PASSIVE ACOUSTIC MONITORING FOR ESTIMATING ANIMAL DENSITY

Len Thomas

Centre for Research into Ecological and Environmental Modeling
University of St. Andrews, The Observatory, Buchanan Gardens, Fife KY16 9LZ, Scotland

and

Tiago A. Marques

Centre for Research into Ecological and Environmental Modeling
University of St Andrews, The Observatory, Buchanan Gardens, Fife KY16 9LZ, Scotland

and

Centro de Estatística e Aplicações da Universidade de Lisboa
Bloco C6, Piso 4, Campo Grande, 1749-016 Lisboa, Portugal

One of the most fundamental questions we can ask about a wildlife population is "How many are there?" Estimates of population size, or a related quantity population density (i.e., animals per unit area), are crucial for effective management, whether the management goal is conservation of a threatened or endangered species, control of a pest species, or optimal harvest of a species used for food. Population estimates are used to prioritize species of conservation concern, to monitor the success of management programs, and to set limits on harvest or incidental bycatch. Although "how many?" is a simple question to ask, it is often a hard one to answer, given that many populations are patchily distributed over very large areas and their lifestyle can make them quite cryptic to human observers. In this article, we introduce an emerging field with great potential—the estimation of wild animal population size and density using passive acoustics.

"Although "how many?" is a simple question to ask, it's often a hard one to answer …we introduce an emerging field with great potential: The estimation of wild animal population size and density using passive acoustics."

The potential for passive acoustic density estimation

Traditional methods for estimating animal numbers most often rely on visual surveys, where animals need to be seen to be counted. Probably the most common is a visually-based distance sampling survey (Buckland *et al.*, 2001), where observers visit a set of randomly placed transect lines or points, recording the distance to all detected animals of the target species. The distances are important because they allow us to estimate the probability of detecting animals, and hence account for the proportion missed (see later). Alternatively, if the animals have uniquely identifiable markings, then an alternative method can be used, called mark-recapture (or capture-recapture; Williams *et al.*, 2002). Here, observers visit the study area on a sequence of occasions, recording each animal detected and, if it was marked, which animal it was. The patterns of missing detections of marked animals can be used to estimate the probability of detection and hence infer how many animals have been missed completely. If animals do not bear natural markings, they can in some studies be caught and given marks (such as tags).

These traditional visual methods have been used in thousands of studies covering every major taxonomic group.[1] However, they do not work well in all circumstances. For one thing, some animals are inherently hard to see, for example because they live underwater or in thick forest, or because they are small, well camouflaged, or only active at night. Many do not have readily distinguishable markings and are hard to trap. In addition, traditional surveys can be very expensive, requiring trained observers and expensive survey vehicles to operate for extended periods in often far-flung and inhospitable environments. Think what it would cost to undertake a visual survey of marine mammals in the Southern Ocean in winter. For these reasons, there has been a keen interest in developing alternatives.

Passive acoustic monitoring offers one such alternative. Many species of animal produce distinctive sounds, either as part of a social display (e.g., to mark a territory or attract mates) or as an aid in navigation and foraging (e.g., echolocation). In many cases, these acoustic signals are detectable at greater distances than visual cues. Indeed, in environments where light does not pass easily, and hence visual methods are ineffective, animals are more likely to use sound as a means of communication, making them ideally suited for acoustic methods. Acoustic methods are also potentially less affected by weather conditions, and can operate under varying light levels, particularly at night. Another advantage is that the science of automated detection and classification of sounds is relatively well developed, as readers of this magazine will know, opening the possibility of automated processing of large volumes of remotely gathered data. By contrast, visual surveys are still almost universally performed by human observers, and even if digital imaging is used, classification is almost exclusively performed by humans (e.g., Buckland *et al.*, 2012).

The potential for passive acoustic data to be used to make inferences about animal populations has been recognized for decades (see, e.g., review by Mellinger *et al*. 2007). However, these inferences have largely been limited to confirming presence, or quantifying spatial and temporal patterns in relative indices such as vocalizations detected per day. Although useful in many contexts, a fundamental limitation of such measures is that they do not account for spatial and temporal variation in the acoustic detectability of animals and hence may not reflect true patterns of density or abundance. We aim, so far as possible, to account for such issues with the methods discussed in this article.

Much of our expertise in this area comes from a three-year research project, DECAF,[2] which focused on developing methods for estimating cetacean (whale and dolphin) density from fixed sensors. The examples presented here retain that bias. However the methods can readily be applied to terrestrial systems, and using mobile sensors, and we return to this at the end of the article. DECAF was a highly collaborative multi-agency, multi-disciplinary project bringing together an international team of statisticians, acousticians, engineers and biologists. We gratefully acknowledge the other team members: David Moretti, Ronald Morrissey, Nancy DiMarzio and Jessica Ward from the Navy Undersea Warfare Centre in Newport, Rhode Island; David Mellinger and Elizabeth Küsel from Oregon State University, Newport, Oregon; Stephen Martin from the Space and Naval Warfare Systems Centre Pacific, San Diego, California; David Borchers and Catriona Harris from the University of St. Andrews, Scotland; and Peter Tyack of Woods Hole Oceanographic Institution, Woods Hole, Massachusetts (now also at the University of St. Andrews).

Density estimation 101

Imagine the following general scenario. We have made acoustic recordings at a set of random locations throughout the region within which we were interested in estimating animal density. Together, these recordings survey an area *a* (the union of a set of circles around each recorder with radius large enough that no call from outside these circles can be detected). We have processed the recordings using detection and classification algorithms to produce a count, *n*, for example the number of a particular type of call.[3] To convert the count into a density, *D*, we use an equation of the form

$$\hat{D} = \frac{n}{a} \times \hat{m} \qquad (1)$$

where *m* represents a set of multipliers that convert count-per-unit-area (i.e., *n/a*) to animal density, and "hats" over quantities indicate that they are estimates.

In general, the multipliers that make up *m* do two jobs. First they account for inaccuracies in the detection process, i.e., false positive and false negative detections. The false positive rate is usually easiest to estimate, by hand-validating a sample of detections (assuming that we view a human analyst as the gold standard). False negative rate is often expressed as its complement, the detection probability; it is generally harder to determine, and will be the focus of much of this article. The second job of multipliers is to convert the object counted (e.g., a call) into the number of animals it represents. The exact nature of these second multipliers depends on the type of object counted. In some cases acoustic processing can yield the number of animals present (e.g., if animals have unique vocalizations or can be otherwise isolated)—hence the count is of animals and no multiplier of this type is needed. In other cases, it may be possible to count groups of animals, in which case the required multiplier is mean group size. Most commonly, however, the count is of sounds, such as calls or clicks, and the required multiplier is the sound production rate. This latter type of surveying is called "cue counting" in the statistical literature (the sound is an acoustic cue), and the multiplier is called the "cue rate."

