An Affordable Recording Instrument for the Acoustical Characterisation of Human Environments

Huub Bakker¹, Bruce Rapley², Rachel Summers³, Mariana Alves-Pereira⁴, Philip Dickinson⁵

1 School of Engineering and Advanced Technology, Massey University, Palmerston North, New Zealand (corresponding author)
2 Atkinson & Rapley Consulting, Palmerston North, New Zealand
3 School of People, Environment and Planning, Massey University, Palmerston North, New Zealand
4 School of Economic Sciences and Organizations (ECEO), Lusófona University, Lisbon, Portugal
5 Retired, Wellington, New Zealand

Corresponding author’s e-mail address: h.h.bakker@massey.ac.nz

ABSTRACT

Characterising acoustical environments where humans may be sleeping requires hi-fidelity recordings of the entire soundscape, which allows relevant post-processing of the data. The on-going Citizen Science Initiative for the Acoustical Characterisation of Human Environments (CSI-ACHE) is dedicated to capturing quality recordings of such human acoustical environments. This paper introduces the recording system created to capture such soundscapes. It provides the specifications for the device, as well as the implementation and the tests used to characterise the performance and calibrate against a Class 1 instrument.

This system has been marketed as SAM Scribe and fundamentally consists of a two-channel device that can measure at sampling rates up to 44.1 kHz, and that delivers the data streams via USB to a Windows notebook computer to store as uncompressed wav files to hard disk. GPS information is stored as metadata in the files, which also include a digital signature. Storage of several months of continuous recordings can be achieved using an external USB 3.0 hard drive. The system can accurately record from 0.5 Hz to 10 kHz and is simple to operate.

INTRODUCTION

This paper describes the creation of a sound recording system to support the Citizen Science Initiative on Acoustical Characterisation of Human Environments (CSI-ACHE) [1]. The CSI-ACHE seeks to gather high-fidelity recordings of various acoustic environments where citizens believe they may be suffering health impacts due to abnormal quantities of residential infrasound and low frequency noise (ILFN). This Initiative is the offspring of the decades-long conundrum regarding the issue of ILFN and public health [2, for example].

As explained in more detail elsewhere [3], the dBA metric used by the majority of regulations and guidelines provides no information on ILFN levels. The numerical information required to
(in the future) determine frequency-dependent dose-responses for ILFN is not being obtained. Families who complain of ILFN-induced health effects due to residential exposure rely on acousticians to quantify the acoustical energy in question. Acousticians, however, are often limited by the legislated guidelines and regulations that preclude the numerical quantification of ILFN [2,3]. As a result, levels of ILFN are not quantified, the residential acoustic environment is not properly characterised, and citizen complaints are (too) frequently labelled as psychosomatic in nature [4], or as some form of misophonia [5]. Many citizens worldwide have taken the matter into their own hands and have purchased Class 1 or Class 2 instrumentation to measure ‘sound pressure levels’ during the periods of time when they are most impacted. Much of the instrumentation on the market is built to satisfy the requirements of legislated guidelines and regulations, and as a result, ILFN is not properly quantified. Moreover, much of the instrumentation is not user-friendly for the layperson, and can cost upwards of USD $6000 per channel.

The instruments available on the market are ill suited for citizens to use in their homes for extended periods of time (months), with the goal of numerically quantifying ILFN. Providing an affordable recording instrument for these purposes became the objective of CSI-ACHE researchers. The resulting unit, now marketed as SAM Scribe, is the outcome of such efforts.

Calibration

As for all research and forensic work, calibration is necessary to ensure that measurements can be made to a known accuracy and precision. In the case of sound level meters, calibration is defined in terms of a ‘typing’ of the instrument to either Class 1 or Class 2 [6], this performed by a suitable accredited body and granted for the given model of instrument. The typing stipulates the allowable limits for a number of factors, most important being the frequency response curve for the instrument, the linearity, including the microphone. The drift in the sensitivity of the instrument with time and conditions is catered for by use of a calibrator that provides an accurate 1 kHz, 94 dB tone. The calibrator has, itself, a typing as Class 1 or Class 2 and will also have (a smaller) variation with time and temperature.

