Procedure description: using AUTEC's hydrophones surrounding a DTAGed whale to obtain localizations

Tiago A. Marques[1], Jessica Shaeffer[2] , David Moretti[2] & Len Thomas[1]

[1] Centre for Research into Ecological and Environmental Modelling, The Observatory, University of St Andrews, St Andrews, KY16 9LZ, Scotland.
[2] Naval Undersea Warfare Center Division, 1176 Howell Street, Newport, Rhode Island 02841, U.S.A.

Contents

1 Introduction

One of the possible uses of DTAG (Johnson and Tyack, 2003) data is to estimate the actual track of an animal based on the measurements made on the DTAG. This is referred to as obtaining a "pseudo-track". The reason this is called a pseudo-track is because the procedure involved is prone to errors. However, if one combines the

DTAG information with independent localizations of the animal, the procedure can become much more robust. This is referred to as "georeferencing" the track.

As part of the process of georeferencing a dive track using DTAG data, AUTEC hydrophones provide localizations of the tagged animal that are used to "adjust" a pseudo-track derived from DTAG data alone. In this report we describe how the AUTEC data is used jointly with the DTAG data to arrive at animal localizations, which can then be used as inputs into the procedure of obtaining a track from DTAG data. The georeferencing itself is elsewhere described (Marques and Thomas, 2012a) and implemented (Marques and Thomas, 2012b). Here we look in detail to the "localization" procedure. This goes from the underlying raw acoustic data, i.e. actual time of arrival (TOA) of clicks at AUTEC hydrophones surrounding the tagged whale, plus the acoustical DTAG data, all the way up to localizations of the animal at each click detected on surrounding hydrophones.

In most of the process described in this report, the only data that gets used from the DTAG is the acoustic data, that is, the time of emission of the clicks from the tagged whale. The exception is the last step, the localization algorithm itself, where as described in section 3.5, data from the pressure sensors, namely depth, is also used.

We explicitly ignore here all the processing of DTAG data that goes on "upstream" to obtain the pitch, heading and roll data, required for the pseudotracking / georeferencing itself. Readers are referred to Johnson and Tyack (2003) for details on this, including an example on how to use the DTAG data to obtain a pseudotrack. We note however that errors introduced by said process might propagate upwards and, hence, be responsible for a large proportion of the measurement error in the georeferenced localizations, and would probably deserve to be looked at in close detail. That is not attempted here, and (despite knowing that they are not!) we take those measurements to be error free.

The pre-processing of data required to obtain the inputs into the localization algorithm is a relatively complex procedure. Due to its complexity it is hard to quantify the uncertainty present in a given click localization (x_n, y_n) estimated for the nth click, $n = 1, 2, ..., N$. Because the acoustic-based localizations are one of the inputs in the current procedure used for georeferencing (see e.g. ISEC 2012 talk), it is important to understand this process thoroughly. In particular, it would be interesting to build on it to come up with an integrated inferential framework which would then allow for estimation of the track given DTAG data and AUTEC hydrophones TOA's. This would in turn allow to propagate the uncertainty in the acoustic localizations into the final track estimate (i.e. incorporating both DTAG data and acoustic localizations).

We note upfront that, to some extent, this might be an overkill: while accounting for the uncertainty in the AUTEC's acoustic localizations is elegant and conceptually appealing, it might be the case that these are very precise when compared to the DTAG data alone, and so, the variance will be nonetheless dominated by the latter.

In the next section we briefly describe the several steps involved, providing an overview of the entire process. This is then followed by a separate section, in which the procedure steps which require further clarification are expanded in separate sub-sections.

2 Overview of localization algorithm

We have a “template” sound file from the DTAG, with clicks time-of-emission (TOE). We assume that all sounds not representing clicks from the tagged whale have been identified and removed, and hence, that we have an error free record of all clicks TOE.

As stated above, the raw data corresponds to the TOA of clicks at each hydrophone. These TOA’s could, a priori, correspond to one of 3 categories:

1. TOA actually corresponds to a click emitted by the tagged animal;
2. TOA actually corresponds to a click emitted by any other animal;
3. TOA actually corresponds to a false positive (i.e. not corresponding to a beaked whale click at all).

For our analysis, we hope to identify which belong to the tagged animal (1) vs those who do not (2 and 3).

The hydrophones are assumed to be perfectly synchronized, so that hydrophone’s time-difference-of-arrival (TDOA) can be calculated from TOAs. On the contrary, note that the DTAG is not perfectly time synchronized with hydrophones and that the DTAG likely suffers from clock drift. However, the recorded “DTAG time-on” is used to time align the sound records from the DTAG and the hydrophones. More than once this created problems, because it was later confirmed that the “DTAG time-on” was incorrect. Usually, an up to 2 seconds difference shows up, which is of little consequence. That difference might be up to say 5 seconds in ‘bad” cases, but still not an issue. For really “pathological” cases it could be much more than this, and such cases require identification as they could, otherwise, mess up the entire procedure. An obvious example of problems would be when the TOA at the hydrophone is before the TOE at the DATG. However, we note that perfect synchronicity is not required because it is the TDOA between pairs of hydrophones that gets used in the end for localization, and any time misalignment between

DTAG and hydrophones gets canceled out because the hydrophones themselves are perfectly synchronous.

