Available online at www.sciencedirect.com
ScienceDirect
Applied Acoustics 67 (2006) 1074–1090
applied acoustics
www.elsevier.com/locate/apacoust

Localisation of sperm whales using bottom-mounted sensors

P.R. White *, T.G. Leighton, D.C. Finfer, C. Powles, O.N. Baumann

Institute of Sound and Vibration Research, University of Southampton, Highfield, Southampton Hants SO17 1BJ, UK

Received 20 March 2006; received in revised form 10 May 2006; accepted 21 May 2006
Available online 17 July 2006

Abstract

This paper describes the results of applying acoustic localisation methods to data from a free ranging sperm whale (*Physeter macrocephalus*) recorded using an array of five bottom-mounted hydrophones. The data analysed were supplied as part of the 2nd International Workshop on Detection and Localisation of Marine Mammals, Monaco, 2005. The localisation technique employed is based on a weighted least squares optimisation procedure applied to estimated delays. The method by which these delays are estimated in a robust manner is detailed. Three analytic acoustic models for the environment are explored and the effects of noise and model mismatch are considered.

Keywords: Passive acoustics; Bioacoustics; Localisation

1. Introduction

Passive acoustics provides a powerful non-invasive tool with which to localise vocalising marine mammals in the oceanic environment. The ability to track such mammals has clear benefits to marine biologists seeking to study the behaviour of these animals but also has growing application as a monitoring tool to help control the exposure of marine mammals to anthropogenic noise. Localisation of an acoustic source can be regarded as a par-

* Corresponding author. Tel.: +44 23 80592274; fax: +44 23 80593190.
E-mail addresses: prw@isvr.soton.ac.uk (P.R. White), tgl@isvr.soton.ac.uk (T.G. Leighton), dcf@isvr.soton.ac.uk (D.C. Finfer).

0003-682X/$ - see front matter
doi:10.1016/j.apacoust.2006.05.002

ticular example of an acoustic inversion problem, and has been studied in the military context for some considerable time [1,2]. There is a growing trend to adapt these principles to the localisation and tracking of marine mammals [3–6]. Particular attention has been paid to the localisation of odondocetes, with the majority of work concentrating on the localisation of sperm whales (*Phyester macrocephalus*), a problem to which this paper is addressed. The localisation of odontocetes benefits from the fact that their echo-location clicks have, through the evolution of the species, become signals which, being broad-band and of short duration [7–10], possess good range resolution [11]. Odontocetes exploit this range resolution to allow them to accurately locate objects, such as prey.

Source localisation can be regarded as inverting an acoustic model, i.e. identifying the location of a source, usually in 3-D space, based on a series of observed delays. The forward acoustic model on which such an inversion is based can be expressed as

$$\mathbf{\Delta}(n) = \mathscr{M}(\mathbf{s}(n); \mathbf{\Theta}) + \mathbf{w}(n), \tag{1}$$

where $\mathscr{M}(\mathbf{s}(n); \mathbf{\Theta})$ is a complete model of the acoustic environment relating the source location, $\mathbf{s}(n)$, through the parameters, $\mathbf{\Theta}$, to the time delays observed between sensor pairs and n represents a discrete time index. The observed delays $\mathbf{\Delta}(n)$ are corrupted by the additive noise, $\mathbf{w}(n)$. For this discussion the model, $\mathscr{M}$, is an abstract concept because we are assuming that it completely captures the acoustic properties of the environment and, as such, requires an unrealistic level of knowledge of physical processes and parameters. We regard practical models used in source location to be simplifications of $\mathscr{M}(\mathbf{s}(n); \mathbf{\Theta})$ in which the parameter space is constrained, i.e. $\mathbf{\Theta}$ is restricted to a manifold within the parameter space. As an illustration of this concept, one can regard an isospeed model as a special case of a linear sound speed profile model in which the gradient of the sound speed profile is constrained to be zero. In turn, a linear sound speed model may itself be regarded as a simplification of a general range independent (or stratified) model. This can be extended to construct a hierarchy of models, the most detailed and complete of which we are denoting $\mathscr{M}$.

The source localisation problem is solved by inverting the model, computing

$$\hat{\mathbf{s}}(n) = \mathscr{M}^{-1}(\mathbf{\Delta}; \mathbf{\Theta}_0), \tag{2}$$

where $\hat{\mathbf{s}}(n)$ is the estimated source location, $\mathbf{\Theta}_0$ is the parameter space constrained to lie on a manifold corresponding to one's chosen simplification of the ideal model. The inverse operator, $\mathscr{M}^{-1}$, is defined so that, the absence of noise and perfectly matched parameters, the estimated source location will coincide with the actual source location, i.e. $\hat{\mathbf{s}}(n) = \mathbf{s}(n)$. The reason for adopting the conceptual model (1) is to provide a common framework within which to discuss the influence of two sources of imprecision (error) that occur when estimating source location (or indeed conducting any form of inversion). The first errors arise as a consequence of the inevitable introduction of the noise, $\mathbf{w}(n)$, during the measurement of delays. The second source of error comes as a result of model mismatch; caused by the fact that $\mathbf{\Theta}_0 \neq \mathbf{\Theta}$. It is unfeasible to measure the oceanic parameters with sufficient spatial and temporal resolution in order to completely model its acoustic properties, so inevitably one adopts a simplified model.

There is a clear route for determining the influence of noise on the estimated source locations, neglecting the effects of model mismatch [12]. This requires one to make assumptions regarding the statistical distribution of the errors and normally entails making a local linearisation of the model. The effects of model mismatch are generally less

easily quantified. For a given model, the effect of small errors in a single model parameter, θ, can be approximated using:

$$\delta\hat{\mathbf{s}}(n) = \frac{\partial \mathscr{M}^{-1}(\mathbf{\Delta}; \mathbf{\Theta}_0)}{\partial \theta} \delta\theta, \tag{3}$$

where $\delta\hat{\mathbf{s}}(n)$ is the change in the estimated source location, $\delta\theta$ is the perturbation to the model parameter, which represents one form of model mismatch, and $\partial \mathscr{M}^{-1}(\mathbf{\Delta}; \mathbf{\Theta}_0)/\partial\theta$ is the model sensitivity. Typically, the sensitivity depends on the true source location, sensor location, as well as the acoustic model. As an example consider a planar array of sensors arranged in a rectangular geometry, and a stratified propagation model. If the source is located at the centre of this array, then the sensitivity of the algorithm, in the plane of the sensors, is zero. This is because there are no differences in the propagation times from the source to each of the sensors, regardless of the acoustic model used, or any of its parameters. More broadly, it is true that models are less sensitive when the source is located toward the centre of the array. Part of the goal of this paper is to assess the relative importance of model mismatch and noise for the given data set.