Just as important as the estimate itself is a reliable characterization of uncertainty in the estimate. Your interpretation of a density estimate of 1 whale per 1,000 km[2] would be quite different if we told you that the 95% confidence interval was 0.8–1.2 versus 0.2–5.1 whales per 1,000 km[2]. Quanitifying uncertainty is also fundamental in testing for trends over time, differences between areas, etc. One common way to report uncertainty is as coefficient of variation, CV, which is the standard error of an estimate divided by the estimate, and usually reported in percent. From here, it's straightforward to calculate quantities like confidence intervals (see, e.g., Buckland *et al*. 2001, p.77). A CV of 10% on a density estimate is very good, corresponding to the kind of 95% confidence interval we gave first, above. A CV of 20% is reasonable, but by the time you get to CVs of 100% the estimate is nearly useless (as in the second confidence interval we gave above). The CV on a density estimate can be calculated easily given the CV on each random component making up the estimate (the *n* and each multiplier in *m*, assuming each is statistically independent).[4] It is our experience, however, that not enough attention is paid to this—all too often estimates are given without corresponding CVs or confidence intervals, or there is no discussion of how reliable the estimates are.

An example—Cue counting beaked whales in the Bahamas

To take a concrete example, Marques *et al*. (2009) estimated the density of Blainville's beaked whales (*Mesoplodon densirostris*) at Tongue of the Ocean, Bahamas over a 6-day period in spring 2005. This area contains a US Navy testing range, the Atlantic Undersea Testing and Evaluation Center (AUTEC) which is instrumented with a wide baseline array of 82 bottom-mounted hydrophones (see Fig. 1) cabled to shore, making it an ideal laboratory for bioacoustic studies. Blainville's beaked whales occur there, and are of particular concern to the Navy because there have been documented strandings of this and related species coincident with Navy exercises (D'Amico *et al.,* 2009). In common with other beaked whale species, they undertake long (~45 minute), deep (600-1200 m) foraging dives, during which they produce high frequency echolocation clicks to locate prey. Running a simple detection and classification algorithm on the sound recordings over the survey period logged *n* = 2,940,521 echolocation clicks. Marques *et al*.

assumed that no clicks at greater than 8 km horizontal range could be detected, so the area surveyed was $a = 82 \times \pi \times 8^2 = 16{,}487$ km². The problem was to build on these numbers to derive a density estimate.

Marques *et al.* (2009) used a cue count approach, treating the echolocation click as the cue. In addition to the above information, four multipliers were required. Two account for the detection process—false positive rate, c, and average detection probability, p. Two convert density of clicks to density of animals—the click production rate, r, and the time spent monitoring T. Hence the formula used was

$$\hat{D} = \frac{n}{a} \times \frac{(1 - \hat{c})}{\hat{p}\,\hat{r}\,T} \qquad (2)$$

This makes intuitive sense if we think of $n \times (1 - \hat{c})$ as an estimate of the *true* number of Blainville's clicks detected, accounting for false positives. Dividing this by $\hat{p}$ gives an estimate of the true number of Blainville's clicks produced, accounting for those missed. The quantity $\hat{r}\,T$ is an estimate of the average number of clicks produced per whale over the time T, so dividing the estimated total number of clicks produced $n \times (1 - \hat{c})\,/\,\hat{p}$ by the estimated number produced per whale $\hat{r}\,T$ gives an estimate of the number of whales. Dividing this by the area a gives an estimate of whale density, and hence the formula shown above.

The problem now becomes one of obtaining reliable values for the multipliers. Time spent monitoring was known ($T = 4{,}961$ minutes, accounting for some data outages), and so does not have a hat. Estimating the other multipliers required some additional data.

For false positives, a systematic random sample of 30 10-minute recordings was hand-screened for false detections, yielding an estimated false positive rate of $\hat{c} = 0.451$. This is rather high, due to the simple nature of the detection and classification system used, where dolphin clicks were often confused with beaked whale clicks. The fact that 30 random recordings were used enabled calculation of uncertainty in this estimate—the CV was 1.99%. In other words, although the average false positive rate was high, its' value was estimated quite precisely and so did not contribute much to the overall uncertainty in the density estimate. Because it was calculated from a random sample taken from the 6-day data set used for density estimation it is also unbiased—i.e., it is correct, on average, for the 6-day period in which we are interested. If we wanted to estimate density for some other time period, the safest thing would be to take a new systematic random sample of times from that period and re-calculate the false positive rate. False positive rates vary according to the density of interfering sounds—in our case dolphins were the main problem, and the density of dolphins could be completely different in another time period. This illustrates a general principle with multipliers—it is best to estimate them using data from the time and place from which you want to estimate density. For false positives this is relatively tractable (so long as we have some raw sound files or other pertinent information from the survey data, and trust human analysts as the gold standard). For other multipliers it may be less so.

A good example of this is the data used to obtain cue rate and detection probability in this study. These data come from a sample of Blainville's beaked whales fitted with digital recording tags, "DTAGs," in fall 2006 and summer 2007 as part of another study within the same area. DTAGs record sound at high sampling rate, as well as kinematic variables that enable the animal's track to be reconstructed.[5] The sound recordings were processed, and clicks produced by the tagged whales were counted, yielding an estimated cue rate of $\hat{r} = 0.407$ clicks/second. Each of the 21 dives was considered to be an independent sample, yielding a CV 9.8%. The 21 dives actually came from 5 whales, so there is a potential issue there with non-independence of dives within whales (biologists call this issue "pseudoreplication"). But a more important issue is that we assumed the cue rate of these whales, tagged in 2006 and 2007, is representative of cue rates of an average whale during the 6-day period we wish to estimate density for in 2005, when there were no tagged animals. This assumption may be reasonable here, since beaked
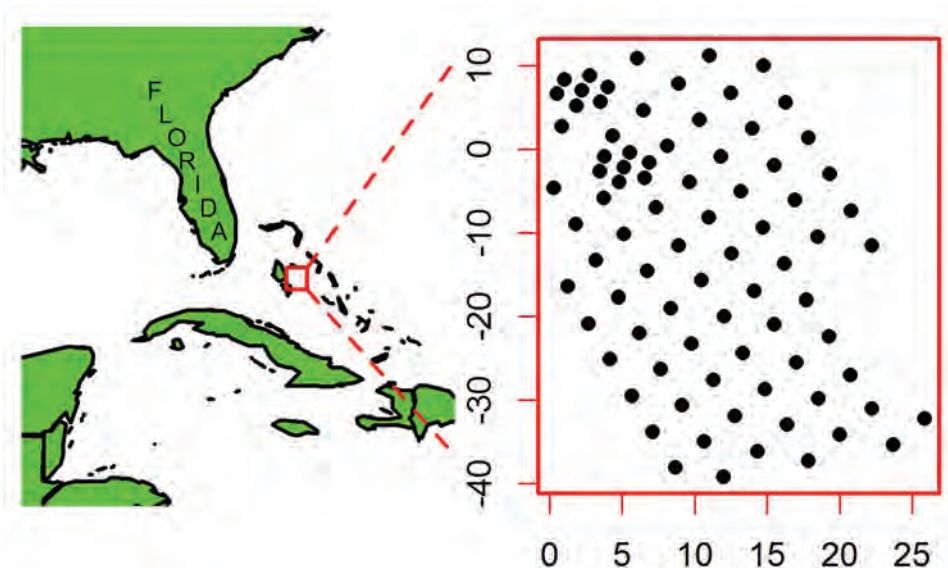