This type of calibration requires that the instrument have high linearity over the range of frequencies to measured as usually only derived measurements are recorded, such as the $L_{eq}$, $L_{peak}$, and $L_{Aeq 10min}$ parameters.

The recording system for the CSI-ACHE will not be typed to this standard, for reasons of cost, time and accuracy of the equipment, but more importantly because it is a sound recorder and not a sound level meter. This means that the actual sound signal, as measured by the microphones as well as the analog and digital processing circuitry, is available for post facto correction of known non-linearities in the system.

Recording of a calibrating tone, from a suitable calibrator, can capture the absolute sensitivity of the system and such a recording should be part of the normal recording process, just as it is for sound level meters. Further corrections can be made for non-linearities in the frequency response of the system—if the response curve of the particular microphone(s) is known—after the recorded signal has been deconstructed into its individual frequency bands by FFT or narrow-band filter bank. Provided that the parameters of the system did not drift too much with time and atmospheric conditions, such a scheme could achieve high-linearity even if the system itself was not of high linearity.
SPECIFICATIONS

From the stated objective for a recording system to support CSI-ACHE, it is possible to derive a set of specifications that must be met by such a device. This can then drive the design, implementation and testing processes. What follows is a list of the main requirements of system including relevant considerations.

Ease of Use

The equipment and software is not intended for use by experts but by citizens who may not have any experience in working with industrial sound recording equipment. Added to this is the fact that industrial equipment is not primarily designed for ease of use but for advanced functionality.

Furthermore, many of the citizens who will use the equipment may be suffering from sleep deprivation caused either by a malignant soundscape, or their response to what they consider to be a malignant soundscape. Therefore, the ease of use becomes the overriding requirement for the equipment and the recording software.

‘Full-Spectrum’ Capture

The recording system must be able to capture the acoustical environment of interest to humans. This goes beyond the audible range of 20 Hz to 20 kHz as low-frequency sound and infrasound have also been posited as having an affect on humans [7-10] although, not necessarily via the standard audio sensory pathway. Therefore the range should extend from about 1 Hz to 20 kHz.

Flat Frequency Response

In order to provide sufficient fidelity in recording, the frequency response of the system should be flat over this frequency range. The current requirements for Class 1 sound level meters have tolerance limits of, for instance, +3 dB and -∞ dB at 10 Hz [6]. This is not stringent enough at this lower end of the range. To faithfully capture infrasound down to 1 Hz and below the ‘flat’ region should extend to here.

Note that the absolute sound level accuracy of the system is not paramount as it would be in Class 1 sound level meter. The intent of the initiative is to provide characterisation of the acoustical environment not to provide legal proof of compliance.

Include Wind Shields

Wind shields are required by a many acoustical standards [11]. Their presence is particularly required at deep-infrasound frequencies [12]; at roughly 1 Hz the microphone is acting as a micro-barometer where the slightest movement of air—even from an impulsive, non-acoustical source such as a window opening—can swamp any underlying acoustic signal.

Onsite Calibration

The system should be capable of calibration using a standard 1 kHz, 94 dB tone. The system does not need to achieve absolute calibration in itself but can simply record the tone, allowing for calibration of the recording post facto.

Twin Channels

The system should have more than one recording channel so that multiple microphones can be used synchronously. This allows synchronous capture of inside/outside comparisons, sound intensity measurements and stereo recording (as experienced by someone standing at the microphone location). The presence of multiple channels also amortises the cost of some of the system components, thereby reducing the system cost per channel.
Flexible Location

A multi-channel system may require that microphones be placed at different locations, tens of metres apart and at similar distances from the recorder itself. Furthermore, the recording system may need to be at some position remote to the microphones for other reasons, e.g., to be kept out of the weather or to allow for GPS reception.

