Methods in Ecology and Evolution

Methods in Ecology and Evolution 2015 doi: 10.1111/2041-210X.12330

REVIEW

Measuring acoustic habitats

Nathan D. Merchant[1, 2, 3]*, Kurt M. Fristrup[4], Mark P. Johnson[5], Peter L. Tyack[5], Matthew J. Witt[6], Philippe Blondel[3] and Susan E. Parks[2]

[1]Centre for Environment, Fisheries & Aquaculture Science (Cefas), Lowestoft, Suffolk NR33 0HT, UK; [2]Department of Biology, Syracuse University, Syracuse, NY 13244, USA; [3]Department of Physics, University of Bath, Bath BA2 7AY, UK; [4]Natural Sounds and Night Skies Division, National Park Service, Fort Collins, CO 80525, USA; [5]Scottish Oceans Institute, University of St. Andrews, St. Andrews, Fife KY16 8LB, UK; and [6]Environment and Sustainability Institute, University of Exeter, Penryn TR10 9FE, UK

Summary

1. Many organisms depend on sound for communication, predator/prey detection and navigation. The acoustic environment can therefore play an important role in ecosystem dynamics and evolution. A growing number of studies are documenting acoustic habitats and their influences on animal development, behaviour, physiology and spatial ecology, which has led to increasing demand for passive acoustic monitoring (PAM) expertise in the life sciences. However, as yet, there has been no synthesis of data processing methods for acoustic habitat monitoring, which presents an unnecessary obstacle to would-be PAM analysts.

2. Here, we review the signal processing techniques needed to produce calibrated measurements of terrestrial and aquatic acoustic habitats. We include a supplemental tutorial and template computer codes in MATLAB and R, which give detailed guidance on how to produce calibrated spectrograms and statistical analyses of sound levels. Key metrics and terminology for the characterisation of biotic, abiotic and anthropogenic sound are covered, and their application to relevant monitoring scenarios is illustrated through example data sets. To inform study design and hardware selection, we also include an up-to-date overview of terrestrial and aquatic PAM instruments.

3. Monitoring of acoustic habitats at large spatiotemporal scales is becoming possible through recent advances in PAM technology. This will enhance our understanding of the role of sound in the spatial ecology of acoustically sensitive species and inform spatial planning to mitigate the rising influence of anthropogenic noise in these ecosystems. As we demonstrate in this work, progress in these areas will depend upon the application of consistent and appropriate PAM methodologies.

Key-words: acoustic ecology, ambient noise, anthropogenic noise, bioacoustics, ecoacoustics, habitat monitoring, passive acoustic monitoring, remote sensing, soundscape

Introduction

The increasing sophistication of passive acoustic monitoring (PAM) – the recording of sound in a habitat – has led to new insights in the study of acoustically sensitive organisms over a wide range of spatial and temporal scales (Van Parijs *et al.* 2009; Blumstein *et al.* 2011). Such studies point towards the fundamental role of sound for many species and ecosystems, mediating processes as diverse as predator–prey interactions (Remage-Healey, Nowacek & Bass 2006), larval settlement (Simpson *et al.* 2005) and coordinated behaviour (Boinski & Campbell 1995).

The acoustical backdrop to these phenomena is by no means a silent world: the evolution of acoustic signalling has taken place in the context of a varying natural background (Wiley & Richards 1978; Brumm & Slabbekoorn 2005) to which organisms adapt their acoustic behaviour (Morton 1975). This background sound is generated by weather processes (wind, rain, thunder), seismic events and competing biotic sound (Hildebrand 2009; Pijanowski *et al.* 2011). However, since the advent of large-scale industrialisation, acoustic habitats have become increasingly disrupted by anthropogenic noise. On land, these sources include road, rail and air transport, and industrial activity (Barber, Crooks & Fristrup 2010), while underwater, shipping, offshore construction, oil and gas exploration, and sonar operations contribute to the soundscape (Payne & Webb 1971; Hildebrand 2009). In both domains, this noise can mask acoustic cues (Brumm & Slabbekoorn 2005; Clark *et al.* 2009) and elicit behavioural responses (Sun & Narins 2005; Nowacek *et al.* 2007), with the potential to cause chronic physiological stress (Rolland *et al.* 2012; Francis & Barber 2013) and wider effects on populations (National Research Council 2005) and communities (Francis, Ortega & Cruz 2009). Awareness of

*Correspondence author. E-mail: nathan.merchant@cefas.co.uk

these human impacts has brought renewed urgency to the study of acoustic habitats and their influences on ecosystem processes (Francis & Barber 2013).

The rapid expansion of this field has been driven by major advances in PAM technology and has led to growing demand for acoustics expertise in the life sciences. Over the past two decades, the development of cost-effective autonomous recording units has revolutionised bioacoustics in both air (Mennill *et al.* 2012; Digby *et al.* 2013) and underwater (Sousa-Lima *et al.* 2013). Marine bioacoustics has been particularly driven by technological innovation, with new perspectives offered by non-invasive recording tags (Burgess *et al.* 1998; Johnson & Tyack 2003), autonomous gliders (Rudnick, Davis & Eriksen 2004; Baumgartner & Fratantoni 2008), drifting platforms (Wilson, Benjamins & Elliott 2013) and region-scale cabled ocean observatories (Favali, Beranzoli & de Santis 2015).

However, there is currently a lack of clear guidance on how to analyse PAM data to produce calibrated measurements. Calibrated data provide absolute measures of biotic, abiotic and anthropogenic sound levels, which are necessary to draw meaningful comparisons of habitats through time and at different locations. Here, we seek to address this deficit through a user-friendly guide to the methods that underpin the study of acoustic habitats. We detail the signal processing steps required to produce absolute measurements of sound pressure and demonstrate the use of analytical techniques to describe variability and trends in sound levels and to characterise discrete acoustic events. In doing so, we emphasise the considerable overlap in acoustic analysis methods in air and underwater, and the potential for both terrestrial and aquatic bioacoustics to benefit from greater integration and knowledge exchange.

Monitoring platforms

Technology has shaped the development of bioacoustics. Advances in instrumentation, data storage capacity and data analysis capabilities have opened up new avenues of research while enriching established areas. To help contextualise historical constraints on data collection and inform the design of future habitat monitoring programs, this section briefly summarises the capabilities and limitations of the main types of PAM platform. We first consider *fixed platforms* – those designed to be deployed at one location for days or longer – and then *mobile platforms* – those that record while in motion or are portable and deployed for short periods.

FIXED PLATFORMS

The recent expansion in the study of acoustic habitats has been facilitated by the development of autonomous acoustic recorders: self-contained digital instruments that can be fixed to terrestrial structures or moored to the seafloor to record the soundscape continuously on the scale of months (Mennill *et al.* 2012; Sousa-Lima *et al.* 2013; see Table 1 for a selection of commercially available devices). These are generally more cost-effective and easier to deploy than cabled systems and can be positioned in arrays to investigate spatial characteristics of acoustic habitats and to localise and track sound sources (Van Parijs *et al.* 2009; Blumstein *et al.* 2011). Each unit consists of battery-powered electronics for digital data acquisition and storage within a weather- or waterproof housing. The acoustic transducer (microphone or hydrophone) may be mounted on the device or attached via a cable. Deployment longevity is limited by power consumption and data storage capacity, both of which continue to improve as the technology evolves (Sousa-Lima *et al.* 2013). Supplementing the inbuilt power supply with an external source (solar panel or additional battery) or duty cycling the ‘on time’ can increase longevity. Habitat monitoring can continue indefinitely if recorders are regularly serviced to replenish batteries and data storage, and systems are also being developed for remote data retrieval over wireless communication networks (e.g. Wildlife Acoustics Song Stream; SMRU Marine PAMBuoy). Seafloor-mounted systems can be recovered using acoustic release devices, although acoustic release malfunction and trawling by fishing vessels are common field hazards (Dudzinski *et al.* 2011). Theft and vandalism are also potential concerns (Clarin *et al.* 2013), as well as damage from wildlife and biofouling.

Cabled systems have long been used to monitor underwater sound (Wenz 1961) and consist of seafloor-mounted hydrophones connected to shore stations providing power and data acquisition. In recent years, cable-mounted systems have seen a resurgence with the expansion of cabled ocean observatories, for example the NEPTUNE, VENUS and RSN networks in the North-east Pacific (Favali, Beranzoli & de Santis 2015). The capacity for real-time acoustic monitoring and the less frequent servicing required (compared to autonomous units) are the principal advantages of cabled systems, although the associated costs of deployment, maintenance and management of these long-term devices and the large volumes of data they generate are correspondingly high.

MOBILE PLATFORMS

In air, a convenient and portable tool for acoustic habitat monitoring is the commercial sound level meter, which makes calibrated measurements of sound pressure level (SPL; Table 2; Appendix S1, Eqn 17). However, many sound level meters apply standardised filters known as A- and C-weightings, which modify the signal to approximate the frequency response of human hearing (Kinsler *et al.* 1999). It is therefore important to consider whether the frequency range of interest coincides with human audibility (which it may, e.g. Brumm 2004) and whether a human frequency-weighted metric is appropriate. Some sound level meters have an unweighted setting (known as Z-weighting, flat or linear), but are still limited to the nominal frequency range of human hearing.

These limitations of the sound level meter can be avoided by instead making calibrated sound level measurements using portable or autonomous field recorders. This has long been standard practice in aquatic bioacoustics (Au & Hastings 2008), where portable field recorders are used to record data from a hydrophone lowered over the side of a drifting vessel.

Table 1. Selection of commercially available integrated acoustic habitat recorders

Manufacturer	Model	Channels	Maximum sampling rate (kHz)	Domain
Wildlife Acoustics	SM2+	2	96	Air
Wildlife Acoustics	SM3	2	96	Air
Jasco Applied Sciences	AMAR G3	9	687·5	Underwater
Loggerhead Instruments	DSG-ST	1	288	Underwater
Ocean Instruments	SoundTrap 202HF	1	576	Underwater
Wildlife Acoustics	SM3M	2	192	Underwater

Table 2. Units and abbreviations for quantities described in the text. Note that some authors have used units of dB re p_{ref} μPa^2 for SPL and TOLs: this notational difference does not affect the numerical levels reported

Term	Abbreviation	Units in air	Units underwater	Short definition
Sound pressure level	SPL	dB re 20 μPa	dB re 1 μPa	Sound level over specified frequency range as a single number
Power spectral density	PSD	dB re (20 μPa)^{2}Hz^{-1}	dB re 1 μPa^{2}Hz^{-1}	Standardised spectrum of sound levels across frequency
1/3-octave band level	TOL	dB re 20 μPa	dB re 1 μPa	Coarse sound level spectrum with logarithmic frequency scaling
Sound exposure level	SEL	dB re (20 μPa)^{2}s	dB re 1 μPa^{2}s	Cumulative measure of sound energy

Unlike data from sound level meters, field recordings also enable post hoc analyses to identify components of the acoustic environment or to produce other acoustical metrics. Calibrating the recordings requires knowledge of the relevant hardware specifications and an understanding of the signal processing steps required for the calibration procedure (Appendix S1, Section 3). Although such calibrated measurements of terrestrial acoustic habitats have been made (Waser & Brown 1986), these are the exception: habitat monitoring studies more commonly report relative (i.e. uncalibrated) sound spectra (Boinski & Campbell 1995), sometimes supplemented by measurements made with sound level meters (Lengagne & Slater 2002).

Portable field recorders are well suited to short-term surveys in favourable weather conditions. A battery-powered digital audio recorder can be rapidly deployed with a microphone mounted on a hand-held boom or tripod, lashed to vegetation, or with a hydrophone lowered to a desired depth in the water. Care should be taken to minimise incidental noise from the presence of the monitoring platform and noise generated by flow of air or water past the acoustic sensor (known as flow noise or 'pseudonoise').

Other mobile platforms have largely been developed for marine bioacoustics applications. For example, to study marine mammal behaviour, acoustic tags have been developed which are temporarily attached to animals via suction cups, recording sound and tracking movement for up to several days (Burgess *et al.* 1998; Johnson & Tyack 2003). While these devices have been effective in studying vocalisations in cetaceans (Johnson *et al.* 2004) and exposure to high-amplitude anthropogenic noise (DeRuiter *et al.* 2013), the ability to record lower-amplitude low-frequency sound is limited: the movement of the animal through the water causes turbulence around the hydrophone, producing flow noise which contaminates low frequencies (Johnson & Tyack 2003), and acoustic time series are periodically disrupted by surfacing events (Fig. 1) and vocalisations of the host animal. Application of similar recording tags to terrestrial wildlife has thus far been limited to one study of wild mule deer (*Odocoileus hemionus*; Lynch *et al.* 2013). Emerging mobile platforms also include autonomous underwater gliders (Rudnick, Davis & Eriksen 2004; Baumgartner & Fratantoni 2008), which can be deployed for up to several hundred days, and freely drifting platforms (Wilson, Benjamins & Elliott 2013), which minimise flow noise in high tidal flow environments by moving with the current.

GENERAL CONSIDERATIONS

The ability of PAM systems to accurately record sound is limited by several factors: dynamic range, frequency range and system self-noise. The dynamic range is the ratio of the highest to the lowest amplitude that can be measured by a given system and can be scaled to higher or lower amplitudes by adding gain to the signal (within the limitations of the device's own self-noise; Johnson, Partan & Hurst 2013; Merchant *et al.* 2013). If the gain is too low, quieter sounds may not be recorded (Rempel *et al.* 2013), and if it is too high, loud sounds can saturate the system, leading to distortion of the signal through clipping (Madsen & Wahlberg 2007; Fristrup & Mennitt 2012). A PAM system should be chosen whose frequency range encompasses the spectral content of all sounds of interest to the study. This range is limited to half the sampling rate of the recordings (the Nyquist frequency), and by the sensitivity of the transducer and any filtering in the recording system (Appendix S1; Madsen & Wahlberg 2007). System self-noise is noise generated by the recording system and acoustic transducer and can limit the ability of a system to record low-amplitude sound (Fristrup & Mennitt 2012; Johnson, Partan & Hurst 2013). It is therefore important to consider self-noise specifications when selecting a PAM system, and where possible to measure self-noise levels by making recordings in a quiet location isolated from sources of vibration.

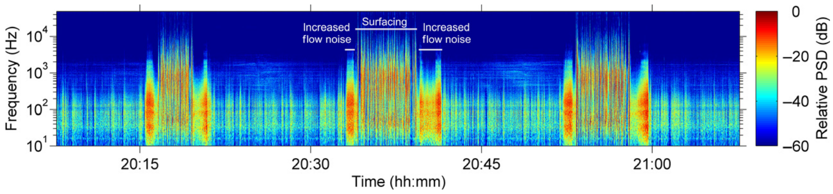


Fig. 1. Recording from an acoustic recording tag (DTAG) attached to a North Atlantic right whale in the Bay of Fundy, Maine, USA, on 3 August 2005. Sampling rate: 96 kHz. Note the periodic surfacing events (three in total) and the increase in flow noise before and after these surfacing events caused by increased travel speed.

A wider consideration is that many species (e.g. many fish, insect and arthropod species) primarily sense sound not through sound pressure, but through particle motion, a directional property of the sound field (Hawkins 1981; Stumpner & von Helversen 2001). The devices described above do not directly measure particle motion, which may be significant in the region near to sound sources (the near field) or close to reflecting surfaces, where variations in sound pressure and particle motion can differ considerably (far from the source and surfaces – in the far field – sound pressure and particle motion are directly related). Although many studies have measured auditory sensitivity to particle motion in controlled experiments, the use of particle motion sensors in habitat monitoring is still in its infancy.