Fig. 1. Approximate location of the bottom-mounted hydrophones within the Atlantic Undersea Testing and Evaluation Center (AUTEC), Tongue of the Ocean, Bahamas.

whales show quite metronomic diving behavior (see discussion in Marques *et al.*, 2009 paper). However, for other species and other contexts it may not be, given that vocalization rates can vary in marine mammals as a function of time of day, year, group size, season, bottom depth, location, etc. If the relationship between cue rate and these factors is known, it may be possible to use a modeling approach to predict mean cue rate during the time of the survey. For example, if cue rate depends on local group size then one would need an estimate of the relationship between cue rate and group size (from a statistical model), and also the distribution of group sizes during the survey period of interest. One important limitation arises when cue rate depends on animal density itself—in this case the only option is to measure it within the study area during the time period of interest.

For estimating the detection probability, Marques *et al.* (2009) were able to match up clicks produced by the tagged whales with those received on the surrounding bottom-mounted hydrophones. The simplest analysis would have been to take all hydrophones within an 8 km radius of the whale and calculate what proportion of the emitted clicks was received. However, because there was a relatively small number of dives (only 15 with tracks), and they did not wish to assume that animals were randomly distributed with respect to hydrophones, Marques *et al.* took a more complex approach. This involved building a regression model of the probability of detecting a click on the bottom-mounted hydrophones as a function of direct-path distance and horizontal and vertical angle of the whale with respect to the hydrophones. The angles were important because the echolocation clicks of beaked whales are highly directional. Given the fitted regression model (see Fig. 2), a Monte-Carlo procedure was used to integrate out distance and angle, and yield an estimated average detection probability of $\hat{p} = 0.032$ with CV of 15.9% (again using the dive as the sample unit).

As with the cue rate multiplier, there is the issue that detection probability was not estimated at the same time and place as the density dataset. There is a particular potential for bias here because tagging studies can only take place under calm sea conditions which may not be representative of average conditions. Indeed, the average wind speed at a nearby weather station during the 6-day period of the survey was 12.4 kn, while the average speed when the tags were on was only 6.1 kn. Higher wind speed could reduce detectability, causing the estimate of detection probability using the tags to be too high. One option to address this would be to include wind speed in the detection probability regression model—however, predicting detectability at wind speeds observed in the 6-day period would then mean extrapolating outside the range of the data from the tag periods, so this approach is unlikely to be reliable. A second option is to analytically model the effect of increased noise on the detection algorithm. Instead, being empiricists, we chose a third option, described in a follow-up study (Ward *et al.*, 2011), that involves measuring the effect. Ward *et al.* (2011) added noise from sound samples taken at the study area under a variety of wind speeds to the original hydrophone recordings, re-ran the detection algorithm, and determined the reduction in performance caused by the additional noise. They found that detection probability was indeed lower at higher noise levels, although not by as much as one would predict from theory.

Putting all of the above together, Marques *et al.* (2009) came up with a density estimate of 25.2 animals per 1000 km², with a CV of 19.5% and corresponding confidence interval of 17.3-36.9. The estimate contains 4 random components: n (number of clicks), $\hat{c}$ (false positive rate), $\hat{p}$ (detection probability) and $\hat{r}$ (cue rate). The percentage contribution to the overall variance of these was, respectively, 8, 1, 66 and 25%. Hence, to reduce uncertainty, the thing to focus on in this study would be detection probability and, to a lesser
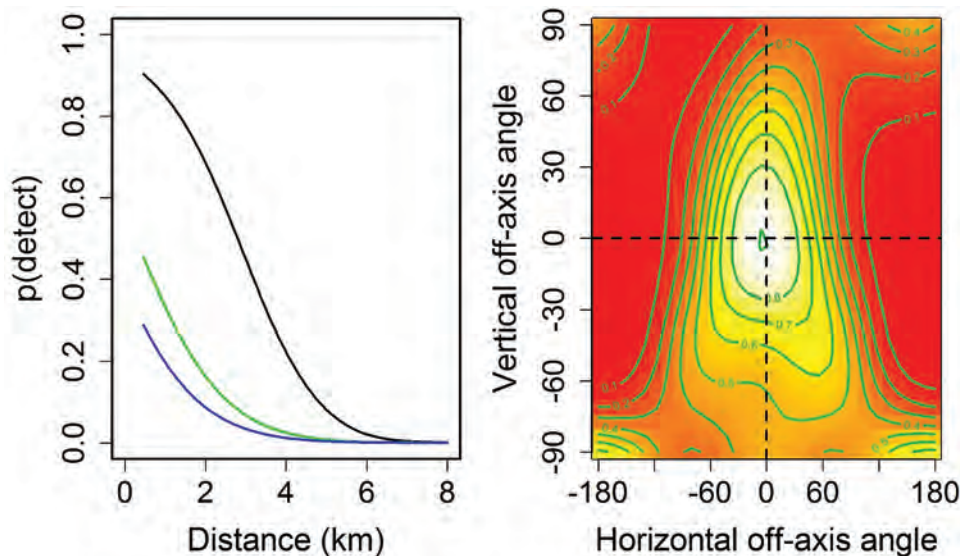

Fig. 2. Estimated probability of detecting a Blainville's beaked whale echolocation click on bottom-mounted hydrophones at the Atlantic Undersea Testing and Evaluation Center (AUTEC), derived from a regression analysis on tagged whales at known locations and orientations relative to the hydrophones. The left plot shows probability of detection as a function of distance for clicks where the animal is pointed directly at the hydrophones (black line) and at off-axis angles of 45 and 70 degrees (green and blue lines). The right plot is a heatmap showing probability of detection as a function of vertical and horizontal off-axis angles, evaluated at the smallest observed distance (0.46km).

extent, cue rate—implying tagging more whales. As we have discussed, performing the tagging study during the main survey would also be better.