2. Methods

Most source localisation algorithms can be regarded as consisting of two distinct processes. The first element is the computation of the delays between sensor pairs, $\mathbf{\Delta}$, and the second aspect is the inversion of these delay estimates in order to infer the source location, $\mathbf{s}(n)$. The following subsections detail the approaches used here to solve these two tasks.

2.1. Time delay estimation

The estimation of the time delay between two sensors is most commonly accomplished through the use of some form of cross-correlation based analysis [13]. There are two main reasons why direct application of cross-correlation to raw acoustic data is prone to poor performance in this application. Firstly, the fact that a sequence of sperm whale clicks has a regular, although not exactly periodic, structure results in a correlation function that has multiple peaks. Robustly determining which of these peaks corresponds to the true delay can represent a challenge. Secondly, the directional character of the sperm whale echolocation clicks [7–10] means that recordings on sensors at different orientations to the animal receive waveforms that maybe quite different in structure, so that the signals on the two sensors may be relatively poorly correlated.

To avoid the second of the problems, it is natural to perform cross-correlation on some form of signal envelope, as opposed to using the raw acoustic data. The processing scheme adopted here employs an optimal detection statistic [14] for detecting transients. This optimal detection statistic represents the energy of the signal after the background noise has been whitened. The detection statistic is smoothed using a 1024 Hanning window, corresponding to a support of 21 ms at 48 kHz. The choice of 21 ms reflects the typical duration of a sperm whale click. An example of this detection statistic computed for a 10 s section of data containing a sequence of sperm whale clicks is shown in Fig. 1.

Cross-correlation of this detection statistic between sensor pairs results in a correlation function which still exhibits multiple peaks and hence may generate erroneous delay estimates. These spurious peaks arise, not only because the clicks occur at near-regular inter-

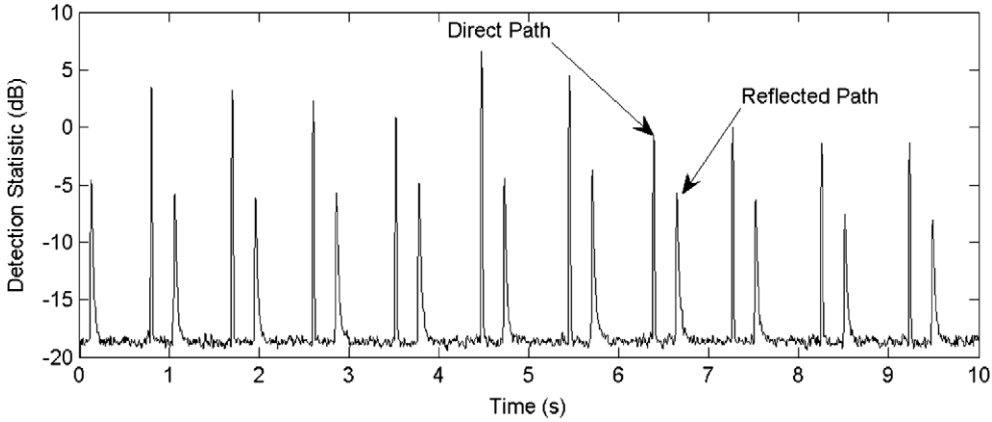

Fig. 1. Example of detection statistic for 10 s of data.

vals, but also because each click is commonly recorded via more than one propagation path. In this particular data set a single secondary arrival from the sea surface is observed throughout the recordings. The fact that the delay between the direct arrivals and the multi-path signal is fixed, at least over modest time-scales, results in additional large peaks appearing in the correlation function. It is these multi-path peaks that are most commonly responsible for generating spurious delay estimates.

In order to generate robust estimates of delay we avoid relying completely on correlation based techniques. Instead we first detect each click on a channel and seek to discriminate between direct path and reflected arrivals. Since the multi-path arrivals pass through the surface layer and are reflected from the sea–air interface, they are subject to significant surface reverberation. This causes an elongation of the measured detection statistic (which is evident in Fig. 1). Consequently, the rates of decay of the detection statistics from direct and reflected paths are significantly different. A relatively crude clustering algorithm allows one to distinguish the majority of reflected signals from the direct arrivals. There is no demand for highly accurate discrimination; the subsequent delay estimation algorithm performs well as long as the majority of events surviving discrimination correspond to direct arrivals.

The processing proceeds by computing the time difference between all pairs of detected direct path clicks on the two data channels, within a time window (30 s being used here). A histogram of these time differences is then computed and the peak value in the histogram is taken to represent the correct delay. The final time delay estimate is computed by taking the mean of the time differences falling within the histogram bin corresponding to the peak. If the pruning of the reflected path signals is only partially successful, then a secondary peak in this histogram will appear, although it will be smaller than the direct path peak. An example of one such histogram, using a bin size of 8 ms, is shown in Fig. 2.

2.2. Source location

As discussed in Section 1, the problem of source location can be regarded as an example of acoustic inversion, and as such requires one to assume a forward model on which to base the inversion. There are various levels of complexity for the forward modelling that

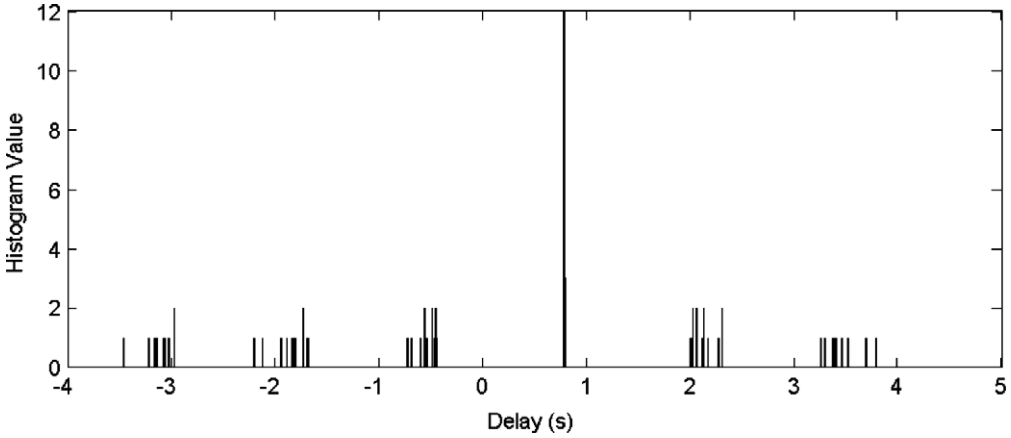

Fig. 2. Delay histogram.

have previously been considered [15–17]. These correspond to different choices for a manifold in $\boldsymbol{\Theta}$-space. In general, as the complexity of the forward model increases, the computational demands of the inversion method also increase. We consider three distinct models for which analytical expressions for the time delays can be obtained, so avoiding the need to employ numerical propagation models. These methods all generate results based on exactly the same delay estimates, obtained by the method described in Section 2.1. Hence by studying the different behaviours of the three algorithms, some insight on the influence of modelling errors can be obtained for this real data set. Although in the absence of ground truth with regards to the animal's location, one cannot be confident as to which model achieves the best absolute performance.

