allows the public to be protected adequately. In this regard there was a greater tendency for Android-based apps to under report noise levels than iOS apps.

The technical specification of a Type 1 SLM includes a floor below which the device cannot measure sound (16.6 dB for A-weighting) [7]. Similar to results emerging from Nast et al's [7] study, our experiments demonstrate that the average differential from true noise level was greater than 5 dB(A) for all apps tested. An exception to this trend was the SLA Lite app which was within 0.15 dB(A) of the true noise level for the background reference condition. Indeed, the differential between measured and true noise levels was greatest for background noise suggesting that apps perform poorest for ambient noise measurement. However, this is not a significant problem given that noise at ambient levels does not typically pose a public health threat.

Finally, our results also suggest that the age of the phone has a bearing on its ability to measure noise accurately; on average, younger phones measure noise more accurately than older phones but with greater volatility. This is a results aspect of smartphone testing which has not been investigated previously in the academic literature. Whether this is due to the deterioration of microphone hardware over time due to everyday wear and tear or due to contemporary versions of noise apps which are coded more accurately for microphones in newer smartphones is unclear and requires more extensive testing. Also requiring further work is the relationship between phone manufacturer and noise measurement accuracy. In this regard, further research is needed to investigate whether microphones used by specific smartphone manufacturers is producing better measurement outcomes as is tentatively implied by the results for our study.

Acknowledgements

E. Murphy acknowledges the Irish Research Council (IRC) Starter Grant award which funded this work under the OASIS project. He also acknowledges the Fulbright Commission for provision of a Fulbright Scholarship to undertake the experimental work at the University of Hartford, CT, USA. E. Murphy also thanks the Department of Mechanical Engineering, University of Hartford for hosting his research stay and for providing the research infrastructure necessary to complete the experiments. E.A. King acknowledges the Department of Mechanical Engineering Faculty grant which provided research assistance for the project. We both thank Jake Schepro, Sean Rahusen and Lane Millar, students of the Acoustics Program at the University of Hartford who provided research assistance for the lab experiments. We also thank Prof. Bob Celmer, University of Hartford for constructive comments on an early draft of the manuscript.

References

[1] Pew Research Center (2014) 'Cell Phone and Smartphone Ownership Demographics', http://www.pewinternet.org/data-trend/mobile/cell-phone-and-smartphone-ownership-demographics/ (accessed 30 April 2015).

[2] Smith, A. (2015) 'U.S. Smartphone Use in 2015', http://www.pewinternet.org/2015/04/01/us-smartphone-use-in-2015/ (accessed 30 April 2015).

[3] GlobalWebIndex (2015) http://cdn2.hubspot.net/hub/304927/file-2301369304-pdf/Reports/GWI_Device_Summary_Q3_2014.pdf?submissionGuid=41ba1017-ee82-435b-8762-9d837a0cc5ac (accessed 30 April 2015).

[4] Kanjo, E. (2010) ‘NoiseSPY: A real-time mobile phone platform for urban noise monitoring and mapping’, Mobile Networks Applications, 15, 562-574.

[5] D’Hondt, E., Stevens, M., Jacobs, A. (2103) ‘Participatory noise mapping works! An evaluation of participatory sensing as an alternative to standard techniques for environmental monitoring’, Pervasive and Mobile Computing, 9, 681-694.

[6] Kardous, C.A. and Shaw, P.B. (2014) ‘Evaluation of smartphone sound measurement applications’, Journal of the Acoustical Society of America, 135, EL186-192.

[7] Nast, D.R., Speer, W., Prell, C.G. (2014) ‘Sound level measurements using smartphone “apps”: useful or inaccurate?’, Noise Health, 16, 251-256.

[8] International Organization for Standardisation (1999) ISO 3741-1999, Acoustics: Determination of sound power levels of noise sources using sound pressure – Precision methods for reverberation rooms.

[9] Murphy, E. and King, E.A. (2010) ‘Strategic environmental noise mapping: methodological issues concerning the implementation of the EU Environmental Noise Directive and their policy implications', Environment International, 36, 290-298.

[10] Murphy, E. and King, E.A. (2014) Environmental Noise Pollution: Noise Mapping, Public Health and Policy. Amsterdam: Elsevier.

[11] King, E.A., Murphy, E. and Rice, H.J. (2011) 'Evaluating the impact on noise levels of a ban on private cars in Dublin city centre, Ireland', Transportation Research Part D: Transport and Environment, 16, 532-539.