Examples of other ways to estimate detectability

We presented our first example in some detail, to show the complexities that can be involved and the potential issues of which to beware, particularly with the multipliers. However, cue counting of echolocation clicks is not the only approach for obtaining a count, and using tagging data is not the only way to estimate detectability. Here, we illustrate the diversity of approaches with a few more examples. We start with methods that rely on having good acoustic data, in the sense of many hydrophones closely spaced and in some cases sophisticated acoustic processing. We end up with approaches that are more applicable in situations where data are sparser. In these latter cases, inferences will be correspondingly less reliable.

Total counts—beaked and sperm whales in the Bahamas

These two examples illustrate the case where we can assume that all objects within some defined spatial boundary are detected, and all outside that boundary can be excluded. Hence, we do not need to worry about false negatives. In both of these examples, the false positive rate is also negligible. Both examples are based on data from the 82-hydrophone bottom-mounted array at Tongue of the Ocean, Bahamas.

The first (Moretti *et al.*, 2010) used a cue-counting approach applied to Blainville's beaked whale, but instead of the cue being an echolocation click, it was the start of the vocal part of a foraging dive. Groups of animals dive together, and at some point during their descent they begin echolocating. The clicks produced can be detected relatively easily on the hydrophone array, and an approximate location determined either by eye or through a simple smoothing algo-

rithm (see Fig. 3). The hydrophone array is dense enough that no diving groups can be missed; we also require that diving groups can be separated with accuracy, and that the localization is good enough to determine whether the group is diving within the area monitored or outside. Given these assumptions, density can be estimated using

$$\hat{D} = \frac{n}{a} \times \frac{\hat{g}}{\hat{d}T} \qquad (3)$$

where n is the number of dive starts recorded within the study area a during time T, and the two other multipliers required to convert density of dives to density of animals are $\hat{g}$, the average group size (obtained from visual surveys) and $\hat{d}$, the average dive rate, obtained from DTAG data. The authors applied both this approach and the click-based approach to data from a 3-day period during navy operations. They found that the density of animals (undertaking foraging dives) decreased substantially during operations, compared with before and after them, but that's another story (see also McCarthy *et al.*, 2011). More relevant here, they found that both methods gave similar results, but that the dive counting method was more precise (CV of 11.9% for dive counting vs. 21.4% for click counting). This was because there were only two random quantities contributing to the variance—the estimates of group size and dive rate. Since the number of dive starts was assumed to be a complete count of everything in the surveyed area, and since the surveyed area covered the whole of the area we wish to estimate density for, then no randomness came from the n. The lesson here is that if it is possible to do a complete count within the area of interest, this is probably better, although it depends on how precisely the various required multipliers can be estimated.

The second example involves density estimation of sperm whales (*Physeter macrocephalus*) from a 42-day period in 2007 (Ward *et al.*, in press 2012). Sperm whales are also
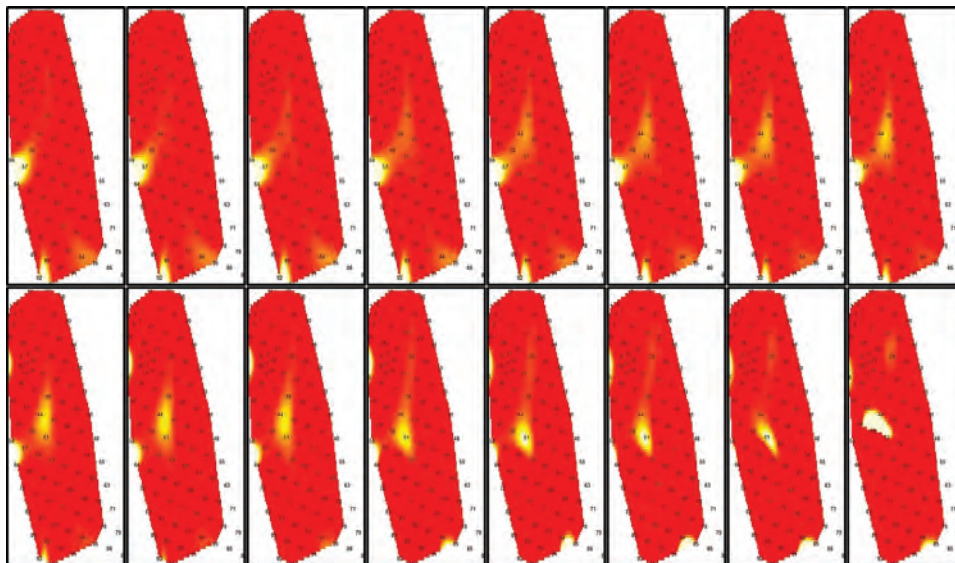

Fig. 3. A set of heatmaps showing the output of a smoothing algorithm applied to the number of Blainville's beaked whale clicks detected at each of 82 hydrophones in the Tongue of the Ocean, Bahamas (cf. Fig. 1 for range location). White indicates the most clicks, then yellow, then red. Each plot includes 10 minutes of data, and successive plots advance by 1 minute from left to right, top to bottom. There appears to be one group diving at the beginning of the time series, at the center left (maybe a second group is present during the first 10 minutes, at the bottom, likely outside the range). A new dive starts (the "objects" counted in the dive counting method) to appear around minute 4, being clearly visible from around minutes 9 onwards.

deep-diving animals that use echolocation to find prey–but compared with beaked whales their echolocation clicks are much louder, so that "trains" of clicks from the same individual can often be detected on multiple hydrophones from the AUTEC range. This means it is possible, with some effort, to localize individuals, or at least groups of individuals (see Baggenstoss, 2011, for details), and to estimate the number of individuals within a diving group as the number of overlapping (direct path) click trains.