Data analysis

There is no standard or widely available software for the production of calibrated PAM measurements. While there are several options for annotation, detection and classification of bioacoustic signals [e.g. Raven (Charif, Clark & Fristrup 2004), PAMGuard (Gillespie *et al.* 2009)], these operations do not demand absolute measures of sound level and so typically use relative levels of signal amplitude. Calibrated measurements require hardware-specific data (the microphone or hydrophone sensitivity and properties of the data acquisition system), and many research groups use custom-developed programs for PAM analysis tailored to their equipment. In the accompanying tutorial (Appendix S1), we present a detailed guide to the signal processing steps required to produce calibrated PAM data in terrestrial and aquatic environments, including *PAMGuide* (Data S1), a template code provided in MATLAB and R versions which implements the equations presented in the tutorial. Here, we present a non-technical outline of the calibration procedure and a description of key acoustical metrics. Application of these metrics to the characterisation of acoustic habitats is illustrated in the following section.

The basic principle of making absolute sound pressure measurements is to reverse the transformations made to the signal along its path into the data acquisition system (Fig. 2). The same steps apply to in-air and underwater PAM devices, which employ the same fundamental components: (i) an acoustic transducer (microphone or hydrophone) to convert the sound pressure into a voltage signal; (ii) a pre-amplifier, which is used to increase the amplitude of the voltage signal before it is recorded and may apply frequency equalisation and anti-aliasing filters; and (iii) an analogue-to-digital converter (ADC), which digitises the analogue voltage signal and then packages the data as audio files with a standardised amplitude range (Au & Hastings 2008; e.g. a 16-bit recording has an amplitude range of $\pm 2^{16-1}$ bits, i.e. integers between $-32\ 768$ and $32\ 767$). To retrieve the original measurement of sound pressure from the audio files, it is therefore necessary to know: (i) the sensitivity of the acoustic transducer, which defines how much voltage is generated per unit of sound pressure; (ii) the amount of voltage gain (if any) applied by the pre-amplifier; and (iii) the ADC input voltage which corresponds to the maximum amplitude that can be represented in the audio files (Robinson, Lepper & Hazelwood 2014). This process is summarised in Fig. 2. Alternatively, an 'end-to-end' calibration can be carried out by inputting a known signal into the transducer, for example with a pistonphone (Appendix S1, Section 3). It is important to note that the transducer sensitivity and the gain of the pre-amplifier may vary significantly with frequency, in which case frequency-dependent corrections should be applied to the signal.

This procedure yields the relationship between the recorded signal and the signal at the transducer, allowing the sound pressure signal (known as the pressure waveform) to be computed from the recordings. The pressure waveform is used to measure the amplitudes of impulsive sounds, such as echolocation calls or gunshots (see next section). Acoustic habitat characterisation generally requires further analysis of the pressure waveform, such as transformation into the frequency domain to analyse frequency characteristics and averaging to represent longer periods (Appendix S1). The output of these analyses is typically one of a number of common acoustical metrics shown in Table 2.

Broadband sound pressure level (SPL) is the most ubiquitous acoustic metric and expresses the root-mean-square (RMS) sound amplitude within a given time window and frequency range as a single decibel (dB) level, for example 64 dB re 20 μPa (Kinsler *et al.* 1999). Most acoustical metrics are expressed as a decibel level relative to a reference pressure, p_{ref}, in the form X dB re p_{ref}. In air, p_{ref} is 20 μPa, which corresponds to the nominal threshold of human hearing at 1 kHz, while in underwater, p_{ref} is 1 μPa. Spectra showing how sound level varies with frequency are given by the power spectral density (PSD), which describes the power in the equivalent of 1-Hz bands in the frequency domain (although the actual frequency resolution may be much coarser than 1 Hz), or by fractional octave analysis

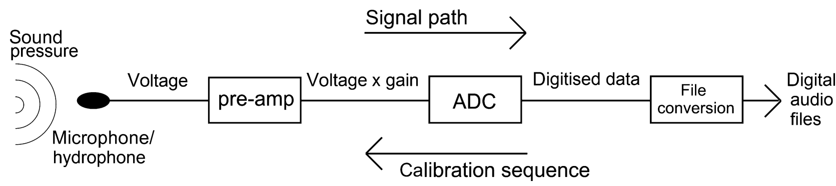


Fig. 2. Signal path and calibration sequence for a typical passive acoustic monitoring system. ADC, analogue-to-digital converter.

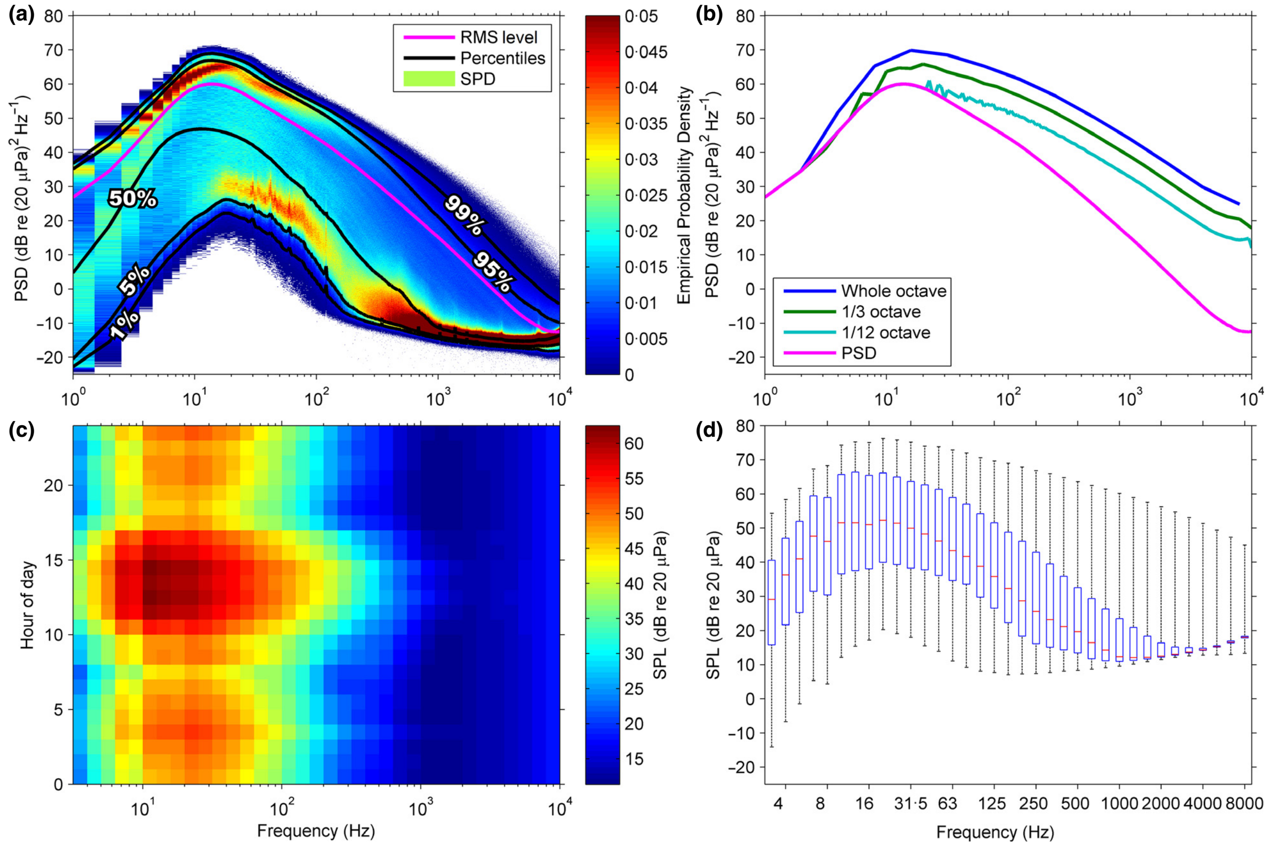


Fig. 3. Statistical analyses of long-term acoustic data recorded using a Wildlife Acoustics SM2+ SongMeter at the Central Plains Experimental Research Station, Colorado, from 19 October to 13 December 2013. Sampling rate: 44.1 kHz. (a) RMS level of the PSD, percentiles and SPD, showing bimodality of sound level distribution and presence of noise floor above ~100 Hz (b) RMS level of the PSD and fractional octave bands (c) Median 1/3-octave level for each hour of the day. (d) Boxplot of 1/3-octave bands – mid-line is median, edges of boxes are first and third quartiles, whiskers are minima and maxima. Note that the noise floor of the instrument limits the range of the higher frequency bands.

(typically 1/3-octave band levels, TOLs), which measures the power in frequency bands that widen exponentially with increasing frequency and are evenly spaced on a logarithmic frequency axis (Fig. 3c). Finally, the sound exposure level (SEL) is a summation of sound energy through time over a specified duration and is often used to assess cumulative exposure to noise.

An emerging area of study is the use of acoustic indices as proxy indicators of biodiversity and species composition (Towsey, Parsons & Sueur 2014). Various new statistical indices (e.g. acoustic diversity index, acoustic complexity index) have been applied in comparative studies of distinct habitats for this purpose, so far with mixed results (Lellouch *et al.* 2014; Towsey, Parsons & Sueur 2014). It remains to be seen to what extent such techniques will complement field observations of biodiversity and automatic detection and classification of species via PAM.

Habitat characterisation

Acoustic habitats can be described by statistical analysis of sound levels and by time-series representations. Based on the metrics outlined above and defined in Appendix S1, this section illustrates how these techniques can be used effectively for various habitat monitoring applications. In general, statistical analyses are more suited to characterising variability and comparing acoustic habitats at differing times or locations, while discrete events and trends in sound levels are better described by time series.

STATISTICAL ANALYSIS

There are several ways to calculate the average sound level from a series of shorter measurements, each of which has particular advantages. The most common metric is the

RMS level (see Fig. 3), which is the mean of the squared sound pressure computed before it is converted to dB (Appendix S1, Eqn 18). The RMS level is the most prevalent averaging metric, perhaps due to the historic centrality of L_{eq} in terrestrial noise characterisation (L_{eq} is the equivalent RMS level during a specified period, which correlates with human noise disturbance), and its direct relation to SPL, which makes it independent of the length of the time segments used in the original analysis (unlike other averaging metrics; Merchant *et al.* 2012). However, the RMS level is also strongly influenced by the highest sound levels (it can be >95th percentile in some cases; Merchant *et al.* 2013) and so should be used with caution if applied to recordings with intermittent high-amplitude events (e.g. pile driving; Madsen 2005).

Alternative averages can give more statistically representative measures of sound levels than the RMS level. The mode, defined as the sound level corresponding to the maximum probability density at each frequency, is (by definition) the most representative metric, although few studies have made use of it (e.g. Parks, Urazghildiiev & Clark 2009; Merchant *et al.* 2012) and there is the potential for multimodality to produce misleading results (Fig. 3a; Merchant, *et al.* 2012, 2013). Median sound levels (Fig. 3d; known as L_{50} in terrestrial noise assessment) can also be used as an indicator of typical sound levels in a habitat (e.g. Klinck *et al.* 2012). These are more robust than the mode and are generally insensitive to limitations in the dynamic range of the recording instrument.

The range of sound levels in a habitat can be assessed by plotting the percentile levels across the frequency spectrum (Fig. 3a; Richardson *et al.* 1995; Castellote, Clark & Lammers 2012). These are alternatively known as exceedance levels, although the percentiles are reversed, for example the 95% exceedance level, L_{95}, is equivalent to the 5th percentile. Percentiles provide an approximate indication of the distribution of sound levels and may be useful in characterising the potential extent of acoustic masking for a particular species (Clark *et al.* 2009). A more comprehensive analysis of the sound level distribution is given by the spectral probability density (SPD; Merchant *et al.* 2013), whereby the empirical probability density of sound levels in each frequency band is presented (Fig. 3a). This shows the modal structure and outlying data in the underlying distribution, which helps to interpret averages and percentiles. It can also reveal limitations in the recording system: for example, in the SPD shown in Fig. 3a, the self-noise of the instrument appears to limit recording of the lowest sound levels in the habitat above *c.* 100 Hz, evidenced by the flattening gradient of the lowest data points and their convergence with the mode above this frequency.

Other statistical analyses include average levels for particular temporal periods to examine cyclical trends, such as diel, seasonal or annual variability (see Fig. 3c, although other configurations are also used, e.g. Radford *et al.* 2008), and box-and-whisker plots (Fig. 3d; Bassett *et al.* 2012) showing the spread of quartiles in each band.

In the frequency domain, fine-scale variations in the sound spectrum can be assessed using the PSD, which is often computed at 1-Hz resolution (Fig. 3b), although this can be computationally demanding when processing and storing large data sets. Frequency resolution can be reduced using either coarser PSDs or fractional octave analysis, most commonly in 1/3-octave bands (Appendix S1, Eqns 13–16), wherein the frequency range of the band is directly proportional to the centre frequency (constant Q). For some taxa (e.g. mammals), constant Q frequency bands are a particularly useful tool for acoustic habitat analysis, as they can approximate the response of the auditory system. It is important to note that the spectral slope of the fractional octave band levels differs from the PSD (Fig. 3b), as the bands are scaled logarithmically with frequency (i.e. they widen with increasing frequency), meaning that higher frequency bands integrate energy over larger frequency ranges. For this reason, spectra with differing frequency bandwidths (e.g. PSD and 1/3-octave) cannot be directly compared.

TIME SERIES

Time series are used to characterise discrete events, such as vocal behaviour or anthropogenic noise events (typically in the form of spectrograms), and to track temporal trends in sound levels, usually in particular frequency bands (e.g. TOLs or broadband level).

A widely used time-series plot is the spectrogram, which comprises a series of PSD measurements showing how the sound level varies with time at each frequency (Fig. 4a). Sound sources can then be identified by visual inspection and by listening to the sound files (depending on the frequency range), or by automatic detection and classification. The time and

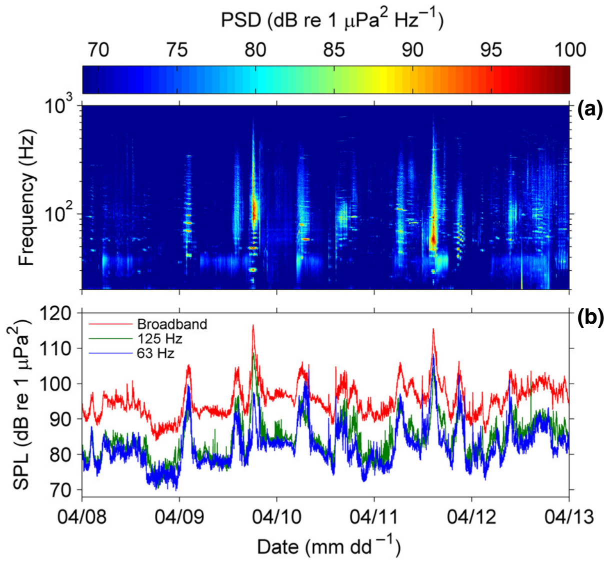


Fig. 4. Time-series analysis of ship passages recorded using an autonomous marine acoustic recording unit (MARU; developed by the Bioacoustics Research Program at Cornell University) in Stellwagen Bank National Marine Sanctuary, Massachusetts Bay, USA, in April 2010. Sampling rate: 2 kHz. (a) Spectrogram composed of PSDs with 1-s time segments; (b) 63- and 125-Hz 1/3-octave bands and broadband (0.2–1 kHz) SPL. Analysis parameters: 1-s time segments averaged to 60-s resolution via the Welch method.