The three models are based on: (1) an isospeed anechoic assumption, (2) an isospeed model that includes surface reflections and (3) an anechoic environment with a linear sound speed variation. In all three cases the propagation times can be expressed in closed analytic form. Consequently, all three models avoid the need to implement a numerical acoustic propagation model and hence they are all extremely efficient to implement.

For each model the same algorithm is used to perform the inversion. It is based on a weighted least squares scheme. Assuming that there is no model mismatch and that the noise on the delay estimates is Gaussian and uncorrelated between sensors, then the maximum likelihood estimator is obtained by the minimisation of the weighted least squares cost function

$$\Psi(\mathbf{s}(n)) = \sum_{k=2}^{N} \frac{\left(\Delta_{1k}(n) - \mathscr{M}_{1k}(\mathbf{s}(n); \boldsymbol{\Theta}_0)\right)^2}{\sigma_{1k}^2}, \tag{4}$$

where $\Delta_{1k}(n)$ is the measured delay between sensor 1 (which is assumed to be the reference sensor), and sensor k, $\mathscr{M}_{1k}$ is the corresponding modelled delay. The term σ_{1k}^2 represents the variance of the errors on the delay estimates $\Delta_{1k}(n)$. The variance of the noise on the delays has to be estimated from the data itself. This can be achieved by spectral analysis of $\Delta_{1k}(n)$. The low frequency components of $\Delta_{1k}(n)$ correspond of the changes in delay resulting from the motion of the animal, whilst the high frequency components relate to the noise. Assuming that the noise on the delay estimates is uncorrelated from sample

to sample, one can estimate the noise variance by computing the energy in the high frequency range. This can be implemented via a finite impulse response filter. Herein we use the energy present in the upper 85% of the frequency band to estimate the variance of the noise over the full band. The weighted least squares cost function in (3) serves to emphasise delay estimates which are contaminated by relatively low levels of noise whilst suppressing the influence of the noisy delay estimates. This use of such an optimisation based method also provides a flexible framework, allowing one to employ any number of sensors (above the minimum necessary for localisation [18]) and to incorporate information from any number of available sensor pairs.

The optimisation of the cost function (3) is achieved using a Nelder–Meade optimisation algorithm [19]. Such optimisation schemes are not immune to convergence to a local (rather than global) optimum, but the cost function in this example exhibits a relatively simply topology. There are various alternatives for overcoming the general problem of locating global rather than local minima [19–21]. Such methods generally involve an increase in the number of function evaluations required, and in practice require considerable effort to ensure that they converge to the correct minimum. The optimisation algorithm is run multiple times using different randomly selected initialisations (herein 10 points are used) centred on the previously estimated source location. The final estimate is taken as the source location which yields the lowest cost function. The cost function exhibits a local minimum at a depth below the sea floor, this is a partial alias and is a consequence of near planar geometry of the array. This local minimum can be automatically avoided by augmenting the cost function with a quadratic barrier penalty term [22].

The fact that minimising (4) can be shown to be a maximum likelihood estimator means that we can ascribe to it the well understood properties of such estimators [23]. Specifically, we know that the variance of the estimates induced by the presence of noise is approximately described by the Cramer–Rao lower bound, and that the performance is asymptotically optimal with regards to the presence of that noise. These properties do not include the errors introduced through model mismatch. As we shall see in this data set (and as is the case in most scenarios of interest), such mismatches are a significant source of error.

2.2.1. Computation of travel times

The source location algorithms require one to use a model that predicts the delay observed between sensors based on different propagation models. This section presents the details of the models used here. The goal is to predict the travel time from a source located at $\mathbf{s}(n)$ to a receiver located at $\mathbf{r}_k = (r_x, r_y, r_z)$. For an isospeed model, where the sound speed is assumed to be c, it is trivial to show that the travel time, τ_k, is

$$\tau_k = \|\mathbf{r}_k - \mathbf{s}(n)\|/c, \tag{5}$$

where $\|\,\|$ denotes the Euclidean norm. The position of a virtual receiver, associated with a surface reflection, is given by $(r_x, r_y, -r_z)$. The delays required by the acoustic model can be expressed as

$$\Delta_{1k} = \tau_1 - \tau_k. \tag{6}$$

The problem of obtaining analytic expressions for the travel times for a linear sound speed environment is significantly more involved. It is well known that the ray paths in a medium with linear sound speed profile are arcs of circles and further the radius of

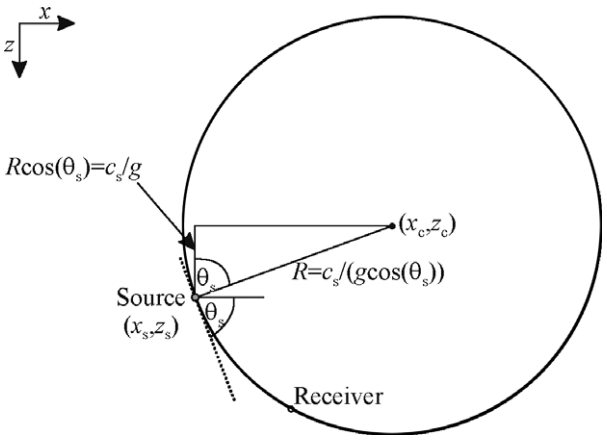

Fig. 3. Geometry for a source and receiver in a linear sound speed profile.

the circle can be computed [24]. The parameter g is the gradient of the sound speed profile, which has the form $c(z) = c_0 + gz$ and c_0 is the sound speed at $z = 0$. Fig. 3 illustrates the appropriate geometry. c_s is the sound speed at the source and θ_s is the launch angle of the ray at the source, measured relative to the horizontal. Note one seeks to determine the launch angle of the ray (θ_s) which will pass through the receiver located at (x_r, z_r). From the geometry shown in Fig. 3, the centre of the circle, (x_c, z_c), along which the ray path is an arc, can be shown to be

$$(x_c, z_c) = \left(\frac{x_s + x_r}{2} + \frac{(z_s - z_r)}{2(x_s - x_r)}(z_r - z_s + 2c_s/g), z_s - \frac{c_s}{g}\right). \tag{7}$$

The standard expression for the travel time of a ray travelling in a medium with a sound speed profile $c(z)$ is [24]

$$\tau = \int_{z_0}^{z_1} \left(c(z)\sqrt{1 - \frac{c(z)^2 \cos(\theta_s)^2}{c_s^2}} \right)^{-1} \mathrm{d}z. \tag{8}$$

For a linear sound speed profile this integral can be evaluated to yield

$$\tau = \frac{1}{g}\left\{ \log\left(\frac{z_c - z_s}{z_c - z_r}\right) - \log\left(\frac{R + x_c - x_s}{R + x_c - x_r}\right) \right\}. \tag{9}$$

Using Eqs. (7)–(9) allows one to compute the propagation time from the source to any receiver and hence allows one to compute the predicted delays.