Ward *et al.* (in press 2012) began by quickly screening the data to exclude long stretches of time when sperm whales were not present. For the remaining times, a systematic random sample of 50 10-minute periods were taken and subjected to more detailed analysis, where the number of animals diving and clicking within the study area was estimated. They assumed that all clicking animals in the sample periods were counted without error. Given this, the density estimator can be written

$$\hat{D} = \frac{n}{a} \times \frac{(1 - p_{np})}{\hat{p}_v k}$$
(4)

where n is the total number of animals counted within the study area a during the $k=50$ 10-minute periods, p_{np} is the proportion of the 42-day period for which sperm whales were not present in the quick screening, and $\hat{p}_v$ is the average probability of a whale vocalizing in a 10-minute period. This last multiplier was estimated from independent DTAG data—not from the same time or place, and so has similar limitations to those we discussed before.

Distance sampling—North Pacific right whales in the Bering Sea

It is quite rare to have a dense enough network of sensors over the area we wish to monitor that we can be certain to detect all animals vocalizing (or related cues such as group dives or vocalizations). Hence, it is almost always necessary to include an estimate of detection probability as a multiplier. The question then becomes how to estimate this probability. The answer depends on what data can be gathered about the detections. If it is possible to estimate the horizontal distance from a sensor to each detection then the distance sampling approach mentioned earlier may be possible. This relies on us assuming that the sensors are distributed randomly with respect to the animals—in this case the pattern of detected distances tells us about the change in detectability with distance. If we further assume that all animals at zero distance can be detected then we can estimate absolute detectability, and hence the average detection probability (See Fig. 4).

An example of this approach is Marques *et al.* (2011), who obtained preliminary density estimates for North Pacific right whales *Eubalaena japonica* in the Bering sea from their calls—i.e., using a cue-count approach. This species is considered one of the world's most endangered, having been subject to catastrophic whaling in the past, and estimates from intensive visual and genetic surveys put the population size in the tens.

In their paper, Marques *et al.* (2011) took advantage of special propagation conditions caused by the shallow water and amenable substrate. That meant that distances to detected calls could be obtained by analyzing the received calls and comparing them to a model of modal dispersal in a shallow water waveguide (see Munger *et al.*, 2011, for details). They used standard distance sampling methods to obtain an estimate of the probability of detecting a call, p, and then calculated density using Eq. 2. Other multipliers required were the false positive rate, in this case assumed to be zero, and the call rate, which was obtained from a separate survey that had been undertaken where groups of whales of known size were followed and call rates measured.

A significant limitation of this work was that only three sampling locations were available, straining the assumption that hydrophones were located at random with respect to the animals and also making extrapolation from local density in the region of the hydrophones to density in some larger area of interest merely speculation. Despite this, when they multiplied the density estimate by the area of Bering Sea shelf thought to contain right whales at the times of year surveyed, they obtained estimates comparable with the much more expensive ship-based surveys (25 whales, with 95% CI 13-47). This suggests that, if the sampling was expanded to a larger number of randomly-selected sites (20-30 ideally) then reliable inferences could be made for this important species.

Spatially explicit capture recapture (SECR)—minke whales off Hawaii

Another method of estimating detection probability is possible if animal vocalizations are detected on multiple hydrophones, and if the same call detected on multiple hydrophones can be accurately matched. If the hydrophones are close enough together and vocalizations frequent enough that animals can be localized and tracked, then the complete count methods described previously can be used. But in many cases only occasional calls are heard, and perhaps only on 1, 2, or 3 hydrophones, so that localization is generally not possible. Nevertheless, the pattern of detection and nondetection on the hydrophones gives us an indication both of where the sound comes from, but also the probability of it being detected (see Fig. 5). The method makes use of this information is called spatially explicit capture recapture (SECR).

Martin *et al.* (in press 2012) used this approach to estimate the call density for minke whales using 12 days of data from 14 bottom-mounted hydrophones located at another Navy testing range, the Pacific Missile Range Facility (PMRF) off Kauai, Hawaii. Minke whales are known to occur seasonally in this area, but are extremely visually cryptic; on the other hand their calls, called "boings," are readily detected. Matching ("associating") boings across hydrophones was done semi-manually, using timing and frequency information, on a subset of the data, and the association information was used in an SECR analysis to estimate the detection probability multiplier. This was then used to estimate a boing density for the whole dataset. Unfortunately, a reliable estimate of boing rate (i.e., the cue rate multiplier) was not available, so once again only a preliminary estimate of animal density
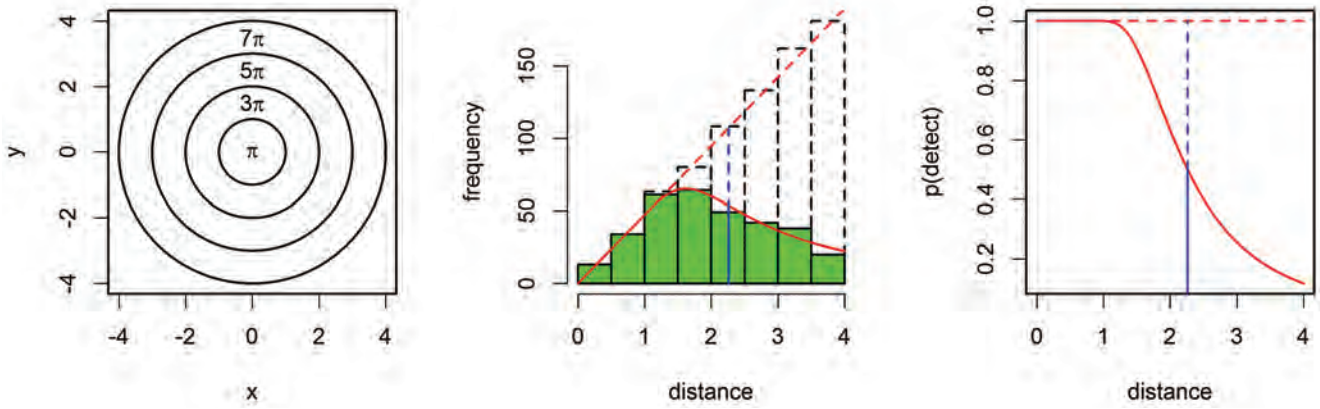
Fig. 4. Illustration of how the distances to detected objects can be used to estimate probability of detection for the case of randomly located point sensors. The area surveyed by bands of fixed width goes up linearly with increasing distance from the point, as shown in the left plot. Therefore, if we detected everything we would expect, on average, the number of detections to go up linearly with distance, as shown by the dashed histogram in the center plot. In practice, the number of detections drops off with distance, as shown by the green histogram. By fitting a curve to this (red solid line) and comparing it with a projection from the origin (red dashed line), we can estimate detectability: for example, at a distance of 2.2 we observed approximately half as many detections (blue solid line) as we expect if everything was seen (blue dashed line), so detection probability is estimated to be 0.5. Detection probabilities for all distances are shown in the right plot (red solid line), which is called a "detection function" in the distance sampling survey literature.
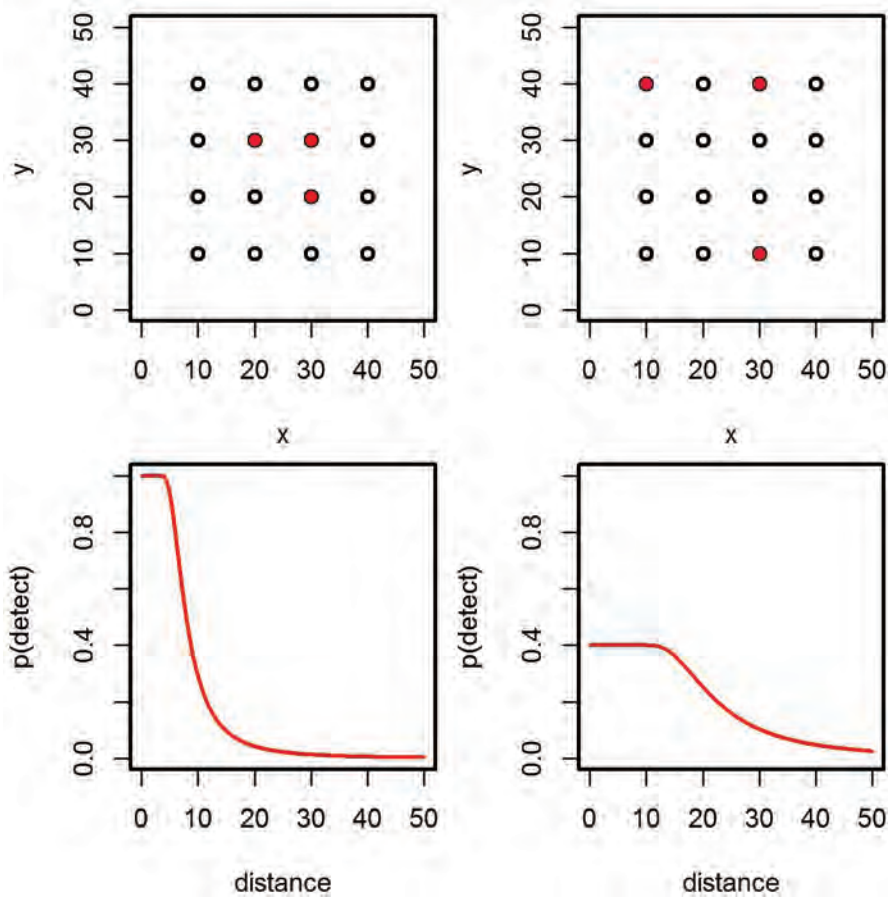