Standard Format

The recording format should be a standard of some form that allows the inclusion of the metadata associated with location, time and other ancillary data. The format should be in an uncompressed or lossless form. The format must be playable by common sound players or analysers (such as the free Audacity program).

Location Recorded

The system should allow confirmation of the location of the recording and the actual time of the recording, independent of the recording system and the operator. The ability to identify, categorically, the time and place of the recording, post facto, is a boon to any researcher and having this determined independently of the citizen, and their possible vagaries, is a must.

Software Version Recording

The recordings should allow confirmation of the system software and hardware used to make the recording. As any developer will know, that ability to identify the software version of software in the field is a prerequisite for diagnosing an issue with the working of the software or equipment.

Chain of Evidence

The recordings should allow verification of the location, timing and integrity of the recording contents at a later date. These are required to provide ‘chain of evidence’ for the recordings, i.e., that they were made at the time and place specified and that the contents of the recording are unaltered after the fact. With even the best intentions, citizens are quite capable of making mistakes in the naming or handling of files after recording, never mind the suspicions they might face that they had tampered with the sound files to their own advantage.

Low Noise Floor

The system should have a noise floor sufficiently low that normal recordings not be compromised. The types of human environments in which the system is to be used include rural areas where the sound levels can be below 15–20 dB. At particular frequencies this can be much lower. It is therefore important that the noise floor of the system is well known throughout the frequency range and situations where this occurs are identified easily.

The noise floor of microphones and electronics varies with frequency so a single sound level may not be the best measure.

Large Storage Capacity

The system should be capable of storing continuous recordings for more than one month at a time at the highest sampling rate. This would be roughly 4 TB at a 44.1 kHz sampling rate for two channels.

Low Cost

The cost of the system should be such that it can be purchased or hired by citizens of average means. The authors set this level as less than US$3000.
DESIGN

The SAM Scribe system is built in two separate sections; 1) the analog section that deals with the measurement of sound, its processing and digitising, and 2) the digital section that deals with the formation of the recording files including their metadata.

Figure 1 shows the SAM Scribe Mk 1 system components. (A Mk 2 version had been prototyped at the time of writing this paper but was not in commercial production.) The system has a notebook computer to perform digital processing and storage (centre, top) and a SAM Scribe FS unit with two microphones to perform the analog processing (left).

Analog Section

The analog section is built around the SAM Scribe FS (full-spectrum) unit. This uses an off-the-shelf, 16-bit, Cmedia chip [13] to provide for the digitisation of the analog sound signal and transmission over USB. The C-Media chip was found to be one of the very few available that does not include an inbuilt high pass filter set to remove frequencies below 20 Hz. The presence of such a filter in a sound chip, of course, invalidates it for this task.

The SAM Scribe FS unit will process frequencies down to 0.1 Hz but includes a switch to add a high pass filter to remove frequencies below 1 Hz. Measurement of sound at such very low frequencies suffers from excessive response to small air movements, 'micro-baroms,' such as the closing of a door. This can lead to excessive clipping of the sound signal.

Figure 1: Main components of the SAM Scribe recording system (Mk 1 version). The SAM Scribe FS unit is shown on the left, with the microphones below and the notebook computer in the centre, top. The calibrator is shown in the centre. The USB hub at the right has an external hard drive (left), a GPS (centre) and the security hasp (right) attached.

All connectors to and from the unit, including the USB connection, use standard ‘microphone’ connectors [14] for robustness. All buttons are vandal-resistant with illuminating LEDs to confirm the position. Unlike normal toggle switches, the buttons can only be operated using
something like the tip of the little finger; therefore they are very unlikely to be inadvertently toggled.

The twin microphones use inexpensive capsules that, nevertheless, have an excellent frequency response. They are colour coded to match the cables and inputs to the Scribe FS unit so there is a high probability that they will be attached to the same channel every time. They use the current-loop protocol, ICP, used in many acoustical microphone systems.