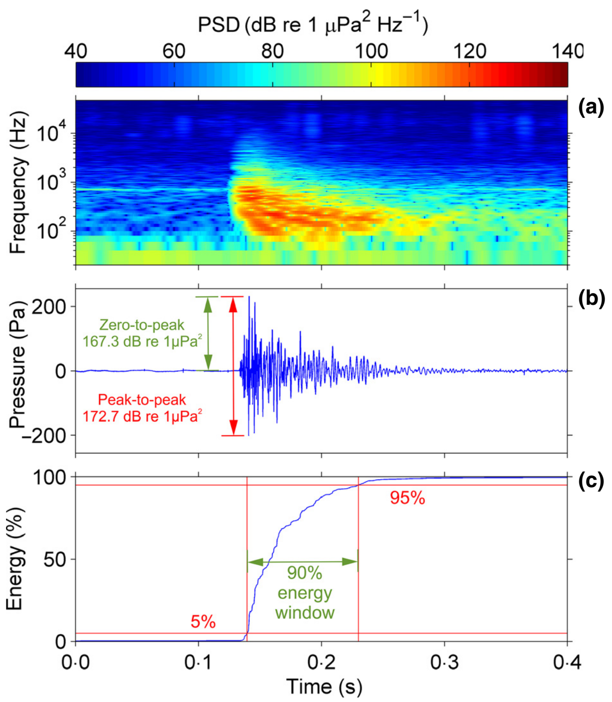


Fig. 5. Seismic airgun array recorded at 6 km distance using a Wildlife Acoustics SM2M deployed off southern Gabon, West Africa, 30 October 2012. Sampling rate: 96 kHz. (a) PSD: 0.05-s Hann window, 99% overlap. (b) Pressure waveform illustrating peak-to-peak and zero-to-peak metrics. (c) Cumulative energy of pulse showing 90% energy envelope.

frequency resolution of the spectrograms must be sufficiently high to resolve the sound events under consideration, and overlapping time windows may be used to smooth data in the time domain and to ensure sounds at the boundary between time windows are represented (Appendix S1, Section 4).

As well as describing discrete events, time series are used to track trends in sound levels in particular frequency bands (Fig. 4b), often in the context of long-term anthropogenic noise studies (e.g. Miksis-Olds, Bradley & Niu 2013).

While the amplitudes of nominally continuous sounds (e.g. tonal vocalisations, ships, wind-generated sound) can be accurately represented in frequency spectra, this is not the case for impulsive sounds (e.g. echolocation calls, snapping shrimp, gunshots, seismic airguns, pile driving; Fig. 5a). If impulsive sounds are identified, their amplitudes should instead be measured using the pressure waveform (Fig. 5b), which gives a much finer time resolution with which to measure the impulse amplitude (Madsen 2005). The peak amplitude of the pulse is described by either the peak-to-peak (SPL_{p-p}) or zero-to-peak pressure (SPL_{0-p}), as illustrated in Fig. 5b (Appendix S1, Eqns 22–23). A further metric, sound exposure level (SEL) describes the acoustic energy contained in the pulse (Appendix S1, Eqns 24–25). The duration of the pulse is commonly defined by the 90% energy envelope: the time window between 5% and 95% of the cumulative acoustic energy (Blackwell, Lawson & Williams 2004; Fig. 5c). To accurately measure acoustic pulses, a recording system should be chosen whose sensitivity does not vary significantly within the peak frequency range of the pulse, and the signal-to-noise ratio (SNR) of the pulse should be ≥ 10 dB (Appendix S1; Madsen 2005; Madsen & Wahlberg 2007).

Future directions

The temporal range of acoustic monitoring has been greatly increased by advances in instrumentation and data analysis. Looking forward, the emerging area of acoustic habitat mapping promises in turn to expand the spatial scope of monitoring studies. Integrating PAM data with spatial models has the potential to produce ground-truthed sound maps of acoustic habitats (Barber *et al.* 2011; Erbe, MacGillivray & Williams 2012; NOAA 2012; Mennitt, Sherrill & Fristrup 2014), which with large-scale monitoring could extend to the regional and national scales relevant to ecosystem-scale assessment (Barber *et al.* 2011; Mennitt, Sherrill & Fristrup 2014) and the migratory ranges of wide-ranging species (NOAA 2012). If such maps are sufficiently predictive of sound levels, they could help to highlight areas of concern for anthropogenic noise impact when overlain with animal distributions, and offer new perspectives on the spatial ecology of acoustically sensitive species. Large-scale mapping inevitably involves pooling measurements made by many groups, potentially using different devices; thus, the success of these modelling efforts will depend in part on the quality and standardisation of passive acoustic measurements they are based on. We hope that the overview of methods presented here contributes to a robust methodological foundation for this progress and helps to make passive acoustic techniques more accessible to newcomers to the field.

Acknowledgements

We thank Sofie Van Parijs and Genevieve Davis (NOAA) for providing passive acoustic recordings made during an NOPP-funded project (Award No. N0001410IP20027), Rodney J. MacAllister for permission to include the seismic airgun recordings from Gabon and Doug Quin for valuable discussions on recording techniques. We also thank the Associate Editor and three anonymous reviewers for helpful comments which improved the manuscript, and Jessica McCordic, Hannah Blair, Leanna Matthews and Sam Denes for useful feedback on the supplementary material. NDM was part-funded by an EPSRC Doctoral Training Award (No. EP/P505399/1). MPJ and PLT were supported by Marine Alliance for Science and Technology Scotland (MASTS) and MPJ by a Marie Curie CIG Award. Recordings from Colorado were made with support from the National Science Foundation under Grant No. 1340669.

Data accessibility

MATLAB and R codes used in the example analyses are available in the Supporting Information.

References

Au, W.W.L. & Hastings, M.C. (2008) *Principles of Marine Bioacoustics*. Springer, New York.

Barber, J.R., Crooks, K.R. & Fristrup, K.M. (2010) The costs of chronic noise exposure for terrestrial organisms. *Trends in Ecology & Evolution* **25**, 180–189.

Barber, J.R., Burdett, C.L., Reed, S.E., Warner, K.A., Formichella, C., Crooks, K.R., Theobald, D.M. & Fristrup, K.M. (2011) Anthropogenic noise exposure in protected natural areas: estimating the scale of ecological consequences. *Journal of Landscape Ecology* **26**, 1281–1295.

Bassett, C., Polagye, B., Holt, M. & Thomson, J. (2012) A vessel noise budget for Admiralty Inlet, Puget Sound, Washington (USA). *The Journal of the Acoustical Society of America* **132**, 3706–3719.

Baumgartner, M.F. & Fratantoni, D.M. (2008) Diel periodicity in both sei whale vocalization rates and the vertical migration of their copepod prey observed from ocean gliders. *Limnology and Oceanography* **53**, 2197–2209.

Blackwell, S.B., Lawson, J.W. & Williams, M.T. (2004) Tolerance by ringed seals (*Phoca hispida*) to impact pipe-driving and construction sounds at an oil production island. *The Journal of the Acoustical Society of America* **115**, 2346–2357.

Blumstein, D.T., Mennill, D.J., Clemins, P., Girod, L., Yao, K., Patricelli, G. *et al.* (2011) Acoustic monitoring in terrestrial environments using microphone arrays: applications, technological considerations and prospectus. *Journal of Applied Ecology* **48**, 758–767.

Boinski, S. & Campbell, A.F. (1995) Use of trill vocalizations to coordinate troop movement among white-faced capuchins: a second field test. *Behaviour* **132**, 875–901.

Brumm, H. (2004) The impact of environmental noise on song amplitude in a territorial bird. *Journal of Animal Ecology* **73**, 434–440.

Brumm, H. & Slabbekoorn, H. (2005) Acoustic communication in noise. *Advances in the Study of Behavior* **35**, 151–209.

Burgess, W.C., Tyack, P.L., Le Boeuf, B.J. & Costa, D.P. (1998) A programmable acoustic recording tag and first results from free-ranging northern elephant seals. *Deep Sea Research Part II: Topical Studies in Oceanography* **45**, 1327–1351.

Castellote, M., Clark, C.W. & Lammers, M.O. (2012) Acoustic and behavioural changes by fin whales (*Balaenoptera physalus*) in response to shipping and airgun noise. *Biological Conservation* **147**, 115–122.

Charif, R.A., Clark, C.W. & Fristrup, K.M. (2004) *Raven 1.2 User's Manual*. Cornell Laboratory of Ornithology, Ithaca, New York.

Clarin, B.M., Bitzilekis, E., Siemers, B.M. & Goerlitz, H.R. (2013) Personal messages reduce vandalism and theft of unattended scientific equipment. *Methods in Ecology and Evolution* **5**, 125–131.

Clark, C.W., Ellison, W.T., Southall, B.L., Hatch, L., Van Parijs, S.M., Frankel, A. & Ponirakis, D. (2009) Acoustic masking in marine ecosystems: intuitions, analysis, and implication. *Marine Ecology Progress Series* **395**, 201–222.

DeRuiter, S.L., Southall, B.L., Calambokidis, J., Zimmer, W.M.X., Sadykova, D., Falcone, E.A. *et al.* (2013) First direct measurements of behavioural responses by Cuvier's beaked whales to mid-frequency active sonar. *Biology Letters* **9**, 20130223.

Digby, A., Towsey, M., Bell, B.D. & Teal. P.D. (2013) A practical comparison of manual and autonomous methods for acoustic monitoring. *Methods in Ecology and Evolution* **4**, 675–683.

Dudzinski, K.M., Brown, S.J., Lammers, M., Lucke, K., Mann, D.A., Simard, P. *et al.* (2011) Trouble-shooting deployment and recovery options for various stationary passive acoustic monitoring devices in both shallow- and deep-water applications. *The Journal of the Acoustical Society of America* **129**, 436–448.

Erbe, C., MacGillivray, A. & Williams, R. (2012) Mapping cumulative noise from shipping to inform marine spatial planning. *The Journal of the Acoustical Society of America* **132**, EL423–EL428.

Favali, P., Beranzoli, L. & de Santis, A. (2015) *Seafloor Observatories: A New Vision of the Earth from the Abyss*. Springer-Praxis, UK.

Francis, C.D. & Barber, J.R. (2013) A framework for understanding noise impacts on wildlife: an urgent conservation priority. *Frontiers in Ecology and the Environment* **11**, 305–313.

Francis, C.D., Ortega, C.P. & Cruz, A. (2009) Noise pollution changes avian communities and species interactions. *Current Biology* **19**, 1415–1419.

Fristrup, K.M. & Mennitt, D. (2012) Bioacoustical monitoring in terrestrial environments. *Acoustics Today* **8**, 16–24.

Gillespie, D., Mellinger, D.K., Gordon, J., McLaren, D., Redmond, P., McHugh, R., Trinder, P., Deng, X.Y. & Thode, A. (2009) PAMGUARD: Semiautomated open source software for real-time acoustic detection and localization of cetaceans. *The Journal of the Acoustical Society of America* **125**, 2547–2547.

Hawkins, A.D. (1981) The hearing abilities of fish. *Hearing and Sound Communication in Fishes* (eds W.N. Tavolga, A.N. Popper & R.R. Fay), pp. 109–137. Springer, New York.

Hildebrand, J.A. (2009) Anthropogenic and natural sources of ambient noise in the ocean. *Marine Ecology Progress Series* **395**, 5–20.

Johnson, M.P. & Tyack, P.L. (2003) A digital acoustic recording tag for measuring the response of wild marine mammals to sound. *IEEE Journal of Oceanic Engineering* **28**, 3–12.

Johnson, M., Madsen, P.T., Zimmer, W.M.X., Aguilar de Soto, N. & Tyack, P.L. (2004) Beaked whales echolocate on prey. *Proceedings of the Royal Society of London. Series B* **271**, S383–S386.

Johnson, M., Partan, J. & Hurst, T. (2013) Low complexity lossless compression of underwater sound recordings. *The Journal of the Acoustical Society of America* **133**, 1387–1398.

Kinsler, L.E., Frey, A.R., Coppens, A.B. & Sanders, J.V. (1999) *Fundamentals of Acoustics*, 4th edn. Wiley, NJ.

Klinck, H., Nieukirk, S.L., Mellinger, D.K., Klinck, K., Matsumoto, H., & Dziak, R.P. (2012) Seasonal presence of cetaceans and ambient noise levels in polar waters of the North Atlantic. *The Journal of the Acoustical Society of America* **132**, EL176–EL181.

Lellouch, L., Pavoine, S., Jiguet, F., Glotin, H. & Sueur, J. (2014) Monitoring temporal change of bird communities with dissimilarity acoustic indices. *Methods in Ecology and Evolution* **5**, 495–505.

Lengagne, T. & Slater, P.J.B. (2002) The effects of rain on acoustic communication: tawny owls have good reason for calling less in wet weather. *Proceedings of the Royal Society of London. Series B* **269**, 2121–2125.

Lynch, E., Angeloni, L., Fristrup, K., Joyce, D. & Wittemyer, G. (2013) The use of on-animal acoustical recording devices for studying animal behavior. *Ecology and Evolution* **3**, 2030–2037.

Madsen, P.T. (2005) Marine mammals and noise: Problems with root mean square sound pressure levels for transients. *The Journal of the Acoustical Society of America* **117**, 3952–3957.

Madsen, P.T. & Wahlberg, M. (2007) Recording and quantification of ultrasonic echolocation clicks from free-ranging toothed whales. *Deep-Sea Research Part I: Oceanographic Research Papers* **54**, 1421–1444.

Mennill, D.J., Battiston, M., Wilson, D.R., Foote, J.R. & Doucet, S.M. (2012) Field test of an affordable portable wireless microphone array for spatial monitoring of animal ecology and behaviour. *Methods in Ecology and Evolution* **3**, 704–712.

Mennitt, D., Sherrill, K. & Fristrup, K. (2014) A geospatial model of ambient sound pressure levels in the contiguous United States. *The Journal of the Acoustical Society of America* **135**, 2746–2764.

Merchant, N.D., Blondel, P., Dakin, D.T. & Dorocicz, J. (2012) Averaging underwater noise levels for environmental assessment of shipping. *The Journal of the Acoustical Society of America* **132**, EL343–EL349.

Merchant, N.D., Barton, T.R., Thompson, P.M., Pirotta, E., Dakin, D.T. & Dorocicz, J. (2013) Spectral probability density as a tool for ambient noise analysis. *The Journal of the Acoustical Society of America* **133**, EL262–EL267.

Miksis-Olds, J.L., Bradley, D.L. & Niu, X.M. (2013) Decadal trends in Indian Ocean ambient sound. *The Journal of the Acoustical Society of America* **134**, 3464–3475.

Morton, E.S. (1975) Ecological sources of selection on avian sounds. *The American Naturalist* **109**, 17–34.

National Research Council (2005) *Marine Mammal Populations and Ocean Noise: Determining when Noise Causes Biologically Significant Effects*. National Academy Press, Washington, DC.

NOAA (2012) CetSound project. Available at: http://cetsound.noaa.gov/index.html [accessed 4 June 2014].

Nowacek, D.P., Thorne, L.H., Johnston, D.W. & Tyack, P.L. (2007) Responses of cetaceans to anthropogenic noise. *Mammal Review* **37**, 81–115.

Parks, S.E., Urazghildiiev, I. & Clark, C.W. (2009) Variability in ambient noise levels and call parameters of North Atlantic right whales in three habitat areas. *The Journal of the Acoustical Society of America* **125**, 1230–1239.

Payne, R. & Webb, D. (1971) Orientation by means of long range acoustic signaling in baleen whales. *Annals of the New York Academy of Sciences* **188**, 110–141.

Pijanowski, B.C., Villanueva-Rivera, L.J., Dumyahn, S.L., Farina, A., Krause, B.L., Napoletano, B.M., Gage, S.H. & Pieretti, N. (2011) Soundscape ecology: the science of sound in the landscape. *BioScience* **61**, 203–216.

Radford, C., Jeffs, A., Tindle, C. & Montgomery, J. (2008) Temporal patterns in ambient noise of biological origin from a shallow water temperate reef. *Oecologia* **156**, 921–929.

Remage-Healey, L., Nowacek, D.P. & Bass, A.H. (2006) Dolphin foraging sounds suppress calling and elevate stress hormone levels in a prey species the Gulf toadfish. *Journal of Experimental Biology* **209**, 4444–4451.