Fig. 5. Illustration of how the spatial pattern of a detected sound (red circles) on a hypothetical array of 16 hydrophones (black circles) contains information about the probability of detection. In the left case, the detection probability is high for small distances, but drops off rapidly. Hence each sound is only heard on a compact set of hydrophones, but tends to be heard on all of them. In the right case, detection probability is low at small distances, but drops off gradually. Here, each sound is heard over widely spaced hydrophones, but not at all of them. This kind of thinking is the basis for spatially explicit capture recapture (SECR) methods.

could be obtained.

The standard SECR method can be extended to use additional information. For example, the time of arrival of each sound at the hydrophones contains information about the sound's location, even when it is only heard on two hydrophones. Another example is that some hydrophones are directional, and so give information about the bearing to the sound source. Members of our research group at St. Andrews are extending the basic method so that it can use this additional information, and we have found in preliminary work that it leads to substantial improvements in precision, especially when sounds are usually heard only on few sensors.

Trials with known-location sounds—harbor porpoise in Denmark

Yet another method of obtaining the detection probability multiplier is possible when you have a set of animals at known locations relative to the hydrophones. You can then record, for each of a set of time periods, whether each animal was detected by each hydrophone, and use these data to construct a regression of probability of detection against distance (and perhaps other factors). This is the approach taken in our first example, above, where the beaked whales were tagged in the vicinity of the AUTEC hydrophones, and hence their location each time they clicked could be derived, as well as whether each of these clicks were detected on the surrounding hydrophones. Another example of this kind of method is the study of Kyhn *et al.* (2012), who were interested in determining the feasibility of estimating the density of harbor porpoise (*Phocoena phocoena*) using commercially-available autonomous porpoise detectors called T-PODs.[6] Kyhn *et al.* (2012) set up a visual monitoring station on cliffs at Fyns Hoved, Denmark, and tracked passing porpoises. They moored T-PODs below the cliffs, and by splitting each track up into a set of short segments, could determine for each segment whether the T-POD registered or not the presence of the porpoise.

Acoustic modeling—beaked whales in the Bahamas, again

The above methods all use additional data, hopefully collected at the time of the survey, to estimate the detection probability multiplier. However, in some cases such data are not available—all we know is how many sounds were detected on each sensor. In this case, it may still be possible to estimate detectability, if we are willing to make some strong assumptions about sound production, propagation, and possibly detection. This involves application of the passive sonar equation, which can be written

$$SNR = SL - DL - TL - NL \qquad (5)$$

where *SNR* is the signal to noise ratio at the receiver, *SL* is source level, *DL* is the directivity loss (i.e., attenuation in the source level due to the direction the animal is pointing relative to the hydrophone when they make the sound), *TL* is transmission loss as a function of the distance between animal and receiver and *NL* is the level of ambient noise. Hence, given a set of assumptions about the distribution these quantities, one can predict the distribution of SNR for a sound produced at a given location. SNR can be related to detection probability either by assuming that sounds above a certain SNR are certain to be detected, or (better) using an empirically-derived relationship between probability of detection and SNR.

An example of this is given by Küsel *et al.* (2011), who derived a density estimate for Blainville's beaked whale based on data from a single hydrophone at the AUTEC range. They obtained distributions of *SL, DL,* and *NL* from the published literature, and used several models of sound propagation to calculate *TL.* The relationship between detection and *SNR*

was estimated using a short sample of data from the study period that was hand-annotated and then run through the detector. From these pieces, average detection probability was derived using a Monte Carlo procedure, sampling from assumed distributions of animal position and orientation, SL, DL, NL and estimated values of the SNR-detection probability regression parameters. The resulting estimate, 0.014 (CV 17.9%) was not too dissimilar from that obtained by Marques *et al.* (2009) using the DTAG data (0.032 with CV 15.9%). Clearly, however, measurements will be preferable to assumptions whenever possible.