3. Results

The methods discussed in Section 2 are applied to recordings supplied as data set 2 for the 2nd International Workshop on Detection and Localisation of Marine Mammals, Monaco, 2005. These data consist of the vocalisations from a single sperm whale recorded at the AUTEC range located at the Tongue of the Ocean, close to Andros Island in the Bahamas. Recordings from five bottom-mounted hydrophones, placed at depths ranging

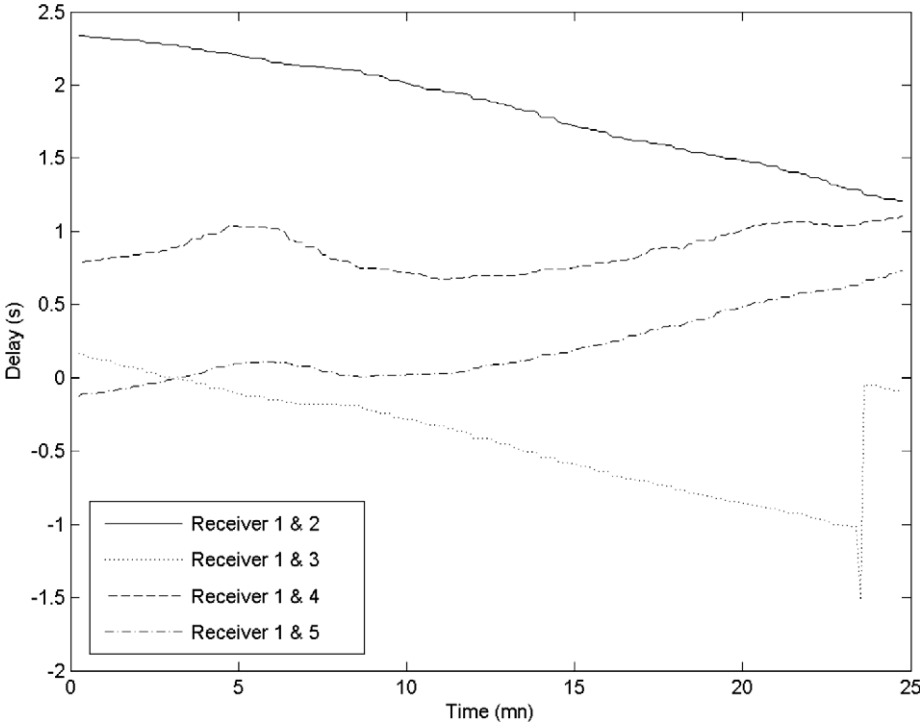

Fig. 4. Delay estimates using channel 1 as a reference.

from 1361 to 1556 m. The recordings were made at close to 11.00, local time, on 30th March 2002. Each recording encompasses a total period of 25 min for each channel, and for each channel these data were provided in the form of five 5 min files sampled at 48 kHz.

The estimates of source location are all constructed based on a common set of delay estimates, obtained using the methods described in Section 2.1. The delay estimates are computed based on 30 s blocks of data. These blocks are overlapped by 75%, so that a new delay estimate is computed for each 7.5 s of data. Thus for each 5 min recording there are a total of 37 delay estimates, resulting in a total of 185 estimates for each channel of data. The estimated delays[1] for each channel relative to channel 1 are plotted in Fig. 4. It is clear that in the majority of cases the delay estimates are consistent, in that the delays evolve in a smooth fashion. Towards the end of the recording the delay estimates between channels 1 and 3 are subject to large errors.[2] There are a total of 11 samples from this sequence that are clearly outliers. Analysis of the data during the portion of the recordings which gives rise to the outliers indicates that it is the low signal to noise ratio (SNR) of channel 3 during this period of the dataset which causes these problems. Analysis of the localisation results presented later shows that the outliers occur when the animal is at

[1] A correction of 2.3395 s to account for synchronisation errors in channels 2 and 3 has been applied.

[2] Such outliers can be readily removed from the final track estimate, either at the delay estimation stage or after the source location has been estimated.

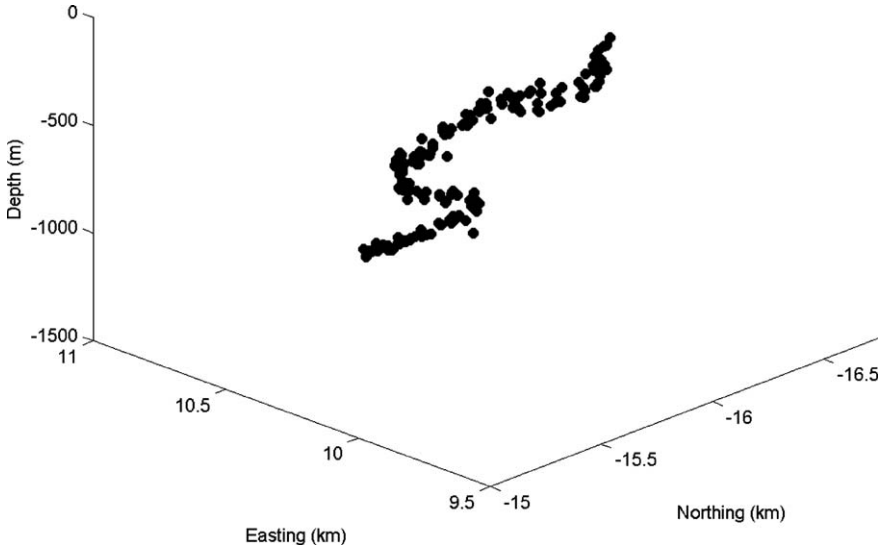

Fig. 5. Three dimension view of estimated locations.

its furthest from sensor 3 at the end of the recording, consistent with the low SNR observed in the data.

The three localisation schemes based on the three propagation models have been implemented. For the linear sound speed model we assume a gradient for the profile, g, of $0.017\ s^{-1}$, the value which reflects a constant temperature, constant salinity medium, in which the only factor affecting the sound speed is hydrostatic pressure. Such characteristics certainly are frequently assumed for the layer in the ocean commonly called the deep isothermal layer [11,24]. Sperm whales are commonly reported vocalising at depths of up to 1000 m [4,17,25], so the vocalisations from such animals recorded on bottom-mounted sensors will spend a significant proportion of their time propagating through the deep isothermal layer. Thus the linear sound speed model, which accurately reflects propagation in that layer, should be well suited to this application.