Rempel, R.S., Francis, C.M., Robinson, J.N. & Campbell, M. (2013) Comparison of audio recording system performance for detecting and monitoring songbirds. *Journal of Field Ornithology* **84**, 86–97.

Richardson, W.J., Greene, C.R., Malme, C.I. & Thompson, D.H. (1995) *Marine Mammals and Noise*. Academic Press, San Diego.

Robinson, S.P., Lepper, P.A. & Hazelwood, R.A. (2014) *Good Practice Guide for Underwater Noise Measurement*. NPL Good Practice Guide No. 133. National Physical Laboratory, UK.

Rolland, R.M., Parks, S.E., Hunt, K.E., Castellote, M., Corkeron, P.J., Nowacek, D.P., Wasser, S.K. & Kraus, S.D. (2012) Evidence that ship noise increases stress in right whales. *Proceedings of the Royal Society of London. Series B* **279**, 2363–2368.

Rudnick, D.L., Davis, R.E. & Eriksen, C.C. (2004) Underwater gliders for ocean research. *Marine Technology Society Journal* **38**, 73–84.

Simpson, S.D., Meekan, M., Montgomery, J., McCauley, R. & Jeffs, A. (2005) Homeward sound. *Science* **308**, 221.

Sousa-Lima, R., Norris, T.F., Oswald, J.N. & Fernandes, D.P. (2013) A review and inventory of fixed autonomous recorders for passive acoustic monitoring of marine mammals. *Aquatic Mammals* **39**, 23–53.

Stumpner, A. & von Helversen, D. (2001) Evolution and function of auditory systems in insects. *Die Naturwissenschaften* **88**, 159–170.

Sun, J.W.C. & Narins, P.M. (2005) Anthropogenic sounds differentially affect amphibian call rate. *Biological Conservation* **121**, 419–427.

Towsey, M., Parsons, S. & Sueur, J. (2014) Ecology and acoustics at a large scale. *Ecological Informatics* **21**, 1–3.

Van Parijs, S.M., Clark, C.W., Sousa-Lima, R.S., Parks, S.E., Rankin, S., Risch, D. & Van Opzeeland, I.C. (2009) Management and research applications of real-time and archival passive acoustic sensors over varying temporal and spatial scales. *Marine Ecology Progress Series* **395**, 21–36.

Waser, P.M. & Brown, C.H. (1986) Habitat acoustics and primate communication. *American Journal of Primatology* **10**, 135–154.

Wenz, G.M. (1961) Some periodic variations in low-frequency acoustic ambient noise levels in the ocean. *The Journal of the Acoustical Society of America* **33**, 64–74.

Wiley, R.H. & Richards, D.G. (1978) Physical constraints on acoustic communication in the atmosphere: Implications for the evolution of animal vocalizations. *Behavioral Ecology and Sociobiology* **3**, 69–94.

Wilson, B., Benjamins, S. & Elliott, J. (2013) Using drifting passive echolocation loggers to study harbour porpoises in tidal-stream habitats. *Endangered Species Research* **22**, 125–143.

Received 23 September 2014; accepted 8 December 2014
Handling Editor: David Hodgson

Supporting Information

Additional Supporting Information may be found in the online version of this article.

Appendix S1. PAMGuide tutorial.

Data S1. PAMGuide.zip - Zipped archive of R and MATLAB codes for PAMGuide.

Measuring acoustic habitats
Appendix S1

Nathan D. Merchant, Kurt M. Fristrup, Mark P. Johnson, Peter L. Tyack, Matthew J. Witt, Philippe Blondel, Susan E. Parks

Methods in Ecology and Evolution

This appendix is a tutorial on how to analyse passive acoustic data to produce measurements of acoustic habitats using *PAMGuide*, the software that accompanies the above article and which can be run in MATLAB or R. Statistical analysis of passive acoustic data and the application of metrics to different monitoring scenarios are discussed in the *Habitat characterisation* section of the main text.

Contents

1 Introduction **3**

2 Quick start guide **4**

2.1 Initialising *PAMGuide* 4
2.2 Running *PAMGuide* 5
2.3 Calibrated analysis 7
2.4 Verifying *PAMGuide* output using test files 9
2.5 Adding time stamps 11
2.6 Analysing large files 13
2.7 Batch processing 13
2.8 *Viewer* 15

3 Calibrating data **16**

3.1 Implementation 16

3.2 Background 18

4 Selecting analysis parameters **21**
4.1 Analysis type 21
4.2 Window type 22
4.3 Window length 23
4.4 Window overlap 24
4.5 Low- and high-frequency limits 25
4.6 Averaging data via the Welch method 26

5 Plot types **27**

6 Technical background **29**
6.1 Computing frequency spectra 29
6.2 TOLs and broadband analysis 32
6.3 Statistical metrics 33
6.4 Reducing time resolution 34
6.5 Computing the acoustic waveform 35
6.6 Impulse metrics 35
6.7 Sound exposure level 35
6.8 Signal-to-noise ratio and the system noise floor 36
6.9 Tonal signals 38

7 *PAMGuide* index **39**
7.1 MATLAB 39
7.2 R 41

References **42**

1 Introduction

The tutorial is structured in three parts:

1. A quick start guide (Section 2), which gives straightforward instructions on getting started with *PAMGuide*.
2. Detailed instructions on calibrating data (Section 3), selecting analysis parameters (Section 4), and available plot types in *PAMGuide* (Section 5).
3. Technical background (Section 6), which provides the underlying signal processing steps and equations that *PAMGuide* uses.

For reference, Fig. 1 summarises the workflow of *PAMGuide* and the plots produced for each output metric. An index of the parameters in *PAMGuide* is provided in Section 7.

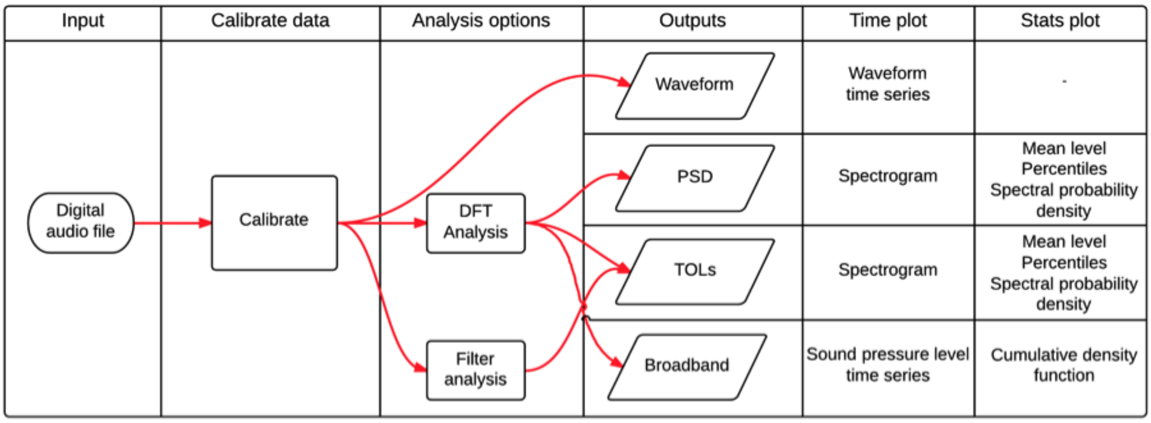


Figure 1: Flow diagram for *PAMGuide* illustrating processing steps and plot types corresponding to each output metric. DFT = discrete Fourier transform; PSD = power spectral density; TOL = 1/3-octave band level.

2 Quick start guide

The section covers the basics to get started with analysing sound files in *PAMGuide*. Subsequent sections contain detailed discussion of calibration procedures (Section 3, p16), selection of analysis parameters (Section 4, p21), and plotting options (Section 5, p27).

Table 1: Inventory of *PAMGuide* files.

MATLAB		**R**
PAMGuide.m	PG_TOL.m	PAMGuide.R
PAMGuide.fig	PG_Viewer.m	Meta.R
PG_Func.m	PG_Waveform.m	PAMGuide_Meta.R
PG_DFT.m		Viewer.R
Test files		
WhiteNoise_10s_48kHz_+-0.5.wav	Sine_10s_48kHz_+-0.5.wav	

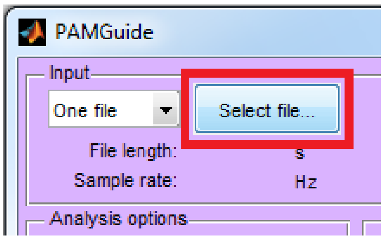
2.1 Initialising *PAMGuide*

Launch MATLAB or R.

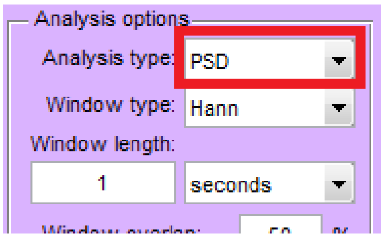
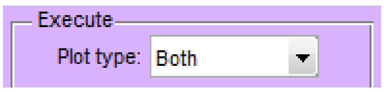
MATLAB	**R**
Set current folder to folder containing *PAMGuide* scripts:	Set current folder to folder containing *PAMGuide* scripts:
`>> cd('`*folderpath*`')`	`> setwd('`*folderpath*`')`
Launch *PAMGuide*:	Initialise *PAMGuide*:
`>> PAMGuide`	**`> source('PAMGuide.R')`**
PAMGuide graphical user interface (GUI) will appear (see Fig. 6, p40).	This command loads the *PAMGuide* function into the workspace, meaning it can then be called from the command line. Note: if not already installed, R will install package 'tuneR' on first execution.

The current folder ***must* be set to the *PAMGuide* folder** for *PAMGuide* to operate.

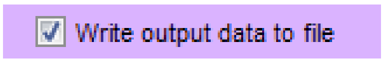
2.2 Running *PAMGuide*

MATLAB	R
Click **Select file...**. A dialogue box will appear. Select the WAV file for analysis. AIFF files can also be analysed if using MATLAB R2014 or later. 	Call *PAMGuide*: `> PAMGuide()` A dialogue box will appear. Select WAV file for analysis. *PAMGuide* will then display settings and analysis progress in the command line.
Click **RUN**. By default, *PAMGuide* will carry out a PSD analysis, and then plot the time representation (spectrogram) and statistical representations of the data. This uncalibrated analysis is plotted in relative units. For absolute levels, calibration data must be input (see Section 3, p16). 	By default, *PAMGuide* will carry out a PSD analysis, and then plot the time representation (spectrogram) and statistical representations of the data. This uncalibrated analysis is plotted in relative units. For absolute levels, calibration data must be input (see Section 3, p16).

Continued over...

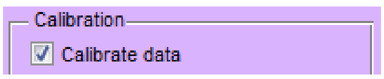
Other analysis types (e.g. broadband level, waveform) are selected in the **Analysis type** drop-down menu, and analysis parameters defined using the other inputs in the **Analysis options** panel. These options are discussed in detail in Section 4, p21.  	Other analysis options can be selected by specifying parameters when calling *PAMGuide*. For example: `> PAMGuide(atype=”13oct”, envi="Wat”,N=2048)` specifies a 1/3-octave band analysis, an underwater measurement, and a window length of 2048 samples, and the default parameters for all others. A full list of these parameters and their defaults is provided in Section 7, p39, and guidance on selecting parameters for different purposes is provided in Section 4, p21.
Plotting of time or statistical representations can be selected using the **Plot type** drop-down menu. Options: **Both**/**Time**/**Stats**/**None**. Default is **Both**.  See Section 5, p27, for examples of each plot type.	Plotting of time or statistical representations can be selected using **plottype**. Options: **Both**/**Time**/**Stats**/**None**. Default is **Both**. E.g.: `> PAMGuide(plottype=”Stats”)` See Section 5, p27, for examples of each plot type.
Figures can be saved from the figure window. Click **File>Export setup...**, then in the **Rendering** tab, choose resolution (300 dpi or greater for publishing quality), and click **Export...**. Image format can be chosen in the subsequent dialogue box.	Figures can be saved as PDF files using **Files>Save As...**

Continued over...

PAMGuide can output analysis files in CSV format for subsequent use by selecting **Write output data to file** in the **Execute** panel.  These files can be read and time and statistical representations plotted using *Viewer* (see Section 2.8). The format of the output files is a heading row with the frequencies of each of the columns, while the first column contains the times of each of the rows, with the data commencing on the second row of the second column. The first cell in the first column stores metadata on the analysis parameters, which is used by *Viewer*.	*PAMGuide* can output analysis files in CSV format for subsequent use using **outwrite**. Default is 0 (i.e. no output file written). Output file is written to same directory as input file. `> PAMGuide(outwrite=1)` These files can be read and time and statistical representations plotted using *Viewer* (see Section 2.8). The format of the output files is a heading row with the frequencies of each of the columns, while the first column contains the times of each of the rows, with the data commencing on the second row of the second column. The first cell in the first column stores metadata on the analysis parameters, which is used by *Viewer*.

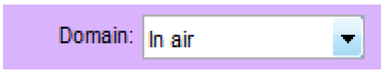
2.3 Calibrated analysis

For calibrated analyses, units will be expressed as absolute units, and relative to the relevant reference pressure if expressed in decibels (see Table 1 of main text).

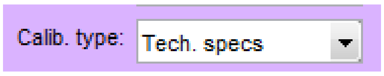
MATLAB	**R**
Select **Calibrate data** in the **Calibration** panel:  	Set calibration parameter, **calib**, to 1: `> PAMGuide(calib=1)`

Continued over...

<table>
<tr>
<td>Select Domain (In Air/Underwater):

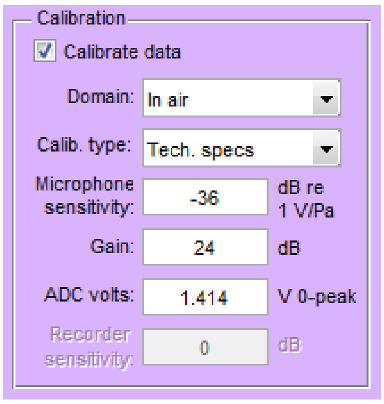

</td>
<td>Specify the environment parameter, envi, if an underwater measurement (default is Air)

<code>> PAMGuide(calib=1,envi="Wat")</code></td>
</tr>
<tr>
<td>Choose Calibration type. Options: Tech. specs/End-to-end/Recorder + Microphone/hydrophone:



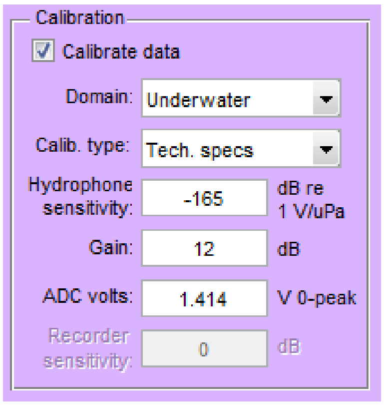
These options are discussed in detail in Section 3 (p16).</td>
<td>Specify calibration type, ctype (see Section 3 for details). Default is MF (manufacturer's technical specifications):

<code>> PAMGuide(calib=1,envi="Wat", ctype="MF")</code></td>
</tr>
<tr>
<td>Define calibration parameters. For example, if using the manufacturer's specifications plus the gain settings used in the deployment, select Tech. specs in the Calib. Type drop-down menu and input the specifications below. For a Wildlife Acoustics SM2+ in air with gain set to +24 dB, the calibration panel options would be:


</td>
<td>Define calibration parameters (see Section 3 for details). E.g. for in-air SM2+ recording with user-defined gain set to +24 dB:

<code>> PAMGuide(calib=1,envi="Air", stype="MF",mh=-36,g=24, vADC=1.414)</code></td>
</tr>
</table>

Continued over...