Simple microphone clamps are provided with the system to allow appropriate clamping to microphone stands (supplied by the user).

Two screened cables are supplied at 5 m and 15 m lengths.

The unit provides two controls on the analog processing of the signal that can be switched for each channel individually. The first is a 1 Hz, high-pass filter that can be used when large ‘micro-baroms’ are present, causing clipping of the signal. The second is a 20 dB gain boost to be used when the sound level is low. This is an important feature given the relatively low dynamic range provided by the 16-bit A/D converter.

Digital Section

The digital section mainly comprises of a notebook computer running Windows 10 and proprietary software. Three peripherals are included; a USB 3 hub, a GPS unit and a security hasp. An MD5 digest, or signature, is created for each file, which is then encrypted—and can be decrypted—by a security hasp. This allows for chain-of-evidence verification for each sound file.

The software itself, SAM Scribe, is written to provide the simplest interface possible that still provides the functionality required, see Figure 2. The only inputs on the main screen are: the base name of the recording that forms part of every filename, the sampling frequency (11025 Hz, 22050 Hz, or 44100 Hz), the Record button and a button to perform a file verification. The rest are indicators of the program status.

![Figure 2: Screenshot of the SAM Scribe recording software. The base name for the recordings is entered at the top right. The sampling rate is chosen at the top right. The other inputs are at the bottom. All indicators are at the right of the screen.](image)

The sound is recorded to disk in 10-minute files, beginning at the hour, 10 past, 20 past, half past, 20 to and 10 to the hour. The first file is an odd length as it fills the time between the start and the next 10-minute mark. It must be at least 2 minutes long to ensure that the calibration
tones will be in the first file of a series. The file names have a base name with the date and time appended, e.g., Scribe Log 2016-12-27 15-16-57.wav. The files are stored in a directory for the year and another for the day.

The files themselves contain metadata including; the GPS-derived location and time at the start of the file recording, the file’s original filename, the software version, the sound device name, and the ‘track number,’ a number incremented for each file from the first in the recording. Also included is an encrypted version of the file’s MD5 digest. Since the files are uncompressed they can be reduced to about 50% by zipping without loss of information.

When post-processing these sound files, the first file in a sequence will contain the calibration tones from the calibrator. These can be found by software such as Matlab, which can be programmed to automatically detect them and use them to calibrate subsequent files in the recording. The CSI-ACHE protocol also requires a calibration at the end of a recording. Again this can be processed automatically.

RESULTS

The following graphs show the results of a number of tests carried out on the Mk 1 and/or Mk 2 versions of the SAM Scribe system.

Noise floor

Tests of the SAM Scribe FS units’ noise floor were conducted. These were produced by attaching 100 \( \Omega \) resistors to the unit, in place of the microphones, to produce a constant-current signal to the microphone inputs. Figures 3 & 4 show the results for both Mk 1 and Mk 2 versions.

Figure 3 shows the noise floor under standard settings, no gain boost and no high-pass filter. The noise floor of both units is fairly flat from 0.2 Hz up to about 100 Hz. From there it rises at a constant decibel increase per frequency decade. The Mk 1 version has a floor of between 10 and 20 dB rising to about 25 dB at 15 kHz. It shows two large peaks at about 9 Hz and 20 Hz. The Mk 2 version has a noise floor about between about 5 dB and 15 dB rising to about 22 dB at 15 kHz. Both versions show a 1 kHz peak with two harmonics at 2 kHz and 3 kHz.