Fig. 5 shows a three-dimensional plot of the result of applying the localisation method based on a linear sound speed profile to the 25 min of data comprising of the second of the Workshop datasets.[3] This plot aims to provide an overview of the full set of estimated source locations, the 174 position estimates are marked as individual points (the 11 outlying samples are not included). An animation showing the full 3D characteristic is available at an associated web page [http://www.isvr.soton.ac.uk/FDAG/uaua/Whale_local.HTM]. In Fig. 6 the same data are shown in plan view and on a larger scale to allow the reader to see the track in relation to the sensor positions which are marked by the symbols '⊕'. The co-ordinate system used in this plot is that supplied with the dataset.

The weighting factors for the least squares cost function (4) are shown in Table 1. These weighing factors are normalised, such that the largest value is unity.

[3] Note that for the purposes of plotting only the vertical scale is (z-direction) is reversed so that negative values represent points below the sea surface.

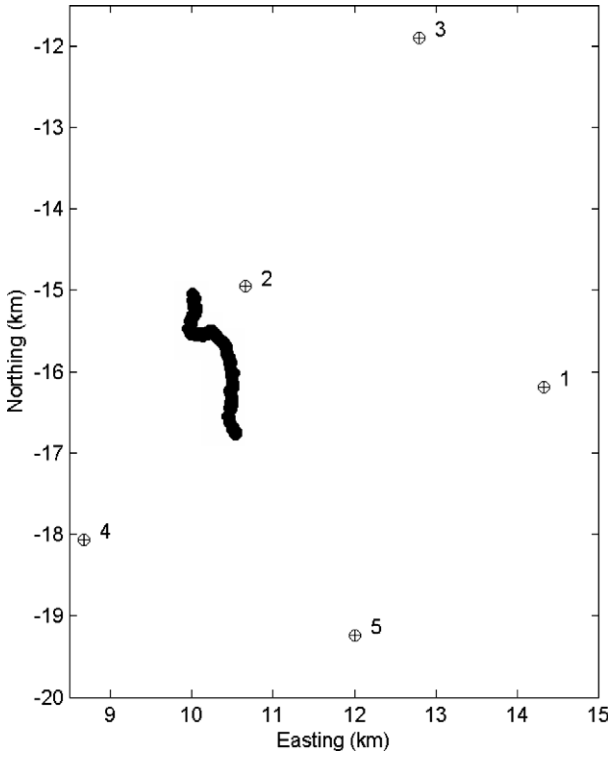

Fig. 6. Large scale plot of estimated whale locations.

Table 1
Normalised weighting least squares factors

Sensor number	Standard deviation (ms)	Normalised weighting factor on direct path	Normalised weighting factor on reflected path
1	–	–	0.0721
2	6.7	0.5692	0.0021
3	8.1	0.3846	0.0688
4	5.8	0.7475	0.0064
5	5.0	1.0000	0.0710

3.1. The effect of noise on delay estimates

The effect of noise on the estimated delays is evaluated using Monte-Carlo techniques. A set of realisations of the delays are created by adding, Gaussian white noise to the delay estimates shown in Fig. 4. The standard deviation of the noise is selected to match the values shown in Table 1. Source location estimates are then computed for each of the realisations and the variance of source locations is computed by averaging the results across the realisations. These results are shown in Figs. 7 and 8, in plan and side view forms, wherein the ellipses surrounding each point represent the 95% confidence intervals. In summary the mean values of the confidence intervals in the x-, y- and z-planes are 8, 10 and 71 m, respectively.

These results confirm that the errors in the vertical plane are significantly greater than those in other axes because of the small aperture in that direction (the maximum difference in the sensor depths is approximately 200 m). With reference to Fig. 6, the animal's direction of motion is in a roughly southerly direction. Examining Figs. 7 and 8 one can see that, as the animal moves towards the centre of the array, the errors in the vertical plane increase, whereas the errors in the horizontal plane reduce; although this latter observation requires close inspection of Fig. 7 since the errors in that plane are small. This observation can be understood by once again considering a symmetric, planar rectangular array. In that case, if the source is located at the centre of the array, then the estimated source locations achieve their maximum accuracy in the plane of the array, whereas it is impossible to localise the source on the perpendicular axis: the relative delays between sensors are identically zero for any source positioned on the axis of symmetry of the array.

3.2. The effect of model mismatch

In this subsection, we shall explore the effects of altering model parameters on the estimated source location. Such an analysis provides insight on the effect of disparities between the assumed model and the actual acoustic environment. Firstly, we consider

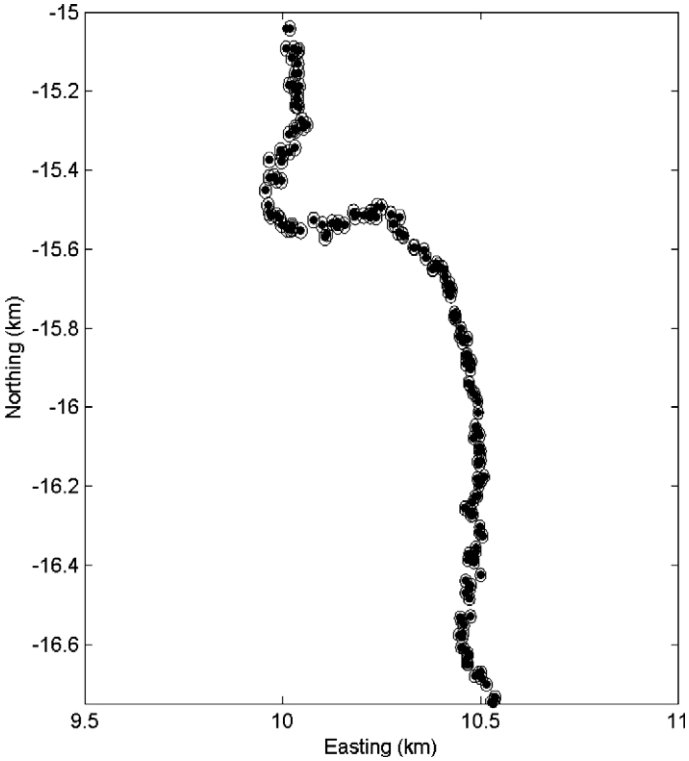

Fig. 7. Plan view of estimated locations. Ellipses around each point define the 95% confidence intervals for the source locations based on errors in the delay estimates.