For an underwater measurement using a Wildlife Acoustics SM2M and user-defined gain of +12 dB, the parameters would be: 	For an underwater measurement using a Wildlife Acoustics SM2M and user-defined gain of +12 dB, the parameters would be: **> PAMGuide(calib=1,envi="Wat",stype="MF",mh=-165,g=12,vADC=1.414)**

2.4 Verifying *PAMGuide* output using test files

Two test files are supplied with *PAMGuide* which users can use to verify that the absolute levels produced by *PAMGuide* are correct: **WhiteNoise_10s_48kHz_+-0.5.wav** and **Sine_10s_48kHz_+-0.5.wav**. The verification process for each file is described below.

White noise

The first test file is a 10-second sample of white noise (random noise with a flat frequency spectrum): **WhiteNoise_10s_48kHz_+-0.5.wav**.

Using the following settings, the RMS level of the white noise test file should be 197.0 dB re 1 μPa. This metric is output in the command line during analysis as: `RMS level (mean SPL) = 197.0 dB re 1 uPa.`

In MATLAB, replicate the settings in Fig. 2; in R, use:

```
> PAMGuide(atype="Broadband",calib=1,envi="Wat",Mh=-200,
vADC=2,plottype="Time")
```

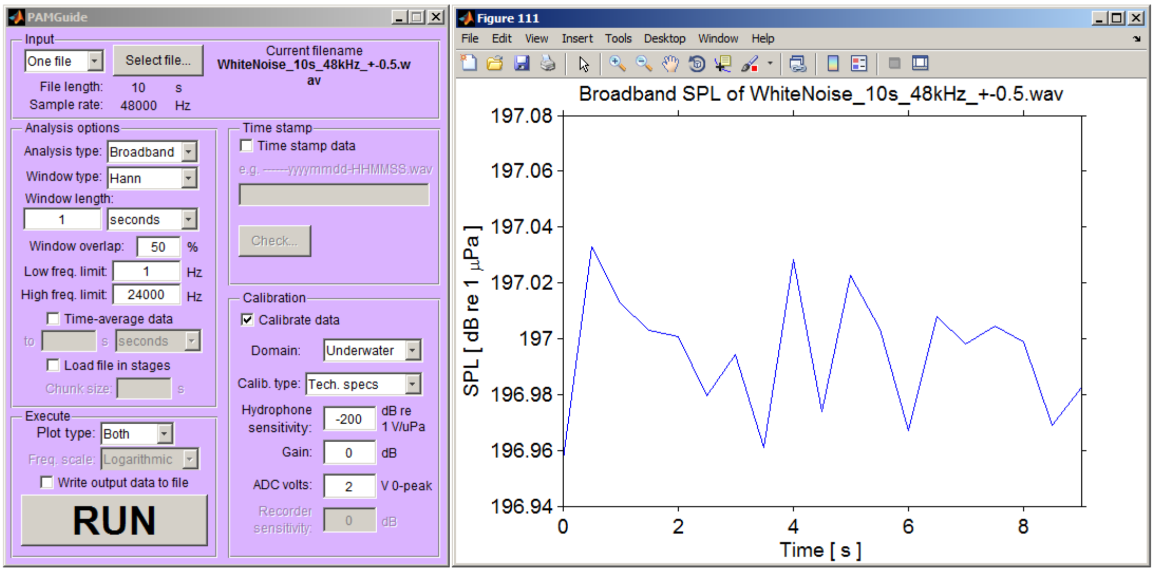


Figure 2: Settings and output for white noise verification file.

Sine wave

The second test file is a 10-second sine wave, or pure tone, at a frequency of 1 kHz: **Sine_10s_48kHz_+-0.5.wav**. For reasons explained in Section 6.9 (p38), the correct analysis to determine the amplitude of a pure tone is the power spectrum, which is closely related to the PSD (see Section 6.1, p29).

Using the following settings, the power spectrum amplitude of the sine wave test file should be 197 dB re 1 μPa^2 Hz^{-1} at 1 kHz. In MATLAB, replicate the settings in Fig. 3; in R, use:

```
> PAMGuide(atype="PowerSpec",calib=1,envi="Wat",Mh=-200,
vADC=2,plottype="Stats")
```

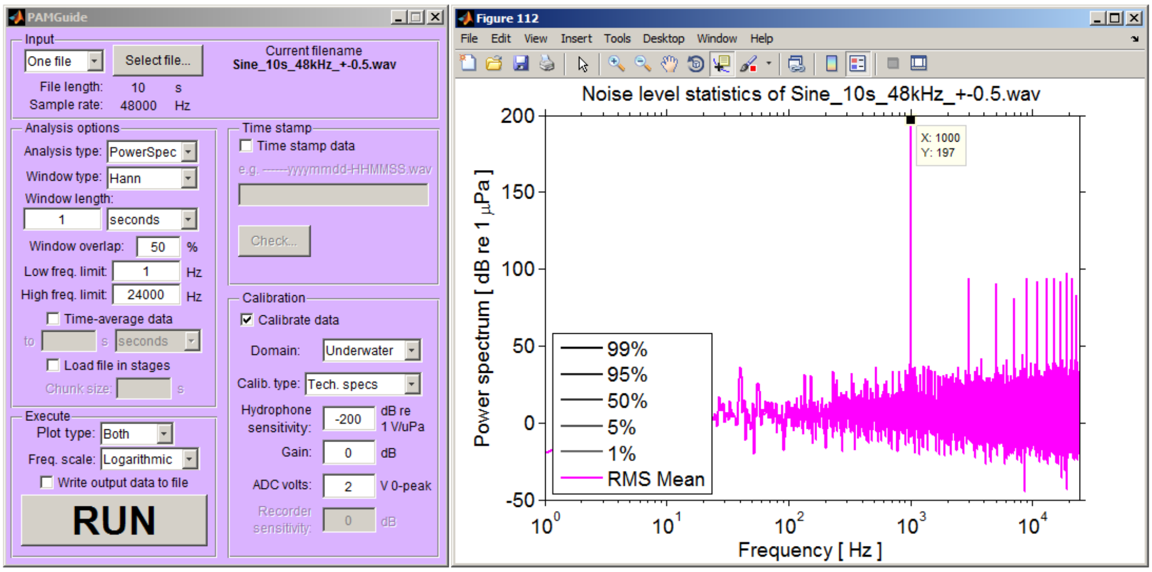
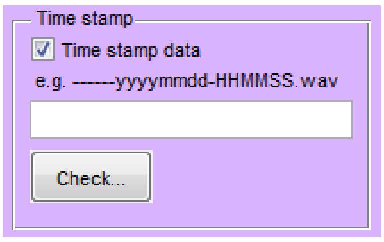


Figure 3: Settings and output for 1 kHz sine wave verification file.

2.5 Adding time stamps

Most autonomous passive acoustic recorders will embed a time stamp in the file name of each audio file. This information can be extracted in *PAMGuide* to produce analyses with absolute time axes. If time stamps are added, time series will be plotted with each file beginning at the time declared in the time stamp.

MATLAB	**R**
Select **Time stamp data** in the **Time stamp** panel: 	Time stamping is implemented if **timestring** is defined when *PAMGuide* is called. E.g. **`MyRecording_Depl456_20140707_141503.wav`** becomes:

Continued over...

In the text entry box, enter the name of the analysis file with digits corresponding to the time stamp replaced with letters (**y** = year, **m** = month, **d** = day, **H** = hour, **M** = minute, **S** = second, **F** = millisecond), and all other characters replaced with dashes, '**-**'. E.g.

`MyRecording_Depl456_20140707_141503.wav`

becomes

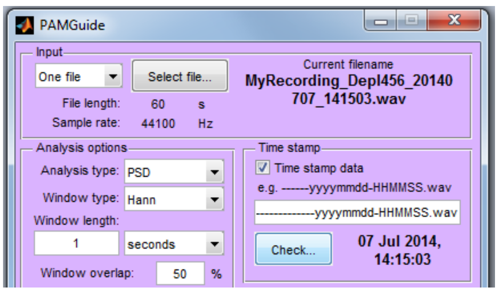
`-------------------yyyymmdd-HHMMSS.wav`

`> PAMGuide(timestring="MyRecording_Depl456_%Y%m%d_%H%M%S.wav")`

where %**Y** = year, %**m** = month, %**d** = day, %**H** = hour, %**M** = minute, %**S** = second.

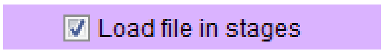
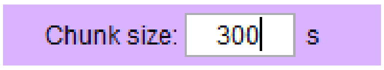
When batch processing multiple files (see Section 2.7), file names must have the same prefix before the time stamp.

If you have already selected a file, its name should be displayed in the **Input** panel. To check that the time stamp format is correct, click **Check...**, and the date and time will be displayed to the right, or else **ERROR: Check format** if the format cannot be interpreted.



2.6 Analysing large files

A common problem with analysing large sound files is that the host computer may not have enough RAM to load and analyse an entire file at once. To get around this, *PAMGuide* has the option to load and analyse files in stages. *PAMGuide* then outputs the plots and output files of the entire file as usual.

MATLAB	R
Select **Load file in stages** in the **Analysis options** panel:  Specify a chunk size in seconds by which to subdivide the input file:  	Specify the chunk size, **chunksize**, in seconds when calling *PAMGuide*: `> PAMGuide(chunksize=300)`

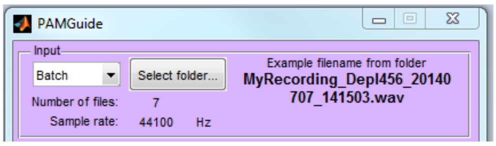
2.7 Batch processing

PAMGuide can batch process multiple files, and will produce contiguous plots of multiple files from the same deployment provided the sampling rate of the files (e.g. 44100 Hz) is the same for all files. Place all audio files for analysis in one folder.

MATLAB	R

Continued over...

Select **Batch** from the drop-down menu in the **Input** panel:

 	
Click **Select folder...** and select the folder for analysis in the dialogue box.

 	
The name of the first file in the selected folder will appear in the **Input** panel under **Example filename from folder**. This is the file that will be used for testing the time stamp format (see Section 2.5), if selected. The other settings used in the analysis can be defined as before. Then click **RUN**.
	In R, batch processing is achieved using *Meta.R*, contained in the *PAMGuide* folder.
	
	Initialise *Meta*:
	
	`> source('Meta.R')`
	
	The analysis settings for *Meta* are input in the same way as *PAMGuide*.
	
	Run *Meta*:
	
	`> Meta()`
	
	A dialogue box will appear. Select any file in the folder you wish to analyse, and *Meta* will gather file names for the whole folder.

If the files all have the same sampling rate, *PAMGuide* will attempt to concatenate them to form a single contiguous spectrogram and analysis file, which will be written to the input folder. If **Write output to data to file** is selected, analysis files for each individual file will also be written to a folder within the target folder. If **Plot type** is not set to **None**, plot(s) of the concatenated dataset will be produced. Plots can alternatively be produced by loading the analysis file into **Viewer** (see next section).	If the files all have the same sampling rate, *PAMGuide* will attempt to concatenate them to form a single contiguous spectrogram and analysis file, which will be written to the input folder. If **outwrite**=1, analysis files for each individual file will also be written to a folder within the target folder. If **plottype** is not set to **None**, plot(s) of the concatenated dataset will be produced. Plots can alternatively be produced by loading the analysis file into **Viewer** (see next section).

2.8 *Viewer*

Analysis files produced by *PAMGuide* (and, in R, *Meta*) can be subsequently viewed as time series and statistical representations using *Viewer*. *PAMGuide* encodes some analysis specifications in its output files so that this metadata can be used by *Viewer*.

MATLAB	**R**

Select **Viewer** from the drop-down menu in the **Input** panel:

 	
Select **Plot type** in the **Execute** menu (**Both**/**Time**/**Stats**/**None**).
 	
Click **RUN**.
	In R, *Viewer* is called from *Viewer.R*, contained in the *PAMGuide* folder.
	
	Initialise *Viewer*:
	
	`> source('Viewer.R')`
	
	Call *Viewer*, optionally specifying **plottype** as for *PAMGuide* (default = **Both**):
	
	`> Viewer(plottype="Stats")`

3 Calibrating data

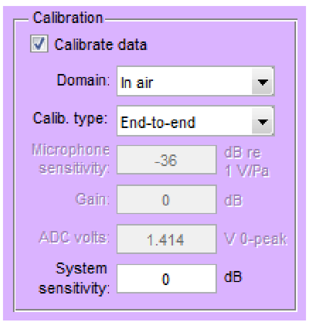
Calibrating data to produce absolute sound levels is essential to produce meaningful measurements of acoustic habitats for comparisons with other sites and other studies. Calibration makes it possible to quantify how much sound pressure a given level in the digital audio files corresponds to. There are three options for achieving this in *PAMGuide*:

1. Providing the end-to-end sensitivity of the entire recording system (from microphone/hydrophone to digital scale).
2. Providing the sensitivity of the recording device and the sensitivity of the microphone/hydrophone separately (a generic sensitivity for the transducer is usually given by the manufacturer).
3. Using the technical specifications of the recorder and transducer as provided by the manufacturer.

3.1 Implementation

This section describes how to input data for the above three options in *PAMGuide*; procedures to measure the parameters described are detailed in the Background section (p18).

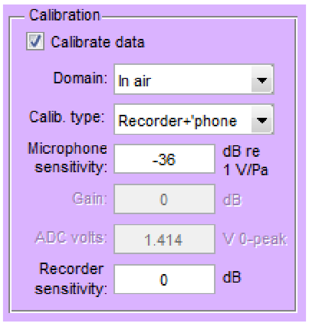
1. End-to-end

MATLAB	**R**
Select **End-to-end** from the **Calib. type** drop-down menu in the **Calibration** panel: 	When calling *PAMGuide*, define the calibration type **ctype** as **EE** (end-to-end) and the measured sensitivity of the entire device, **Si**, in decibels, e.g. `> PAMGuide(ctype=”EE”,Si=-159)`

Continued over...

Input measured sensitivity of entire device in the **System sensitivity** text box in decibels.	

2. Recorder + transducer

<table>
<tr><th>MATLAB</th><th>R</th></tr>
<tr><td>Select Recorder+'phone from the Calib. type drop-down menu in the Calibration panel:


Input microphone/hydrophone sensitivity in the text box and the measured sensitivity of the recorder in the Recorder sensitivity text box in decibels.</td>
<td>When calling PAMGuide, define the calibration type, ctype, as RC (recorder+'phone) and the measured sensitivity of the recorder, Si, in decibels, e.g.

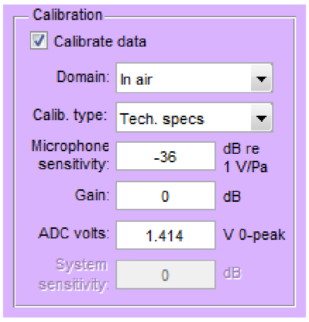
<code>> PAMGuide(ctype=”RC”,Mh=-36, Si=-20)</code></td></tr>
</table>

3. Manufacturer's specifications

MATLAB	R

Continued over...

Select **Tech. specs** from the **Calib. type** drop-down menu in the **Calibration** panel:



Input microphone/hydrophone sensitivity in the text box, user-defined **Gain** for the deployment in dB, and zero-to-peak voltage of the analogue-to-digital converter in the **ADC volts** text box.

When calling *PAMGuide*, define the calibration type **ctype** as **TS** (manufacturer's technical specifications), the sensitivity of the transducer, **Mh**, the user-defined gain settings for the deployment in decibels, **G**, and the zero-to-peak voltage of the analogue-to-digital converter, **vADC**:

```
> PAMGuide(ctype=”MF”,Mh=-36,
G=24,vADC=1.414)
```

3.2 Background

The general form of the calibration procedure converts the uncorrected signal, b, into a calibrated sound level, a, via a correction factor, S:

$$a = b - S \tag{1}$$

Here, we describe how S can be measured and/or calculated. Fig. 4 shows the signal path and conversely the calibration sequence for a typical PAM recorder, which helps to contextualise the equations presented below.