![Figure 3: Noise floor of the Mk 1 and Mk 2 versions of the SAM Scribe FS unit. These were produced by attaching 100 \( \Omega \) resistors to the unit instead of microphones. This graph shows the results with the with no gain boost and without the 1 Hz high-pass filter. (These noise floors are unrelated to those of the microphones.)](image)

Figure 4 shows the noise floor with the 20 dB gain boost on. Both Scribe units have fairly constant noise floors in the infrasound region that drop to a minimum at about 100–200 Hz.
and then rise again in a constant deciBel increase per frequency decade. There is a peak at 60 Hz for both units and another peak at 1 kHz with harmonics at 2 kHz and 3 kHz for all but the red channel of the Mk 2 unit.

The Mk 2 unit has a minimum noise floor of about -10 dB at about 40–100 Hz, rising to about 5 dB at 15 kHz.

Figure 4: Noise floor of the Mk 1 and Mk 2 versions of the SAM Scribe FS unit. These were produced by attaching 100 Ω resistors to the unit instead of microphones. This graph shows the results with the 20 dB gain boost on and without the 1 Hz high-pass filter. (These noise floors are unrelated to those of the microphones.)

In-field comparison

A comparison is shown in Figure 5 of field recordings where both a Mk 1 SAM Scribe and a SVAN 979 were present.

Figure 5: Comparison of the Mk 1 SAM Scribe recording with a SVAN 979 recording. Included are the noise floor plots of a Mk 1 and a Mk 2 SAM Scribe FS unit.

It can be seen that the SVAN and Scribe recordings both agreed as to the sound level from 0.2 Hz through to 6 Hz. Above 6 Hz the SVAN shows significantly lower values of sound level than the Scribe, averaging about 10–12 dB less throughout. All but one of the peaks in the SVAN spectrum have matching peaks in the Scribe spectrum. The exception is the peak at about 120 Hz.
The noise floors of the Scribe Mk 1 and Mk 2 models are included in Figure 4. The peak in the Scribe spectrum at 1000 Hz matches that in the noise floor of the Mk 1. The curve of the Scribe spectrum matches that of the noise floor from about 6 Hz through to 4000 Hz.

**Linearity**

The manufacturer’s specification includes the following frequency-response curve (Figure 6) for the Primo microphone capsules used in the SAM Scribe microphones [15]. The microphone sensitivity is constant to within ±1 dB for the frequency range from about 0.5 dB to 5 kHz. Beyond this range the sensitivity increases and becomes variable between 0 and +5 dB.

![Figure 6: Frequency-response curve for the Primo microphone capsule used in the SAM Scribe microphones.](image)

**DISCUSSION**

**Noise floor**

Results of the noise-floor experiments (Figs 3 & 4) indicate that the Mk 2 version has a noise floor of less than 15 dB through most of its range and below 25 dB beyond 1 kHz. When in the more sensitive 20-dB-gain-boost mode, the infrasound region has a similar noise floor; the low-frequency range has a noise floor of about 0 dB; and about 5 dB for the high-frequency range.

The peak at 1 kHz in all the noise floors is from internal circuitry in or near the Cmedia chip. The location of this chip on the PCB of the Mk 2 version results in this peak appearing in the blue channel but not in the red channel. The peak at 60 Hz in the sensitive mode may be from a similar source or from the USB power supply. Another test (using an iMac) produced a number of large peaks in the infrasound region suggesting that these frequencies are not adequately suppressed in the USB power supply. A further test using only the notebook’s battery did not show any significant difference from those shown above where the notebook was connected to mains power.

While the noise floor of the complete system was not available at the time of writing this paper, these graphs indicate that the SAM Scribe FS unit should be capable of recording the acoustical environments of interest to the CSI-ACHE.
In-field comparison

To calibrate from 1 Hz to 20 kHz requires a \( \frac{1}{2} \)" capsule with a detachable screen to connect to a standard calibration system. This was not possible with the SAM Scribe microphones. Instead calibration should be done by simultaneous comparison with a calibrated microphone in a test chamber. The test chamber was under construction at the time of writing this paper.

Instead a comparison is presented from field recordings where both a Mk 1 SAM Scribe and a SVAN 979 were present. The microphones from the two instruments were placed on either side of a bed with the bed occupying the centre of the room. This comparison is shown in Figure 5.