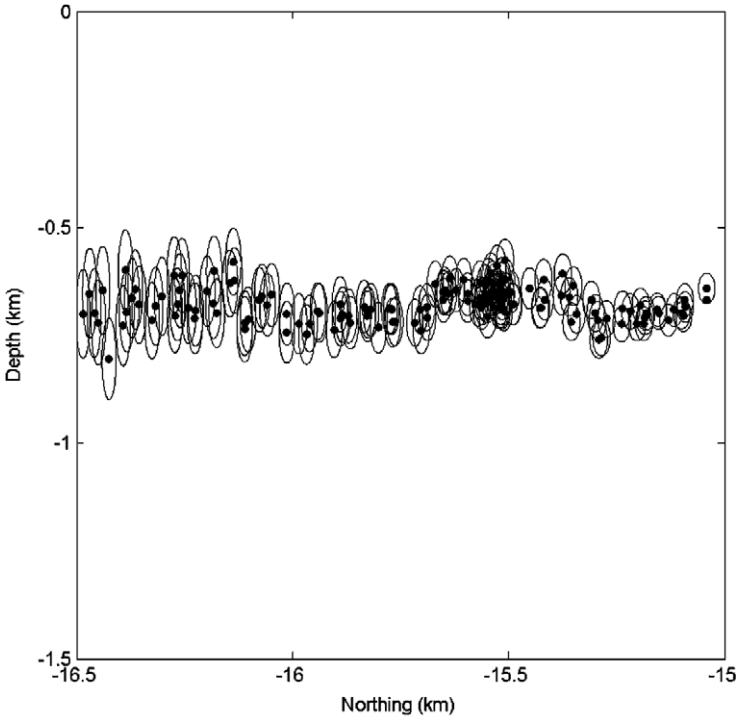

Fig. 8. Side view of estimated locations. Ellipses around each point define the 95% confidence intervals for the source locations based on errors in the delay estimates.

the effect of perturbations of parameter within the same model and then briefly consider the differences between results using different models.

Fig. 9 shows the sensitivity of the linear sound speed model to variations in c_0. These values are computed using a finite difference approach. Results are only shown for sensitivity in two planes for clarity: the results for the y-direction are of a very similar magnitude to those in the x-direction. There are considerable changes in the sensitivity of the model as the source moves within the array. As with the effect of noise, the model becomes more prone to uncertainty in the vertical axis as the source moves towards the centre of the array, whereas the sensitivity in the horizontal axis reduces. The mean sensitivities in these two directions are 0.22 and 4.5 m in the x- and z-directions, respectively.

The model sensitivity and effects of noise can be compared. Herein we choose to achieve this by comparing the mean 95% confidence intervals with the mean sensitivities. In the x-direction the 95% confidence interval equates to a model mismatch of 36 m/s, that is to say that to achieve a perturbation equivalent to the 95% confidence interval one would need to alter the sound speed profile by 36 m/s. In the z-direction a similar analysis reveals that the 95% confidence interval is equivalent to a 15 m/s change in the sound speed profile. This analysis suggests that there is some robustness to the choice of model parameters and that, in this case, the influence of noise on the delay estimates is probably comparable to errors due to inaccurate sound speed estimates.

The above analysis explores the effects of perturbations of the parameters within a particular model. In order to explore the effect of errors between models of different forms we

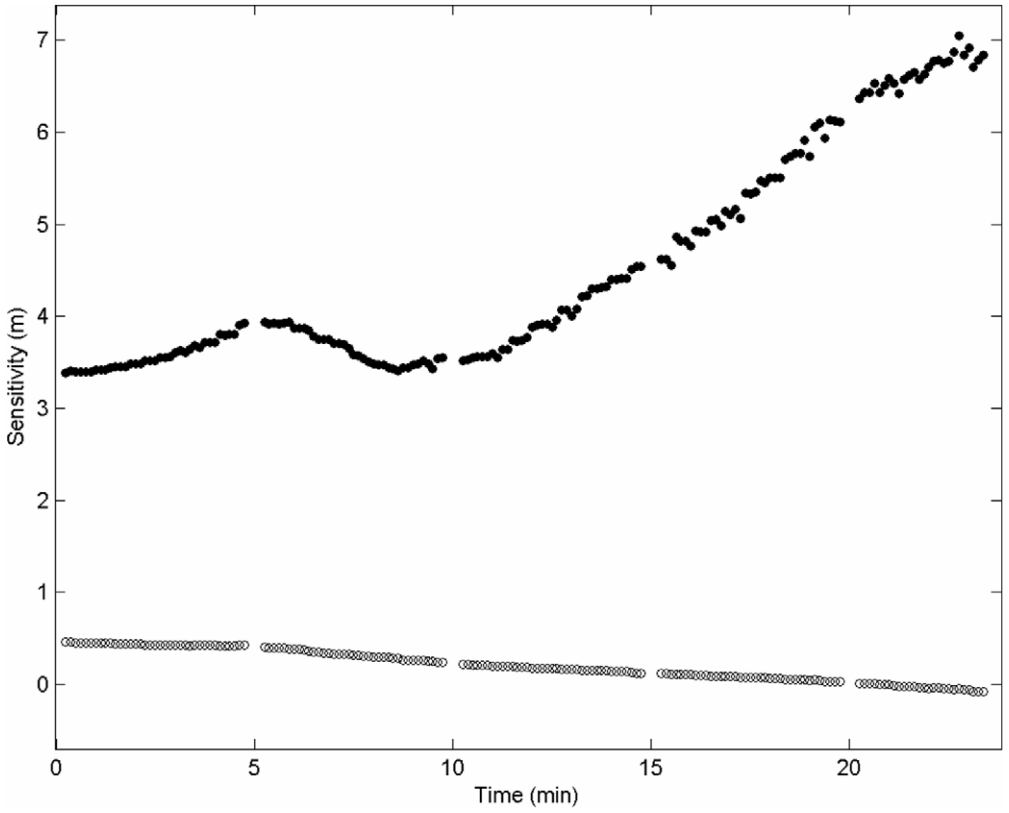

Fig. 9. Model sensitivity in the vertical plane: "○" sensitivity in the x-direction, "●" sensitivity in the z-direction.

compute the source locations using a model based on an isospeed assumption and compare those to the source location computed using a linear sound speed model. The sound speed used for the isospeed model corresponded to the depth-averaged sound speed of linear model used. The depth averaging was performed between 700 and 1500 m, which is the range of depths which the rays spend the majority of their time. Fig. 10 illustrates the differences between the estimated source locations between the two models; only errors in the z-direction are shown. These errors are significantly larger than those seen for simple parameter perturbations. This is unsurprising, because the local differences in the assumed sound speed profiles are greater between the two models than they are in the preceding discussion. Note that we have attempted to ensure that the global properties of the sound speed profiles are in some way equivalent, so it is only the local variations that are contributing to the large errors observed.