The most precise method of producing calibrated sound level measurements is to perform an end-to-end calibration of the system using a calibrated acoustic source, e.g. a reference microphone/hydrophone or a pistonphone (also known as a sound calibrator). The source is used to generate a sinusoidal sound pressure signal of known frequency, f, and zero-to-peak pressure amplitude, p_{peak}, at the system transducer (microphone or hydrophone), and this known level is then compared to the analysed signal from the system to calculate the correction factor, $S(f)$, in dB:

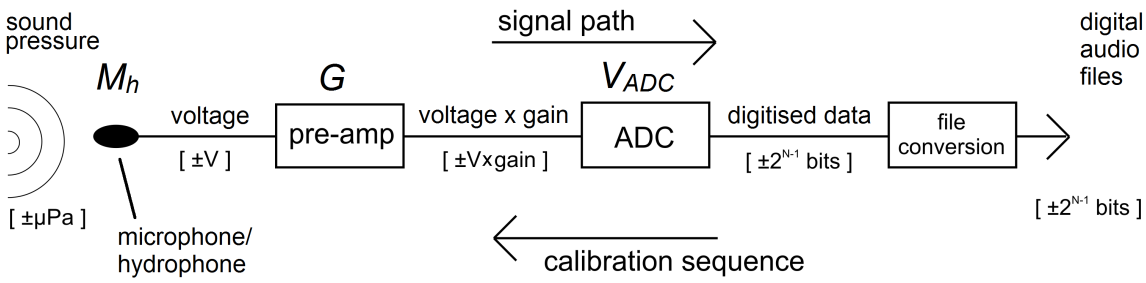


Figure 4: Signal path and calibration sequence for a typical PAM system. Values in square brackets indicate signal units at each processing stage.

$$S(f) = PP(f) + 3 - 20\log_{10}(p_{peak}) \tag{2}$$

where $PP(f)$ is the power spectrum (see Eq. 10 for details) of the digital signal, x_{bit}, at the test frequency, f, and the factor of 3 accounts for the 3 dB difference between the peak pressure amplitude, p_{peak}, and the root-mean-square (RMS) amplitude given by the power spectrum.

Alternatively, the recording system can be calibrated separately from the transducer using a signal generator. The sensitivity of the acoustic transducer is then either measured separately using a standard method (see ANSI 2005, 2012), or is taken to be the manufacturer's declared sensitivity at the test frequency, $M_h(f)$, which describes how much voltage is generated by the transducer per unit of sound pressure. To calibrate the recording system, a sinusoidal voltage of known frequency, f, and zero-to-peak voltage amplitude, V_{peak}, is applied to the transducer input of the system, and the correction factor, $S(f)$, is then

$$S(f) = PP(f) + 3 - 20\log_{10}(V_{peak}) + M_h(f) \tag{3}$$

where $M_h(f)$ has units of dB re 1 V/μPa. Note that microphone sensitivities are usually given in dB re 1 V/Pa, so a correction of -120 dB (equivalent to the 10^6 difference between Pa and μPa) is needed to convert to dB re 1 V/μPa, e.g. a microphone sensitivity of -36 dB re 1 V/Pa = -156 dB re 1 V/μPa.

Finally, the correction factor can also be computed directly from the system specifications, though this is the least reliable of the three methods described here. The specifications required are: the transducer sensitivity, $M_h(f)$, in dB re 1 V/μPa; the system gain, $G(f)$, in dB, at the frequency of interest, which describes any preamplifier gain applied to the signal; and the zero-to-peak voltage, V_{ADC}, of the analogue-to-digital converter (ADC). The correction factor, $S(f)$, is then

$$S(f) = M_h(f) + G(f) + 20\log_{10}\left(\frac{1}{V_{ADC}}\right) + 20\log_{10}(2^{N_{bit}-1}) \tag{4}$$

where N_{bit} is the bit-depth of the digital signal (e.g. 16, 24). Note that some software (e.g. MATLAB) normalises the amplitude to ± 1 when loading digital audio files, though the native format contains integers in the amplitude range $-2^{N_{bit}-1}$ to $2^{N_{bit}-1} - 1$. If such normalisation has been applied to the digital signal, the final term in Eq. 4 is omitted.

As an example of obtaining the manufacturer's specifications for Eq. 4, on p49 of the Wildlife Acoustics SM2+ user manual*, the manufacturer provides the microphone sensitivity, M_h, at 1 kHz as "Sensitivity: -36±4dB (0dB=1V/pa@1KHz)", meaning $M_h(f) = -36$ dB re 1 V/μPa at 1 kHz and within the surrounding flat frequency range of the microphone. On p47 of the manual, the ADC voltage is quoted as: "ADC: 1V rms full-scale 16-bit", meaning that $N_{bit} = 16$ and the *rms* voltage of the ADC is 1 V. V_{ADC} is the *zero-to-peak* voltage, which is given by multiplying the rms voltage by a conversion factor of $\sqrt{2}$, yielding $V_{ADC} = 1.41$ V. The system gain, $G(f)$, is defined by the user. Using these values and assuming a user-defined system gain $G(f) = 0$ dB, Eq. 4 yields $S = 51.3$ dB for this example.

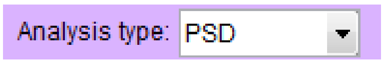
While the above correction factors have been defined for specific frequencies, in practice, the frequency variation of the system sensitivity over the frequency range of the measurements may be considered sufficiently low (e.g. ± 1 dB) that a single correction factor can be applied at all frequencies within the range. Such measurements are said to be within the 'flat frequency response' of the system.

*http://www.wildlifeacoustics.com/images/documentation/SM2plus1.pdf

4 Selecting analysis parameters

PAMGuide can compute a range of acoustic metrics: broadband sound pressure level (SPL), power spectral density (PSD), 1/3-octave band levels (TOLs), and the acoustic waveform. This section describes how to define the parameters used in these analyses and how to select parameters to yield analyses with appropriate time and frequency resolutions (and file sizes) for the application. The merits of using particular acoustic metrics for specific applications are discussed in the main text (section: *Habitat characterisation*).

4.1 Analysis type

MATLAB	**R**
Select the desired **Analysis type** from the drop-down menu in the **Analysis options** panel:  Options: **PSD/TOL/Broadband/Waveform/TOLf**. Default = **PSD**.	When calling *PAMGuide*, define analysis type, **atype**. Options: **PSD/TOL/Broadband/Waveform/TOLf**. Default = **PSD**. E.g.: `> PAMGuide(atype="Broadband")`

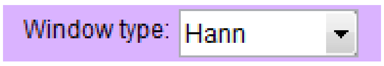
The main difference between PSD, TOLs, and broadband noise levels is in frequency resolution. PSDs yield the sound spectrum at linearly-spaced frequency intervals, and typically offer fine-scale frequency resolution of the spectrum (depending on the window length chosen, see below). As such, they are useful for characterising fine-scale frequency variation, such as in animal vocalisations and tonal components of anthropogenic sound sources.

TOLs are spaced logarithmically in frequency, such that the frequency doubles every three 1/3-octave bands (hence 1/3-octave). Particularly at high frequencies, TOLs have lower frequency resolution than PSDs, but are efficient to compute and produce a much smaller volume of data. The logarithmic frequency spacing (constant Q) is also characteristic of the mammalian auditory system, and so TOLs or other fractional octave bands may be particularly useful for habitat characterisation (see main text). The TOLf option implements the standard 1/3-octave computation method using filter analysis (rather than based on Fourier analysis, which is faster).

The broadband level reduces the frequency content over a given range (which can be specified in *PAMGuide*, see below) to a single dB level, and so is useful when simplicity is required, though the relevant frequency range should always be reported. The broadband option also outputs the median SPL, mode SPL, and SEL for each analysis in the command line.

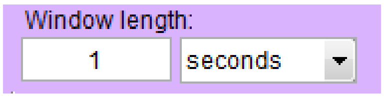
The pressure waveform is used to describe impulsive sounds (see main text) and sometimes animal vocalisations. As such, this analysis is typically applied to short clips of data.

4.2 Window type

MATLAB	R
Select the desired **Window type** from the drop-down menu in the **Analysis options** panel:  Options: **Hann/Rectangular/ Hamming/Blackman**. Default = **Hann**.	When calling *PAMGuide*, define analysis type **winname**. Options: **Hann/Rectangular/Hamming/ Blackman**. Default = **Hann**. E.g.: `> PAMGuide(atype="Hamming")`

For analyses based on the discrete Fourier transform (DFT), namely the **PSD**, **TOL** and **Broadband** options (see Fig. 1), the acoustic data is divided into a series of time segments. These segments may or may not be overlapping in time (see below). It is necessary to apply a window function to these time segments to control spectral leakage, an effect whereby energy from each frequency band spreads erroneously into other frequencies. This effect is caused by the finite length of the time segments, and is an inherent feature of discrete Fourier analysis. If no window function is applied, the window is effectively a rectangular (Dirichlet) window. However, for some signals a rectangular window can introduce broad spectral leakage and correspondingly high amplitude inaccuracies, so it is generally not a good choice for monitoring of acoustic habitats, where the nature of the sound recorded is often unknown *a priori*. Instead, a 'general purpose' window such as the Hann window is advisable (Cerna and Harvey, 2000). *PAMGuide* also provides Hamming and Blackman windows as alternatives. For a detailed discussion of windows for DFT analysis, see the review by Harris (Harris, 1978). More robust windowing techniques also exist, such as multitapering, whereby errors in spectral estimation are reduced by applying multiple windows to each segment (Thomson, 1982).

4.3 Window length

MATLAB	**R**
Enter the desired **Window length** in the text box in the **Analysis options** panel:  **Window length** can be input in seconds or samples (use drop-down menu to switch). If the input file is already loaded, *PAMGuide* will convert between these values an update the text box automatically.	When calling *PAMGuide*, define the window length, N, measured in samples. E.g.: `> PAMGuide(N=2048)` Default is the input file sample rate, Fs (= 1 second).

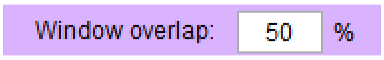
For PSD analysis, longer window lengths give higher frequency resolution but lower temporal resolution, and vice-versa. A compromise therefore needs to be made regarding how much time and frequency resolution is required. As a guide, if N is the frequency resolution in samples, the frequency bins are separated by Fs/N Hz, where Fs is the sampling rate in Hz. Consequently, $N > Fs$ yields frequency resolution < 1 Hz, and $N < Fs$ yields frequency resolution > 1 Hz. The default value for *PAMGuide* is $N = Fs$, yielding a frequency resolution of 1 Hz. Note that varying the window length changes the amplitude of the DFT on which the PSD is based, but the PSD is a stand measure which scales these levels to the equivalent level given by $N = Fs$.

If N is a power of 2, the DFT can be computed more rapidly, and so historically, 512- and 1024-point DFTs have been common. Such DFTs are known as Fast Fourier Transforms (FFTs), though the term is often applied to DFTs of any length. Advances in computational capabilities mean the relative inefficiency of $N{\neq}2^n$ is often no longer an obstacle.

Since the **TOL** and **Broadband** analysis options in *PAMGuide* are based on the

DFT, the window length can have some effect. If the window length is too short there will not be enough frequency bins to compute the TOLs, particularly at low frequencies. For this reason, *PAMGuide* limits the window length for TOL analysis to $\geq Fs$, which yields accurate TOLs for frequencies down to 25 Hz (Mennitt and Fristrup, 2012). The precision of the frequency range for broadband analysis (see below) will also depend on the frequency resolution of the DFT. For $N < Fs$, the true low frequency limit will be the next highest bin in frequency to the input value, and the true high frequency limit will be the next lowest bin in frequency. For $N > Fs$, the input values will be accurate to 1 Hz.

4.4 Window overlap

MATLAB	**R**
Enter the desired **Window overlap** (as a percentage) in the text box in the **Analysis options** panel: 	When calling *PAMGuide*, define the window overlap, **r**, as a percentage. E.g.: `> PAMGuide(r=75)` Default is 50%.

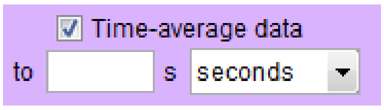
Overlapping the time windows ensure that brief sounds which would otherwise be split between consecutive time segments are represented, particularly if a window function which tapers to zero at its extremities is applied (e.g. the Hann window, see Section 4.2). A 50% overlap is typically sufficient for this purpose; higher overlaps can be used to smooth spectrograms in the time domain, but do not actually improve the temporal resolution, which is determined by the length of the time segment (see previous section). Since higher overlaps produce more data points in the time domain, they also increase the volume of data produced by the analysis.

4.5 Low- and high-frequency limits

MATLAB	R
Enter the desired **Low frequency limit** (in Hz) in the text box in the **Analysis options** panel:  Default is 1 Hz.	When calling *PAMGuide*, define the low frequency limit, **lcut**, in Hz. E.g.: `> PAMGuide(lcut=25)` Default is Fs/N, the lowest DFT frequency bin.
Enter the desired **High frequency limit** (in Hz) in the text box in the **Analysis options** panel:  Default is the Nyquist frequency, $Fs/2$ Hz, which will appear once the input file is loaded (since the sample rate, Fs, in unknown *a priori*).	When calling *PAMGuide*, define the high frequency limit, **hcut**, in Hz. E.g.: `> PAMGuide(hcut=10000)` Default is $Fs/2$, the Nyquist limit.

Upper and lower frequency limits can be applied to limit the analysis to the frequency range over which the system sensitivity is constant (the 'flat frequency response', see Section 6), or to produce a broadband analysis over a particular frequency range of interest (e.g. in the vocal range or auditory range of a particular species). The default high-frequency limit is the Nyquist frequency, $Fs/2$, which is the highest frequency that can be resolved for a given sample rate. If high-frequency content is not of interest for a particular analysis, reducing the high frequency limit from this default will reduce the data volume of the plots and output files.

4.6 Averaging data via the Welch method

MATLAB	**R**
Select the **Time-average data** tick box in the **Analysis options** panel:  Enter the desired averaging time in seconds, or select **Factor** from the drop-down menu and enter an integer number of spectra to average.	When calling *PAMGuide*, define averaging factor, **welch**, as an integer. E.g.: `> PAMGuide(welch=10)` This will average 10 spectra to create each output spectrum, reducing the time resolution by a factor of 10, i.e., for 1-s windows with a 50% overlap, the time resolution becomes 5 s. Data is not lost: the data within the longer time frame is averaged via the standard Welch method (see Section 6). Default is no averaging.

Time-averaging data produces more reliable measures of random signals such as sound levels in a habitat. It also saves storage and plotting time as it produces spectra with lower time resolution: this can be very significant for larger files. Time resolution can also be reduced by using longer time windows, but this does not save storage or plotting time since there is correspondingly greater frequency information (see Technical Background, p29). *PAMGuide* uses the Welch method (Welch, 1967), a standard method for time-averaging spectra.

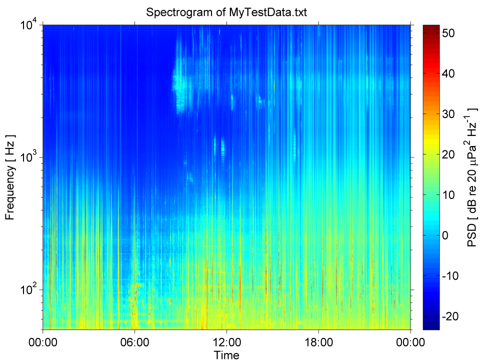
5 Plot types

Each analysis type (PSD, TOL, Broadband, Waveform) in *PAMGuide* has pre-defined plot types for time-domain and statistical representations (see Fig. 1 in the Introduction). This section illustrates what to expect for each analysis type, for calibrated vs. uncalibrated analyses, and for time-stamped and non-time-stamped analyses.