It would appear that the major disagreement between the two recordings from about 6 Hz to 4500 Hz may be due to the difference in noise floors. The sound level up to 6 Hz is well above the noise floor of the Scribe and both instruments have matching sound levels in this region. Above this frequency the sound level drops about 20 dB and the Scribe curve follows the same curve as the Mk 1 noise floor. This suggests that the noise floor of the microphones is lifting the overall noise floor by about 10 dB. If this is the case, then the Mk 2 Scribe would provide a better match in this situation.

The drop-off of the SVAN noise level below 1 Hz may be a consequence of the SVAN’s design as it is not intended for such low frequencies. It will allow a small bleed of air between both sides of the microphone diaphragm to equalise the long-term pressure on both sides. This would show just such a drop-off as is seen here.

The sound level is very low at these frequencies. In the region of 1 kHz, a sound level of 0 dB is the lower limit of good human hearing and rises from there. Were the issue solely based on human perception it would be questionable whether having a lower noise floor would be of any advantage for the SAM Scribe system. This is not the case however and it is not yet known what the minimum sound level is at any frequency that can cause a response in humans.

The Scribe overestimates the size of the peak at about 75 Hz by about 10 dB. Huson [16] states that sound levels close to the noise floor of an instrument will be overestimated and can be corrected for to some extent. This may explain this discrepancy and suggests that it can be partially alleviated.

As it stands, it would appear that the Mk 2 Scribe can provide reasonable recordings in these quiet rural soundscapes but the Mk 1 Scribe cannot.

Linearity

For the purposes of the CSI-ACHE the linearity of the Primo microphone capsules appear to be satisfactory in the range 0.5 Hz to 5 kHz. A correction of the sensitivity could stretch this to 10 kHz. The variability beyond this would suggest that the measurement setup does not approximate a free field condition at these high frequencies. Nor is there an indication from the manufacturer of the manufacturing tolerances of their microphones. This must be answered by calibration of the SAM Scribe system in a suitable testing chamber. This has been designed was entering commercial production at the time of writing.

Once the calibration chamber has been set up, each microphone and Scribe unit can be calibrated in-house before shipping. The calibration curves can be used by the CSI-ACHE library software to correct individual recordings when they are processed.
CONCLUSIONS

A recording system has been created to support the Citizen Science Initiative on Acoustical Characterisation of Human Environments (CSI-ACHE) with the aim of collecting high-fidelity recordings from a diverse set of environments.

The system is capable of accurately recording from 0.5 Hz to 10 kHz. Accurate recording from 0.1 Hz to 20 kHz has not yet been demonstrated through calibration.

The noise floor of the Scribe Mk 1 version (excluding microphones) does not appear to cater well to quiet rural soundscapes although the Scribe Mk 2 version is likely to be able to.

A field comparison of a Scribe Mk 1 with a SVAN 979 sound level meter appeared to show this lack of depth in the noise floor. The Scribe Mk 2 noise floor should be able to capture reasonable recordings in quiet rural soundscapes.

The manufacturer’s specifications for the microphone capsule suggest that they are satisfactory for the range of about 0.5 Hz to 5 kHz without correction for the frequency response and 0.5 Hz to 10 kHz with correction. The response outside this region has yet to be verified by calibration tests.

Financial Disclosure

Due to their efforts in the creation of the SAM Scribe system, authors HHCB, BIR and SRS have a financial interest in the SAM Scribe system.

Acknowledgements

The authors wish to acknowledge the work of Rod Elliott of ESP, Sydney, Australia for his work on the design of the SAM Scribe, the Waubra Foundation, Australia for their support and assistance, Colin Wheeler of Two Wheels Engineering, Palmerston North, New Zealand for his valued assistance, and also the feedback and support of Steven Cooper of The Acoustic Group, Sydney.

REFERENCES