A half-way-house between this and the distance-based approach discussed above would be to use the received level of calls to estimate their distance, by using an assumed transmission loss model and source level distribution, or a regression relationship built using a sample of sounds where received level and actual distance are known. One example of estimating distance from received level is Širović *et al.* (2009), using blue whales in the Antarctic Ocean.

Conclusions
Those pesky multipliers

We have shown that it is possible to estimate animal density from passive acoustic data from fixed sensors, and have demonstrated a variety of approaches. The best method depends on the type of data available, and what you can most reliably count. If you can count all of the individuals within the study area and exclude all of those outside, then reliable estimation is straightforward. In most situations, however, this is not possible, and you then need multipliers to convert the count from calls or groups into animals, and to deal with false positive and false negative rates. We have demonstrated several methods for estimating detectability (the complement of false negative rate). The most reliable use data collected at the same time as the main survey, such as distances to detections, or associations in detections among sensors. A less-satisfactory option is to undertake a secondary survey, such as tagging some animals—again it is best if this is done at the same time and place as the main survey but this is often not possible. Deriving detection probability estimates from acoustic modeling alone is a last resort, but often constraints will mean that other approaches are not possible.

Often the object being counted is an individual vocalization: an echolocation click or a call. We then need an estimate of the vocalization rate. This is often the main impediment to obtaining a density estimate—either the rate is completely unknown, or it has been measured in a very different circumstance and could reasonably be expected to vary over time or space. Knowledge of the basic acoustic biology of our study species is often a fundamental limitation.

Other applications

We have focused here on fixed detectors, but there is plenty of potential for estimating density from towed acoustic sensors, behind ships or gliders for example. If the sensors are designed to get the bearing to detected calls, and animals vocalize repeatedly, then intersecting bearings can be

used to obtain a (left-right ambiguous) localization—this is then amenable to analysis using a different flavor of distance sampling called line transect analysis—see, e.g., Lewis *et al.* (2007).

There are many possible applications in the terrestrial environment, where animals are hard to see but easier to hear. Examples include forest elephants, gibbons and frogs.

Future directions

Currently, few surveys are designed with passive acoustic density estimation in mind. All of the above case studies made use of data originally gathered for other purposes. One example of a designed survey is the SAMBAH project[7] where around 300 autonomous detectors have been deployed on a random grid of locations throughout the Baltic Sea with the goal of yielding an abundance estimate of the extremely low-density population of harbor porpoise resident there. Fixed passive acoustic methods are potentially highly cost-effective compared with other methods for long-term monitoring programs, because costs are relatively low after initial set up. Hence we expect to see increasing uptake where long-term monitoring is required, such as over the lifetime of oil production fields.

We have shown that the methods work best when a large number of sensors can be distributed through the survey area, and where at least a subset collect auxiliary information such as distances to detected sounds. There is great potential for the development of inexpensive commercial hardware and associated algorithms and software to facilitate this.

There are plenty of statistical developments left to pursue, including methods for combining different sources of information to estimate detectability better, methods for sensors that use adaptive duty-cycling to increase their longevity and optimal survey design.

Lastly, improvements in our knowledge of the acoustic biology of animal species, coupled with advances in our ability to detect, classify and localize sounds, will make passive acoustic density estimation ever more feasible in a wider variety of situations.**AT**

Endnotes

1   For example, our research group at St. Andrews develops free software, Distance, which can be used to design and analyze distance sampling surveys. The program has acquired over 26,000 registered users covering all major taxa since the windows-based version was released 13 years ago—see http:// www.ruwpa.st-and.ac.uk/distance/

2   DECAF stands for Density Estimation for Cetaceans from Acoustic Fixed Sensors. The project ran from 2007–2011 and was funded under the National Oceanographic Partnership Program jointly by the Ocean Acoustics Program of the US National Marine Fisheries Service, Office of Protected Resources and the Joint Industry Program of the International Association of Oil and Gas Producers on Sound and Marine Life. In addition to the other team members, we also thank the project steering committee for their assistance: Jay Barlow (National Marine Fisheries Service), Stephen Buckland (University of St. Andrews) and Walter Zimmer (NATO Undersea Research Center). All project outputs are available at http:// www.creem.st-and.ac.uk/decaf/. Our work in this area is also supported by the US Navy, Chief of Naval Operations, Energy and Environmental Readiness Division (Code N45).

3   We omit all details of the acoustic processing techniques required to yield the raw materials for density estimation: the detected and classified (and in some cases associated and localized) vocalizations. Other articles in this issue describe some of the work in this area, as does an article from last year (Tiemann *et al.*, 2011). Recent reference texts are Zimmer (2011) and Au and Hastings (2009).

4   The CV on a quantity that is the product of a set of independent random quantities can be calculated using the "delta method," which just involves adding the squared CVs of each of the independent quantities—see e.g., Marques *et al.* (2009).

5   This is something of an over-simplification. DTAG data alone yields an imprecise "pseudotrack," which needs to be combined with positions estimated from, e.g., acoustic localization to give a useable track—see Ward *et al.* (2008) and references therein for details.

6   The newer versions are called C-PODs; they are manufactured by Chelonia Limited (www.chelonia.co.uk).

7   Static Acoustic Monitoring of Baltic Sea Harbour Porpoise—see www.sambah.org.

References

Au, W. W. L., and Hastings, M. C. (**2009**). *Principles of Marine Bioacoustics* (Springer, New York).

Baggenstoss, P. M. (**2011**). "An algorithm for the localization of multiple interfering sperm whales using multi-sensor time difference of arrival," J. Acoust. Soc. Am. **130**, 102–112.

Buckland, S. T., Anderson, D. R., Burnham, K. P., Laake, J. L., Borchers, D. L., and Thomas, L. (**2001**). *Introduction to Distance sampling–Estimating Abundance of Biological Populations* (Oxford University Press, Oxford).

Buckland, S. T., Burt, M. L., Rexstad, E. A., Mellor, M., Williams, A. E.. and Woodward, R. (**2012**). "Aerial surveys of seabirds: The advent of digital methods," J. Appl. Ecol. **49**, 960–967.

D'Amico, A., Gisiner, R. C., Ketten, D. R., Hammock, J. A., Johnson, C., Tyack P. L., and Mead, J. (**2009**). "Beaked whale strandings and naval exercises," Aquatic Mammals **35**, 452–472.