The example data shown is 24 hours of terrestrial recording made using a Wildlife Acoustics SM2+ in Harvard Forest, Massachusetts, USA, at 44.1 kHz.

<table>
<tr><th>Time</th><th>Stats</th></tr>
<tr><td>

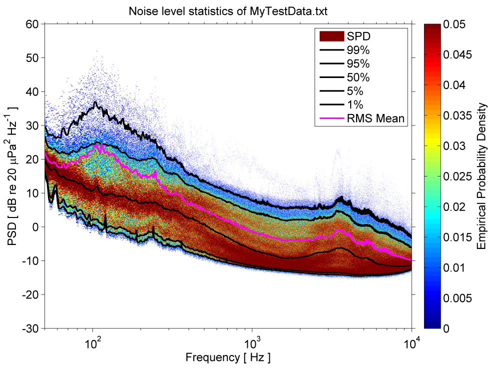
PSD, calibrated, time-stamped



Notes: 60-s Welch averaging applied, frequency limited to 50 Hz – 10 kHz.

</td><td>

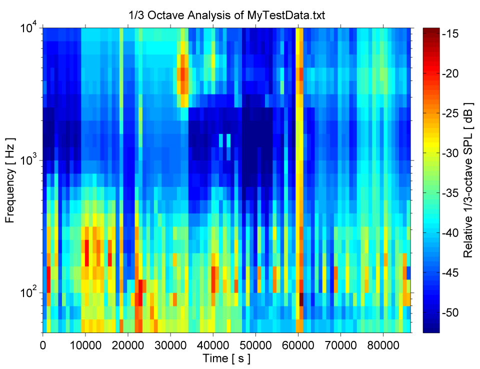
PSD, calibrated



Notes: *PAMGuide* only makes SPD plots for analyses with 1000 or more data points in the time domain.

</td></tr>
<tr><td>

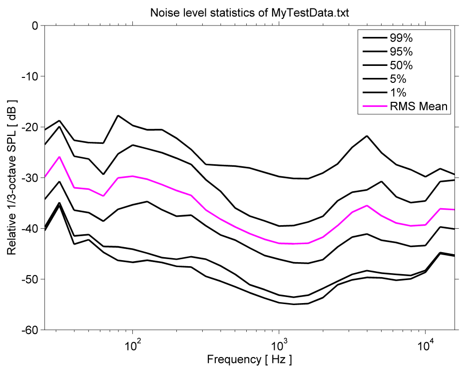
TOL, uncalibrated, not time-stamped



Notes: 900-s Welch averaging applied, frequency limited to 50 Hz – 10 kHz.

</td><td>

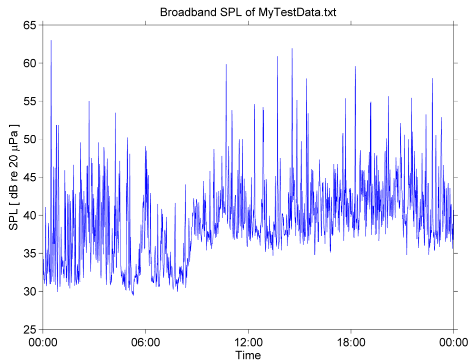
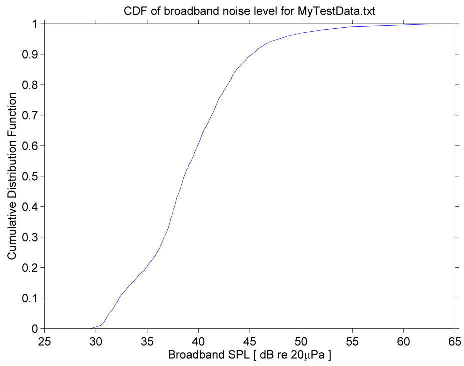
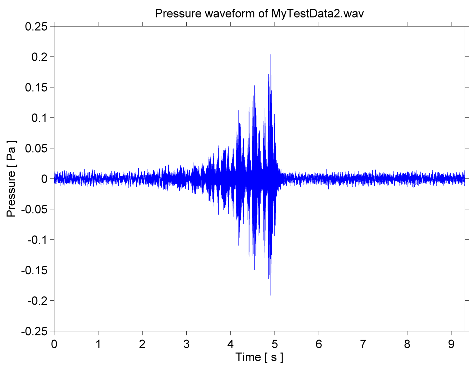
TOL, uncalibrated



Notes: as there were less than 1000 data points in the time domain, *PAMGuide* did not plot the SPD as in the PSD example above.

</td></tr>
</table>

Continued over...

Broadband, calibrated, time-stamped  Notes: frequency range is 50 Hz — 10 kHz.	*Broadband, calibrated*  Notes: frequency range is 50 Hz — 10 kHz. The broadband option also outputs the median SPL, mode SPL, and SEL for each analysis in the command line.
Waveform, calibrated, not time-stamped  Notes: this is a short extract of a bird call recorded at the site.	*Waveform* Notes: there is no statistical representation of the acoustic waveform in *PAMGuide*.

PSD plots have a logarithmic frequency scale by default, which emphasises lower frequencies. A linear frequency scale can be selected by changing **Frequency scale** to **Linear** (in MATLAB) or specifying **linlog**="Lin" (in R).

6 Technical background

This section describes the signal processing steps that *PAMGuide* is based on, detailing the equations used in the code to compute acoustic metrics. Additionally, important considerations such as signal-to-noise ratio (SNR), sound exposure level (SEL), and the analysis of tonal signals are discussed. The signal processing steps to calibrate digital audio files were detailed in Section 3, above.

6.1 Computing frequency spectra

PAMGuide uses two methods to analyse the frequency characteristics of sound: discrete Fourier transforms (DFTs) are used to derive the PSD and can be used to produce 1/3-octave band levels (TOLs), while filter banks can be used to implement the standard definition of TOLs (ANSI, 2009). For the PSD, the time-domain signal, x, is transformed into the frequency domain using an N-point DFT which yields the signal energy in N linearly-spaced frequency bins between 0 and the sampling frequency, F_s. To do this, the time-series signal is divided into segments of consecutive samples, each of which is then multiplied by a window function and then transformed via the DFT. This yields a frequency spectrum for each segment. These can then be combined and plotted as a spectrogram, showing the variation in frequency content through time, or can be averaged – as will be described below – to produce a two-dimensional plot of amplitude versus frequency, forming a longer term estimate of the spectrum.

The first step in computing the PSD is to divide the time-domain signal, x, into segments, which may be overlapping in time. The m^{th} segment is given by

$$x^m[n] = x[n + (1 - r)mN] \tag{5}$$

where N is the number of samples in each segment, $0 \leq n \leq N - 1$ (Marple, 1987), and r is the segment overlap expressed as a decimal (e.g. 0.75 for 75$\%$ overlap; see Section 4.4 for usage of overlap).

A window function is then applied to each data segment (see Section 4.2). Denoting the m^{th} windowed data segment $x^m_{win}[n]$

$$x^{(m)}_{win}[n] = \frac{w[n]}{\alpha} x^{(m)}[n] \tag{6}$$

where w is the window function over the range $0 \leq n \leq N-1$, and α is the coherent gain factor of the window (also known as the scaling factor), which corrects for the reduction in amplitude introduced by the window function (Cerna and Harvey, 2000).

The DFT of the m^{th} time segment is given by

$$X^{(m)}(f) = \sum_{n=0}^{N-1} x_{win}^{(m)}[n]\text{exp}\left(\frac{-j2\pi fn}{N}\right) \tag{7}$$

where $0 \leq f \leq F_s - F_s/N$, and F_s is the sampling frequency.

The power spectrum is computed from the DFT, and corresponds to the square of the amplitude spectrum (DFT divided by N), which for the m^{th} segment is given by

$$P^{(m)}(f) = \left|\frac{X^{(m)}(f)}{N}\right|^2 \tag{8}$$

For real sampled signals, the power spectrum is symmetrical around the Nyquist frequency, $F_s/2$, which is the highest frequency which can be measured for a given F_s. The frequencies above $F_s/2$ can therefore be discarded and the power in the remaining frequency bins are doubled, yielding the single-sided power spectrum

$$P_{SS}^{(m)}(f') = 2P^{(m)}(f') \tag{9}$$

where $0 \leq f' \leq F_s/2$. This correction ensures that the amount of energy in the power spectrum is equivalent to the amount of energy (in this case the sum of the squared pressure) in the time series. This method of scaling, known as Parseval's theorem, ensures that measurements in the frequency and time domain are comparable.

The power spectra for each of the data segments are then combined to form an array in frequency and time, PP

$$PP(f', m) = 10\text{log}_{10}\left(\frac{P_{SS}^{(m)}(f')}{p_{ref}^2}\right) - S(f') \tag{10}$$

where, for in-air measurements, p_{ref} is a reference pressure of 20 μPa, and for underwater measurements, p_{ref} is 1 μPa. The power spectrum is used to measure the amplitude of discrete frequency components (e.g. tonal vocalisations; see Section 9). However, for general characterisation of sound levels, a standardised version of the power spectrum, the PSD, is used (see below).

It is important to note that the number of samples, N, in each data segment determines number of frequency points in the DFT, and therefore the frequency resolution of the power spectrum ($\Delta f = F_s/N$ Hz). If N is smaller, there are fewer frequency bins (number of bins is N/F_s), and each bin encompasses a wider frequency range (F_s/N Hz), meaning more spectral energy is integrated per bin. For sound with a continuous frequency distribution such as the overall sound level in a habitat, the spectral power estimates from Eq. 10 vary with N. To account for this relationship, spectra are often standardised to the levels which would result from a data segment length of F_s (i.e. one second), which yields a frequency bin width of 1 Hz. This quantity is termed the power spectral density, PSD, and is defined by:

$$\text{PSD}(f', m) = 10\log_{10}\left(\frac{1}{B\Delta f}\frac{P_{SS}^{(m)}(f')}{p_{ref}^2}\right) \tag{11}$$

with units of dB re p_{ref} μPa^2 Hz^{-1}, where $\Delta f = F_s/N$ is the width of the frequency bins, and B is the noise power bandwidth of the window function, which corrects for the energy added through spectral leakage:

$$B = \frac{1}{N}\sum_{n=0}^{N-1}\left(\frac{w[n]}{\alpha}\right)^2 \tag{12}$$

For the Hann window, B=1.5 (see Harris (1978) for other windows). The PSD gives an estimate of the power in each frequency band of 1 Hz width even if the spectral analysis has a much coarser frequency resolution (i.e., if $N < F_s$). This is appropriate for the smoothly-varying spectra frequently encountered in ambient noise monitoring and the use of a standardized 1 Hz bandwidth ensures that measurements made with different frequency resolutions are comparable. However, the scaling applied to the PSD for the purposes of standardisation mean it will not give accurate amplitude estimates for narrowband signals such as tonals. To make accurate measurement of tonal signals, the unscaled power spectrum (Eq. 11) is instead used (see Section 6.9).

For some bioacoustics applications, the absolute amplitude of the signal is often not important, and spectra are displayed with arbitrary units (relative dB, often scaled to give a 0 dB maximum, such that amplitudes have negative dB values; see Fig. 1 in main text). To produce such measurements for a particular metric, the correction factor, S, should be the maximum decibel level measured.

For long-term monitoring, averaging of spectra in frequency and/or time is frequently used to produce more manageable data. One approach is to compute PSD with the desired frequency resolution and then average these to the desired time resolution (see Section 6.4). Alternatively, the frequency resolution can be reduced by computing TOLs (see Section 6.2).

6.2 TOLs and broadband analysis

Besides the PSD, the most common acoustic metrics for ambient noise analysis are TOLs and broadband SPL. These can be computed directly from the time-domain signal using filters, or by integrating the power spectrum, P_{SS}, over the frequency range of interest before it is converted to dB.

Historically, TOLs were computed using analogue filtering equipment, and the current standard maintains this approach using digital filters (ANSI, 2009). In the digital context, it is typically more computationally efficient to instead sum the relevant frequency bins in the power spectrum (Mennitt and Fristrup, 2012) which does not deviate appreciably from the standard except at low frequencies (difference is <1 dB for frequencies above 25 Hz; Mennitt and Fristrup 2012) which are below the range of interest for most species. In *PAMGuide*, either approach can be implemented, though only the DFT-based approach is implemented in the otherwise comparable R code (Appendix S3). The filter-based approach is specified in the relevant standard (ANSI, 2009); here, we describe the frequencies at which TOLs are assessed, and illustrate the power spectrum approach.

The 1/3-octave centre frequencies, f_c, are defined as

$$f_c = f_{ref} 10^{\frac{i-1}{10}} \tag{13}$$

where f_{ref} is a standardised reference frequency of 1 kHz, and i is an integer, such that $i \geq 1$ corresponds to $f_c \geq 1$ kHz, and $i < 1$ corresponds to $f_c < 1$ kHz (ANSI, 2009). The upper and lower bounds of each band are then given by

$$f_{upper} = f_c 10^{\frac{1}{20}} \tag{14}$$

$$f_{lower} = f_c 10^{\frac{-1}{20}} \tag{15}$$

Although the exact 1/3-octave frequencies are used for computation of noise levels, they are commonly referred to by nominal values, e.g. 1250 Hz for the band centred on 1258.9 Hz (ANSI, 2009). The SPL of the m^{th} data segment in the 1/3-octave band centred on f_c can then be estimated by summing the outputs of the power spectrum that are within this band:

$$\text{SPL}(m, f_c) = 10\log_{10}\left(\frac{1}{p_{ref}^2}\sum_{f'=f_{lower}}^{f'=f_{upper}}\frac{P_{SS}^{(m)}(f')}{B}\right) - S(f_c) \quad (16)$$

where $P_{SS}^{(m)}(f')$ is as defined in Eq. 9.

Similarly to 1/3-octave bands, broadband SPL is the integral of acoustic energy over a given frequency range. Rather than a frequency spectrum of levels, however, broadband SPL reduces the noise level to a single number, typically over a wider frequency range than 1/3-octave bands. This simplicity is useful, but can also be problematic, since broadband SPL is often reported without indicating the frequency range covered in the analysis, confounding interpretation of the results. Its single-dimensionality can also disguise problems with system sensitivity that would be apparent in spectral representations. To ensure the validity of broadband SPL measurements, the same care should be taken in correcting for system sensitivity and gain as detailed above for PSD measurements, and the frequency range of the measurement should be given. The broadband SPL of the m^{th} data segment (see Table 1 for units), is defined by:

$$\text{SPL}(m) = 10\log_{10}\left(\frac{1}{p_{ref}^2}\sum_{f'=f_{low}}^{f'=f_{high}}\frac{P_{SS}^{(m)}(f')}{B}\right) - S \quad (17)$$

where f_{low} and f_{high} are the lower and upper bounds of the frequency range under consideration.

6.3 Statistical metrics

Spectrograms can be averaged in time to produce two-dimensional plots of PSD vs. frequency. These are commonly used to summarise and compare noise levels at different times or locations, and to view specific frequency components with greater clarity. Though there are a number of averaging methods in use for ambient noise spectra (Merchant et al., 2012), the most common is the RMS level, where the mean is computed before it is converted to dB (i.e. the mean of P_{SS}, Eq. 9). This metric is sometimes termed 'RMS mean', 'linear-space mean', or 'arithmetic mean', and is the method used by commercial sound level meters. Here, we retain the notation used above, and define the mean of the PSD spectrogram, $\overline{\text{PSD}}$, by

$$\overline{\text{PSD}}(f') = 10\log_{10}\left(\frac{1}{p_{ref}^2}\frac{1}{M}\sum_{m=1}^{m=M}\frac{P_{SS}^{(m)}(f')}{B\Delta f}\right) - S(f') \quad (18)$$

The median level is also widely used, and is independent of whether it is computed before or after conversion to decibels. Mode levels can be computed as the maximum of the PSD (Eq. 11) probability distribution in each frequency bin. Note that these average measures can be applied to other spectral measures such as TOLs and to broadband levels in the same way.