3.3. The use of surface reflections

Historically many passive acoustic monitoring systems used in bioacoustics consist of two sensors. Two sensors provide insufficient data to locate a source the three dimensions. In this case surface reflections offer the opportunity to measure delays from a virtual pair of sensors, effectively producing an array of four sensors with which localisation in three

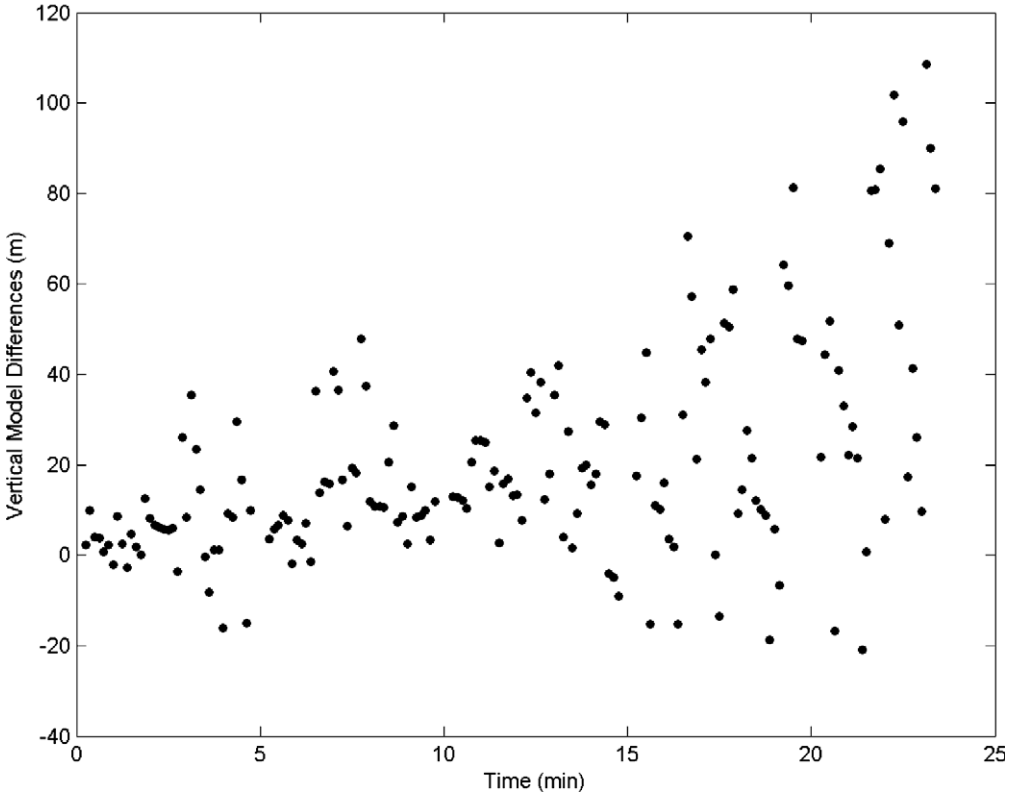

Fig. 10. Vertical differences between source locations estimated using a linear sound speed profile and an isospeed model.

dimensions can be obtained [4]. There remains an inherent left-right ambiguity because of the planar geometry of the four element array. It is to allow localisation of sources with a two element array that has provided the primary motivation of previous work exploiting surface reflections.

Herein the motivation for seeking to exploit surface reflections is rather different. It is based on the observation that the accuracy of a localisation routine is enhanced by increasing the number of sensors. Furthermore, the robustness of a localisation routine to modelling errors improves if the source is located within the sensor field, i.e. the sensors surround the source. In this application we have data from a set of sensors located on the sea floor, in a geometry that is close to planar. When viewed in the x–y plane the source is within the sensor field, so the results are robust to modelling mismatch, in that plane, the results (Section 3.2) are consistent to illustrate this. But in the vertical plane the source is located above all the sensors and so the solution becomes sensitive to modelling errors in that plane. This argument suggests that using the virtual sensors one can provide a set of sensors located above the source and thus create a situation in which the robustness of the source locations is greatly increased and it is to exploit this robustness that we seek to use surface reflections in this application.

In order to explore the use of surface reflections we revert to an isospeed model. Whilst a linear model can be adapted to model surface reflection, the validity of the model is

greatly reduced since the upward travelling rays pass through the upper ocean which is more complex than the deep isothermal layer and is not suitable for modelling with a single gradient. Evidently an isospeed model is an even cruder approximation, but the objective of the analysis in this section is to demonstrate the effectiveness of the technique, rather than to provide accurate locations for this particular dataset.

Fig. 11 shows the sensitivity of the model, in depth, of the isospeed model with and without including the surface reflections. Clearly the model in which the echo has been included is more robust (less sensitive) to model mismatch, the mean sensitivities for these models are 4.2 and 3.2 m for the cases without and with echoes, respectively. Towards the centre of the array (at later times in the recording) the advantage of using the surface reflections increases.

However, this potential offered by the use of surface reflections is limited because of the high levels of variability associated with such delay estimates. This is not a result of the increased path length associated with the reflected rays, because some of the direct paths are longer than some of those for the reflected rays. We believed it to be a consequence of the volatile nature of the acoustic environment in the near-surface region. Factors contributing to such observations include variations in the path lengths caused by ocean swells