Küsel, E. T., Mellinger, D. K., Thomas, L., Marques, T. A., Moretti, D., and Ward, J. (**2011**). "Cetacean population density estimation from single fixed sensors using passive acoustics," J. Acoust. Soc. Am. **129**, 3610–3622.

Kyhn, L., Tougaard, J., Thomas, L., Duve, L., Steinback, J., Amundin, M., Desportes, G., and Teilmann, J. (**2012**). "From echolocation clicks to animal density–Acoustic sampling of harbour porpoises with static dataloggers," J. Acoust. Soc. Am. **131**, 550–560.

Lewis, T., Gillespie, D., Lacey, C., Matthews, J., Danbolt, M., Leaper, R., McLanaghan, R., and Moscrop, A. (**2007**). "Sperm whale abundance estimates from acoustic surveys of the Ionian Sea and Straits of Sicily in 2003," J. Marine Biol. Assn. of the UK **87**, 353–357.

Marques, T. A., Munger, L., Thomas, L., Wiggins, S., and Hildebrand, J. A. (**2011**). "Estimating North Pacific right whale (*Eubalaena japonica*) density using passive acoustic cue counting," Endangered Species Res. **13**, 163–172.

Marques, T. A., Thomas, L., Ward, J., DiMarzio, N., and Tyack, P. L. (**2009**). "Estimating cetacean population density using fixed passive acoustic sensors: An example with Blainville's beaked whales," J. Acoust. Soc. Am. **125**, 1982–1994.

Martin, S. W., Marques, T. A., Thomas, L., Morrissey, R. P., Jarvis, S., DiMarzio, N., Moretti, D., and Mellinger, D. K. (**in press, 2012**). "Estimating minke whale (*Balaenoptera acutorostrata*) boing

sound density using passive acoustic sensors," Marine Mammal Sci. doi: 10.1111/j.1748-7692.2011.00561.x

McCarthy, E., Moretti, D., Thomas, L., DiMarzio, N., Morrissey, R., Jarvis, S., Ward, J., Izzi, A., and Dilley, A. (**2011**). "Changes in spatial and temporal distribution and vocal behavior of Blainville's beaked whales (*Mesoplodon densirostris*) during multiship exercises with mid-frequency sonar," Marine Mammal Sci. **27**, E206–E226.

Mellinger, D. K., Stafford, K. M., Moore, S. E., Dziak, R. P., and Matsumoto, H. (**2007**). "An overview of fixed passive acoustic observation methods for cetaceans," Oceanography **20**, 36–45.

Moretti, D., Marques, T., Thomas, L., DiMarzio, N., Dilley, A., Morrissey, R., McCarthy, E., Ward, J., and Jarvis, S. (**2010**). "A dive counting density estimation method for Blainville's beaked whale (*Mesoplodon densirostris*) using a bottom-mounted hydrophone field as applied to a mid-frequency active (MFA) sonar operation," Appl. Acoust. **71**, 1036–1042.

Munger, L. M., Wiggins, S. M., and Hildebrand, J. A. (**2011**). "Right whale "up-call" source levels and propagation distance on the southeastern Bering Sea shelf," J. Acoust. Soc. Am. **129**, 4047–4054.

Širović, A., Hildebrand, J. A., Wiggins, S. M., and Thiele, D. (**2009**). "Blue and fin whale acoustic presence around Antarctica during 2003 and 2004," Marine Mammal Sci. **25**, 125–136.

Tiemann, C. O., Jaffe, J. S., Roberts, P. L. D., Sidorovskaia, N. A., Ioup, G. E., Ioup, J. W., Ekimov, A., and Lehman, S. K. (**2011**). "Signal and image processing techniques as applied to animal bioacoustics problems," Acoust. Today **7**(3), 35–43.

Ward, J. A., Thomas, L., Jarvis, S., DiMarzio, N., Moretti, D., Marques, T. A., Dunn, C., Claridge, D., Hartvig, E., and Tyack, P. (**in press, 2012**). "Passive acoustic density estimation of sperm whales in the Tongue of the Ocean, Bahamas," Marine Mammal Sci. DOI: 10.1111/j.1748-7692.2011.00560.x

Ward, J., Jarvis, S., Moretti, D., Morrissey, R., DiMarzio, N., Thomas, L., and Marques, T. A. (**2011**). "Beaked whale (*Mesoplodon densirostris*) passive acoustic detection with increasing ambient noise," J. Acoust. Soc. Am. **129**, 662–669.

Ward, J., Morrissey, R., Moretti, D., DiMarzio, N., Jarvis, S., Johnson, M., Tyack, P., and White, C. (**2008**). "Passive acoustic detection and localization of *Mesoplodon densirostris* (Blainville's beaked whale) vocalizations using distributed bottom-mounted hydrophones in conjunction with a digital tag (DTag) recording," Can. Acous. **36**, 60–66.

Williams, B. K., Nichols, J. D., and Conroy, M. J. (**2002**). *Analysis and Management of Animal Populations* (Academic Press, San Diego, CA).

Zimmer, W. M. X. (**2011**). *Passive Acoustic Monitoring of Cetaceans* (Cambridge University Press, Cambridge, UK).

Tiago A. Marques is a Research Fellow at the University of St. Andrews, from where he obtained his Ph.D. in Statistics in 2007. His first degree was in Biology at the Universidade de Lisboa, and he is still affiliated to the university as a member of CEAUL, Centro de Estatística e Aplicações da Universidade de Lisboa. His research focuses on statistical methods applied to estimating animal abundance, in particular the use and development of distance sampling. Through the Density Estimation for Cetaceans from Acoustic Fixed Sensors (DECAF) project he got involved in developing and applying methods to estimate animal abundance from passive acoustic data, and lately has started to think about methods to model cetacean movement.

Len Thomas is a Reader in Statistics at the University of St. Andrews. He holds degrees in biology (BSc, Sheffield, UK, 1990), biological computation (MSc, York, UK, 1991) and forest science (Ph.D., University of British Colombia, Canada, 1997). He is interested in many aspects of statistical ecology including methods for estimating animal abundance, modeling population dynamics and linking anthropogenic disturbance of animals to population outcomes. His involvement in passive acoustic density estimation began in 2005 at a meeting about monitoring marine mammals on Navy testing ranges, when Stephen Martin (Space and Naval Warfare Systems Command, SPAWAR) and John Hildebrand (Scripps Institution of Oceanography) suggested it would be an interesting topic to look into. They were right.