The spectral probability density (Merchant et al., 2013) shows the distribution of noise levels based on the empirical probability density. It can be computed from the PSD as an array of normalized histograms corresponding to each frequency bin:

$$\mathrm{SPD}(f') = \frac{1}{hM} H\left(\mathrm{PSD}(f', m), h\right) \tag{19}$$

where $H\left(\mathrm{PSD}(f', m), h\right)$ is the histogram of M noise levels in the PSD at frequency f', and h is the histogram bin width in units of dB re 1 μPa^2 Hz^{-1}.

6.4 Reducing time resolution

For many applications, and particularly when plotting long time-series spectrograms, the PSD may produce an unmanageable volume of data. This can be remedied by time-averaging the spectrogram to a lower time resolution. This approach is generally more computationally efficient than carrying out the original analysis with longer time segments (which does not reduce the data quantity since the frequency resolution is correspondingly higher), and is known as the Welch method (Welch, 1967). The reduced time resolution spectrogram, PSD_W, of a full-resolution spectrogram of (time) length M is given by

$$\mathrm{PSD}_W(f', q) = 10\log_{10} \frac{1}{p_{ref}^2} \frac{R}{M} \sum_{m=(q-1)R+1}^{m=qR} \frac{P_{SS}^{(m)}(f')}{B\Delta f} - S(f') \tag{20}$$

where M is the total number of data segments, R is the downscaling factor (e.g. R = 10 to yield 10 times fewer data segments), $Q = M/R$ is the number of averaged time segments produced, and $1 \leq q \leq Q$. Each segment of PSD_W thus consists of the mean of R full-resolution segments averaged in linear space. The same process can be applied to 1/3-octave levels and broadband SPLs.

6.5 Computing the acoustic waveform

Some acoustic measurements (e.g. peak-to-peak pressure measurements of impulses; see next section) are performed on the *pressure waveform*: the time-domain acoustic pressure signal, $x_{\mu Pa}$, that was recorded by the hydrophone.

The acoustic pressure waveform, $x_{\mu Pa}$, can be computed from the digital audio signal, x_{bit}, using the calibration correction factor (see Section 3):

$$x_{\mu Pa} = \frac{x_{bit}}{10^{\frac{S}{20}}} \tag{21}$$

where S is the correction factor and x_{bit} is the digital (bit-scaled) signal, such that the amplitude range is $-2^{N_{bit}-1}$ to $2^{N_{bit}-1} - 1$.

6.6 Impulse metrics

Impulsive sounds are typically described using peak-to-peak or zero-to-peak sound pressure levels (here denoted SPL_{p-p} and SPL_{0-p}), which are computed directly from the acoustic pressure waveform, $x_{\mu Pa}$ (Eq. 4):

$$\text{SPL}_{p-p} = 10\log_{10}\left(\frac{(\max(x_{\mu Pa}) + |\min(x_{\mu Pa})|)^2}{p_{ref}^2}\right) \tag{22}$$

$$\text{SPL}_{0-p} = 10\log_{10}\left(\frac{\max(x_{\mu Pa}^2)}{p_{ref}^2}\right) \tag{23}$$

If necessary, the waveform should be cropped such that the impulse of interest has the highest amplitude in the time series. When measuring the levels of impulsive sounds, it is important to ensure that the measurements are made within the flat frequency response of the system, and to report the frequency range of the measurements so that they can be usefully interpreted.

6.7 Sound exposure level

Sound exposure level (SEL) is a cumulative measure of acoustic energy over a period of time. For impulsive sounds, it is computed from the pressure waveform, $\text{x}_{\mu Pa}$:

$$\text{SEL} = \sum_{0}^{t} \frac{x_{\mu Pa}^2}{p_{ref}^2 s} \tag{24}$$

where s is a reference time of 1 s. A time interval corresponding to the period between the occurrence of the 5^{th} and 95^{th} percentiles of cumulative $x^2_{\mu Pa}$ is often used (see main text, Section 4).

For continuous sound, the SEL can be computed as a summation of the power spectrum over a specified bandwidth of frequencies:

$$\text{SEL} = 10\log_{10}\left(\sum_{m=1}^{M}\sum_{f'=f_{low}}^{f_{high}}\frac{P_{SS}^{(m)}(f')/B}{p_{ref}^2 s}\right) - S = \text{SPL} + 10\log_{10}(T) \tag{25}$$

where f_{low} and f_{high} are the lower and upper bounds of the frequency bandwidth under consideration, and T is the total duration of the signal. To be meaningful and enable comparisons with other studies, SELs should be quoted with the period of time over which they were calculated.

6.8 Signal-to-noise ratio and the system noise floor

The signal-to-noise ratio (SNR) describes the difference between the amplitude of a sound and the background noise (or the system noise floor, see below). It is important to calculate the SNR since low SNR can affect the accuracy of amplitude measurements. The SNR is measured in dB and is approximated as the difference between the level of the signal plus noise and the noise level that would have been recorded in the absence of the signal. The noise level is usually approximated by the noise level immediately preceding or following the signal if it is transient of from a nearby frequency band if the signal is continuous. As a general rule, an SNR $\geq$ 10 dB yields reliable amplitude measurement.

The system noise floor is the sound level measured by a PAM device in the absence of sound. This property of PAM systems can have an important influence on the quality of sound measurements and so is important to estimate. The noise floor varies with the gain and filter settings, the sampling rate, and any processing that is performed in the device prior to storing the data and so should be estimated for each configuration. Although it may be possible to infer the noise floor from manufacturers' data, it is preferable, if possible, to measure it directly with the settings that will be used in the field. To do this, the device is set to record in a quiet location with the transducer isolated from sources of vibration. For in-air recorders, an acoustically isolated space (e.g., an anechoic chamber) is usually required for this measurement. As hydrophones are relatively insensitive in air, a quiet room will usually suffice for underwater sound recorders. The PSD of the resulting recording can be used to determine the level of the noise floor.

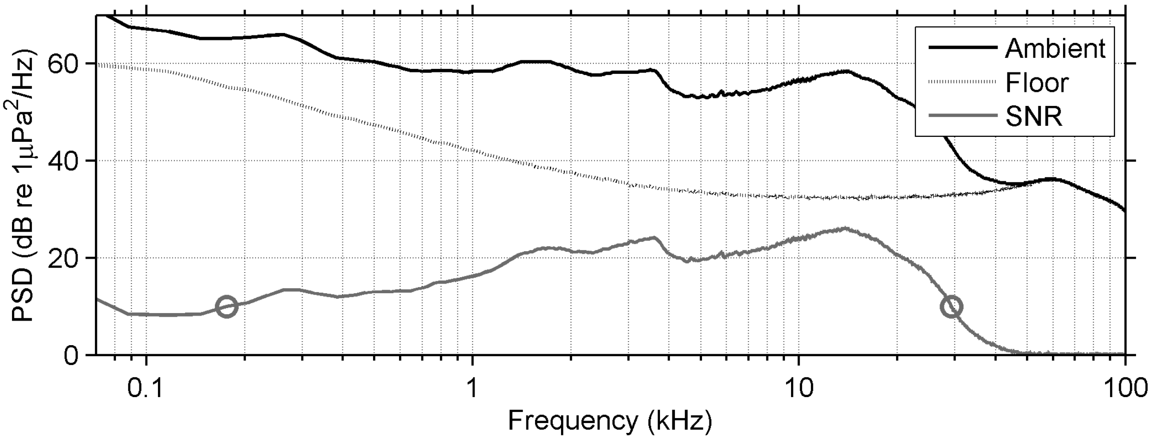


Figure 5: Power spectral density (PSD) and signal-to-noise ratio (SNR) of a 5-minute underwater noise recording in deep water off the Azores using a 8192-point FFT. Sampling rate: 240 kHz. 'Ambient' denotes the ambient noise signal (which contained vessel sounds and sperm whale clicks); 'Floor' is the system noise floor; 'SNR' is the difference between these in decibels. The ambient noise levels are only reliable where the SNR $\geq$ 10 dB, between 200 Hz and 29 kHz (marked by circles).

The system noise is one factor determining the suitability of a PAM device to make measurements in a given environment. A rule of thumb is that the system noise should be at least 10 dB below the signals of interest in each frequency band (e.g., each bin of the PSD) for spectral measurements to be reliable. For habitat monitoring applications, this implies that the system noise should be 10 dB lower than the lowest anticipated ambient noise at all frequencies of interest. This can be especially challenging to achieve at high frequencies, and most aquatic sound recorders will be system noise-limited at frequencies above about 20 kHz (Johnson, Partan, and Hurst, 2013). Because of this, the signal-to-noise ratio (SNR), i.e., the decibel difference between the signal and the noise floor, is a key metric to assess data quality that should be reported for all sound measurements. It is good practice to show the system noise floor or SNR on plots of ambient noise spectra (e.g., Fig. 5) to demonstrate the frequency range over which the system noise has little impact on the recorded levels. Whether the system noise floor has constrained the ability to record sound levels is often evident in plots of the spectral probability density (Merchant et al., 2013), wherein a sharp cut-off in sound levels recorded may be observed (see Fig. 3, main text).

6.9 Tonal signals

Unlike measurements of the overall sound level in a habitat, the amplitudes of discrete tonal signals (such as tonal vocalisations, blade-passing tones in ship or windmill noise, or sinusoidal tones used for equipment calibration) are measured either from the filtered time series or using the unscaled power spectrum (Eq. 10). If the power spectrum length, N, is chosen so that the tone frequency falls within a single frequency bin, the power spectrum level in that bin will represent the power amplitude of the plus any noise present in that bin. If the SNR within the bin is greater than 10 dB, the power spectrum level is a close estimate of the tonal power and the noise contribution can be neglected. If the tone frequency falls near the edge of a DFT band, it will contribute to the power spectrum level in more than one bin, and the power in both bins must be summed (before conversion to dB) to estimate the overall power of the tone.

7 *PAMGuide* index

This section provides an exhaustive list of *PAMGuide* parameters and their relevant page references in the tutorial.

7.1 MATLAB

Panel	Parameter	Page
Input	One file	5
	Batch	13
	Viewer	15
Analysis options	Analysis type	21
	Window type	22
	Window length	23
	Window overlap	24
	Low freq. limit	25
	High freq. limit	25
	Reduce time resolution	26
	Load file in stages	13
Time stamp	Time stamp data	11
Calibration	Calibrate data	7, 16
	Domain	8
	Calib. type	16
	'phone sensitivity	16
	Gain	16
	ADC voltage	16
	Recorder sensitivity	16
Execute	Plot type	6, 27
	Frequency scale	28
	Write output data to file	7

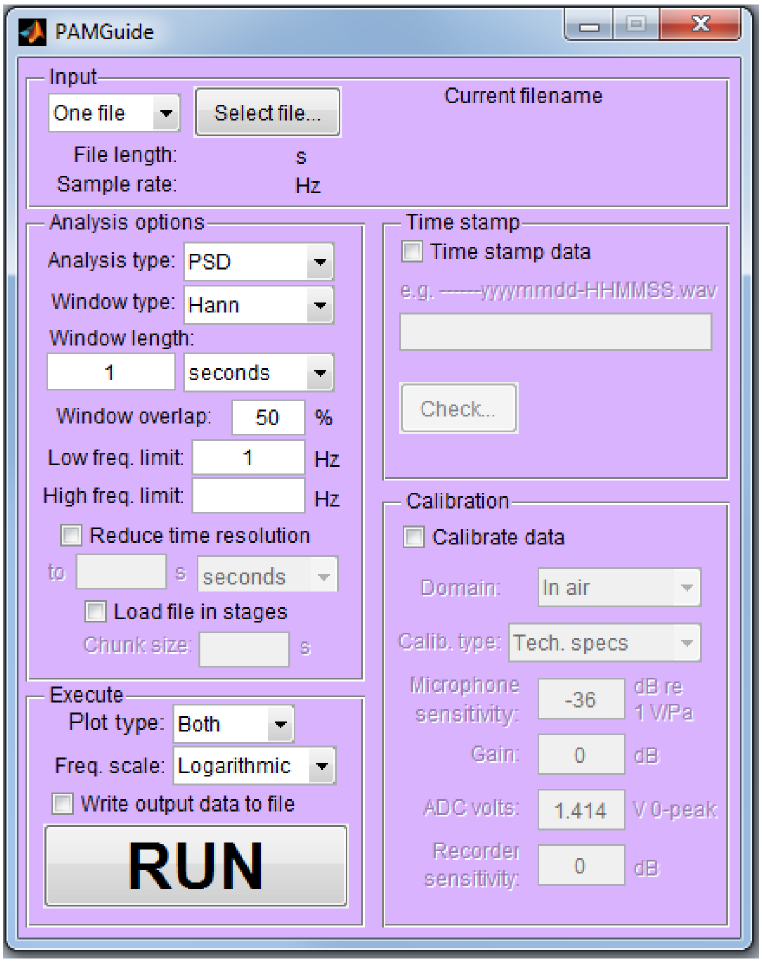


Figure 6: *PAMGuide* graphical user interface in MATLAB.

7.2 R

Parameter	**Default**, options	**Page**
atype	**"PSD"**, "TOL", "Broadband", "Waveform"	21
plottype	**"Both"**, "Time", "Stats", "None"	6, 27
envi	**"Air"**, "Wat"	8
calib	**0**, 1	7, 16
ctype	**"TS"**, "EE", "RC"	16
Si	-159	16
Mh	-36	16
G	0	16
vADC	1.414	16
r	50	24
N	Fs	23
winname	**"Hann"**, "Rectangular", "Hamming", "Blackmann"	22
lcut	Fs/N	25
hcut	$Fs/2$	25
timestring	""	11
outwrite	**0**, 1	7
welch	""	26
chunksize	""	13
linlog	**"Log"**, "Lin"	28

References

ANSI (2005) ANSI/ASA S1.15-2005/Part 2 (R2010) Measurement Microphones — Part 2: Primary Method for Pressure Calibration of Laboratory Standard Microphones by the Reciprocity Technique. American National Standards Institute.

ANSI (2009) ANSI/ASA S1.11-2004 (R2009) Specification For Octave-band And Fractional-octave-band Analog And Digital Filters. American National Standards Institute.

ANSI (2012) ANSI/ASA S1.20-2012 Procedures for Calibration of Underwater: Electroacoustic Transducers. American National Standards Institute.

Cerna, M. & Harvey, A.F. (2000) The fundamentals of FFT-based signal analysis and measurements, Application Note 041. Tech. rep.

Harris, F.J. (1978) On the use of windows for harmonic analysis with the discrete Fourier transform. *Proceedings of the IEEE* **66**, 51–83.

Johnson, M., Partan, J. & Hurst, T. (2013) Low complexity lossless compression of underwater sound recordings. *The Journal of the Acoustical Society of America* **133**, 1387–1398.

Marple, S.L. (1987) *Digital Spectral Analysis with Applications*. Prentice Hall Inc, NJ.

Mennitt, D.J. & Fristrup, K.M. (2012) Obtaining calibrated sound pressure levels from consumer digital audio recorders. *Applied Acoustics* **73**, 1138–1145.

Merchant, N.D., Barton, T.R., Thompson, P.M., Pirotta, E., Dakin, D.T. & Dorocicz, J. (2013) Spectral probability density as a tool for ambient noise analysis. *The Journal of the Acoustical Society of America* **133**, EL262–EL267.

Merchant, N.D., Blondel, P., Dakin, D.T. & Dorocicz, J. (2012) Averaging underwater noise levels for environmental assessment of shipping. *The Journal of the Acoustical Society of America* **132**, EL343–EL349.

Thomson, D.J. (1982) Spectrum estimation and harmonic analysis. *Proceedings of the IEEE* **70**, 1055–1096.

Welch, P. (1967) The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms. *IEEE Transactions on Audio and Electroacoustics* **15**, 70–73.