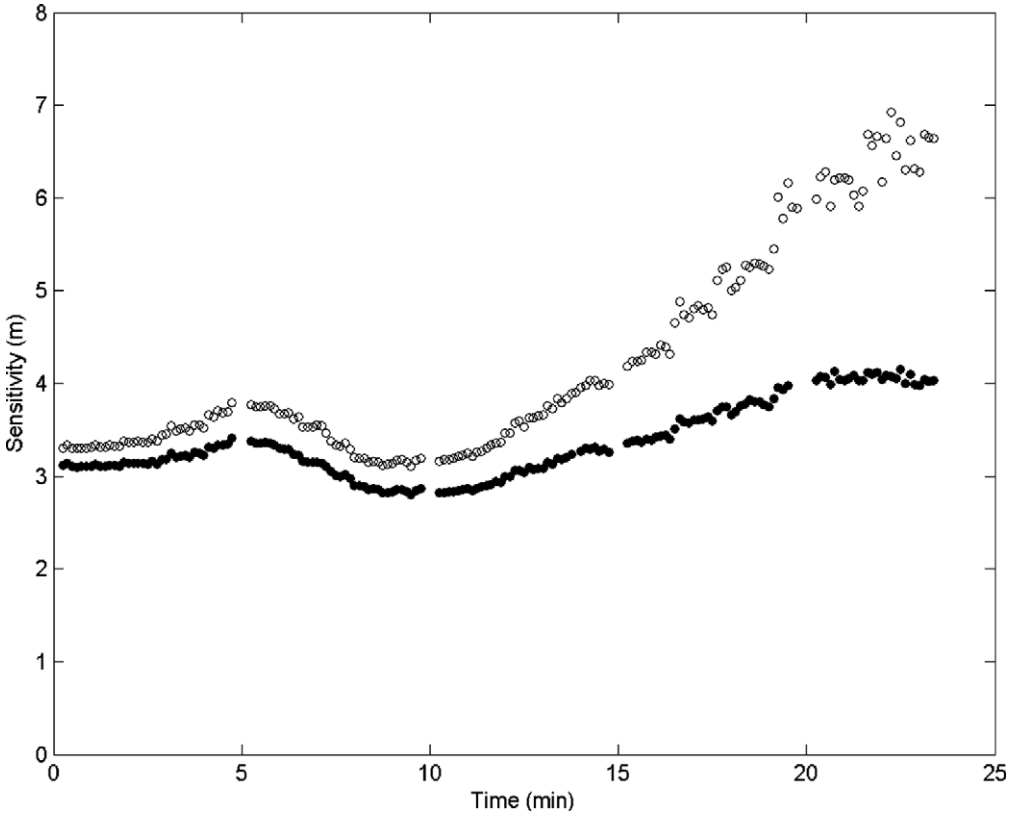

Fig. 11. The depth sensitivity of isospeed models, "○" without using surface reflections, "●" with surface reflections.

and small scale variability in the local sound speed in the near surface region, partly caused by the ambient bubble population and partly by the time-varying roughness of the sea surface [25–29].

The increased variability of delay echoes should highlight the need for care in the use of surface reflections in the two element array case. In such cases the real and virtual sensors may be subject to very different uncertainties which should be accounted for in the algorithm, for example by using a weighted least squares cost function.

4. Conclusions

The results of the preceding section suggest that one can localise sperm whales using bottom-mounted sensors. In particular, the indications are that the location in the horizontal plane is comparatively robust, especially if the animal is surrounded by sensors. Indeed under such conditions the results are comparatively insensitive to both the propagation model and the acoustic parameters used as input for the model. Conversely the results in the vertical plane are strongly influenced by the chosen propagation model. The errors due to model mismatch and noise on the delay estimate are both significant. The robustness of the model in the vertical plane can be increased by the inclusion of data collected at virtual sensors, i.e. measurements of delay along surface reflected paths.

References

[1] Bucker HP. Use of calculated sound fields in and matched-field detection to locate sound sources in shallow water. J Acoust Soc Am 1979;59(2):368–73.
[2] Caiti A et al., editors. Experimental acoustic inversion methods for exploration of the shallow water environment. Kluwer Academic Press; 2000.
[3] Wahlberg M, Mohl B, Madsen PT. Estimating source position accuracy of a large-aperture hydrophone array for bioacoustics. J Acoust Soc Am 2001;109(1):397–406.
[4] Thode A. Tracking sperm whale (*Physeter macrocephalus*) dive profiles using a towed passive acoustic array. J Acoust Soc Am 2004;116(1):245–53.
[5] Tiemann CO, Porter MB, Frazer LN. Localization of marine mammals near Hawaii using an acoustic propagation model. J Acoust Soc Am 2004;115(6):2834–43.
[6] Zimmer WMX et al. Combining data from a multisensor tag and passive sonar to determine the diving behavior of a sperm whale (*Physeter macrocephalus*). IEEE Trans Ocean Eng 2003;28(1):13–28.
[7] Mohl B et al. Sperm whale clicks: directionality and source level revisited. J Acoust Soc Am 2000;107(1):638–48.
[8] Mohl B et al. The monopulsed nature of sperm whale clicks. J Acoust Soc Am 2003;114(2):1143–54.
[9] Zimmer WMX et al. Three-dimensional beam pattern of regular sperm whale clicks confirms bent horn hypothesis. J Acoust Soc Am 2005;117(3):1473–85.
[10] Zimmer WMX et al. Off-axis effects on the multipulse structure of sperm whale usual clicks with implications for sound production. J Acoust Soc Am 2005;118(5):3337–45.
[11] Burdic WS. Underwater acoustic systems. 2nd ed. Englewood Cliffs (NJ): Prentice-Hall; 1991.
[12] Zanolin M et al. Asymptotic accuracy of geoacoustic inversions. J Acoust Soc Am 2004;116(4):2031–42.
[13] Carter GC. Coherence and time-delay estimation. Proc IEEE 1987;75(2):236–55.
[14] White PR. Detection algorithms for underwater acoustic transients. In: Leondes CT, editor. Control and dynamic systems advances in theory and applications, vol. 77. New York: Academic Press; 1996. p. 193–224.
[15] Thode A et al. Localization using Bartlett matched-field processor sidelobes. J Acoust Soc Am 2000;107(1):278–86.
[16] Spiesberger JL, Wahlberg M. Probability density functions for hyperbolic and isodiachronic locations. J Acoust Soc Am 2002;112(6):3046–52.
[17] Thode A. Three-dimensional passive acoustic tracking of sperm whales (*Physeter macrocephalus*) in ray-refracting environments. J Acoust Soc Am 2005;118(6):3575–84.

[18] Spiesberger JL. Hyperbolic location errors due to insufficient numbers of receivers. J Acoust Soc Am 2001;109(6):3076–9.
[19] Press WH et al. Numerical recipes in C: The art of scientific computing. Cambridge: Cambridge University Press; 1993.
[20] Kirkpatrick S, Gelatt CD, Vecchi MP. Optimisation by simulated annealing. Science 1983;220(4598):671–80.
[21] Man KF, Tang KS, Kwong S. Genetic algorithms. Berlin: Springer; 1999.
[22] Boyd S, Vandenberghe L. Convex optimisation. Cambridge: Cambridge University Press; 2004.
[23] Van Trees HL. Detection estimation and modulation theory. New York: Wiley; 1968.
[24] Clay CS, Medwin H. Acoustical oceanography: principles and applications. New York: Wiley; 1977.
[25] Miller PJO, Johnson MP, Tyack P. Sperm whale behaviour indicates the use of echolocation click buzzes 'creaks' in prey capture. Proc R Soc Lond B 2004;271:2239–47.
[26] Leighton TG. The acoustic bubble. New York: Academic Press; 1997. p. 263–86.
[27] Medwin H. Acoustic fluctuations due to microbubbles in the near-surface ocean. J Acoust Soc Am 1974;56(4):1100–4.
[28] Dashen R, Henyey FS, Wurmser D. Calculations of acoustic scattering from the ocean surface. J Acoust Soc Am 1990;88(1):310–23.
[29] McDonald BE. Echoes from vertically striated subresonant bubble clouds: a model for ocean surface reverberation. J Acoust Soc Am 1991;89(2):617–22.