Naturally, the time resolution is much higher on the DTAG, sampling at a very high frequency, compared to the hydrophones. After processing the data through the detector, Morrissey *et al.* (2006) refers the use of a 10.7 ms window, which stems from the FFT time resolution being 10.67 ms. Further, the FFT uses 50% overlap. This should also be considered a quantifiable error associated with detection time, which is then propagated into localization error.

A number of parameters need to be set before implementing the procedure:

- DTAG time window of 6 seconds; this corresponds to the "chunk" of time on the DTAG that gets processed: for each click considered we use as data for cross-correlation the recording time from TOE and TOE + 6 s;
- Hydrophone time window, Morrissey *et al.* (2006) refers the use of a 10 s window. We consider the current click TOA at the DTAG to be time 0. Then clicks before and after 10 seconds in the hydrophone are searched for. This value was considered to make sure that the click in question would be, if detected, present in this interval on all surrounding hydrophones;
- the bin width over which one considers whether there is a click or not. Note there are two different "units" to be considered here:

  1. the bin that gets slided each time, to get the TDOA, referred to as the comb sieve bin width. This is set to 5 ms.
  2. the bin over which a click is considered to be present if it is in the same time bin in the DTAG and the hydrophone data. This is set, in practice, to 10 ms

  This is a consequence of how the algorithm is coded. The construct sieve function ("con_sieve.m") automatically adds an additional 5ms bin subsequent to the detection bin, e.g. a "fudge factor", which means that a click is considered to be present if it is in the same 10 ms bin in the DTAG and the hydrophone data. JW found that using a 10 ms comb sieve bin width to begin with, which would lead to effectively 20 seconds bin for matching, resulted in many instances where more than one time delay would have the same number of "matches" due to the presence of noise, conspecific clicks, and steady click repetition rate.

Using the data and the above parameters, for each of the K hydrophones surrounding the DTAGed whale, the following procedure is then repeated for all the N clicks present in the DTAG record.

For each hydrophone ($k, k = 1, 2..., K$):
For each click ($n, n = 1, 2, ..., N$):

1. time align DTAG and hydrophone data using the "DTAG time-on" and the AUTEC's hydrophone clock time; For each possible time alignment over the 20 seconds window (i.e. sliding the hydrophone sound file 5 ms at the time):
   (a) calculate and store the number of bins (of 10 ms) which simultaneously have clicks on them
   (b) go to the next TDOA; if it is the last within the 20 second window go to next step
2. assign a TDOA to the n^{th} click, corresponding to the time difference that results in the largest correlation between the DTAG data and the current hydrophone data (further details on section 3.1);
3. go to the next click; if it is the N^{th} click, go to the next step, else go back to point 1 above
4. go to the next hydrophone; if it is the last hydrophone, procedure ends.

Then, using all the hydrophone-DTAG TDOA's for all the click and hydrophone combinations obtained as described above:

1. plot the resulting TDOA data, with time along the dive in the x axis and TDOA on the y axis, which results in a "TDOA plot" (further details on section 3.2);
2. "clean up the TDOA plot" from clearly imconsistent TDOAs, resulting from spurious correlations (further details on section 3.3);
3. Interpolate the available TDOA measurements to obtain a TDOA for (essentially) every click and every hydrophone (further details on section 3.4);
4. transform these TDOA's between DTAG and hydrophone into hydrophone TOA and feed the resulting TOA data into a localization algorithm, finally obtaining each click's localization (further details on section 3.5).

3 Zooming in on key algorithm steps

3.1 Creating TDOA's

Note at this stage TDOA's are between the hydrophone and the DTAG, not between pairs of hydrophones, although one can get the later from the former.

For hydrophone k, the outcome of this previous step can be seen as a vector of 0's and TDOA's, represented by $TDOA_{k.}$, where say

$$TDOA_{k.} = (0, TDOA_{k2}, TDOA_{k3}, 0, 0, TDOA_{k6}, ...)$$

means that of the first 6 clicks, only the 2nd, 3rd and 6th clicks were detected in that hydrophone, and so only for these TDOA's are available.

Note that in practice there is a very small number of 0's (if any), because there is no "correlation quality" check in this step. Even if all correlations are really low, there is always one which is the highest. This contributes for adding noise (i.e. spurious correlations) to TDOA plots (see below).

However, in the MATLAB code, several quality variables are created and outputed as part of the process, and hence could be used in the future to provide some validation. These quality check quantities include variables like number of TOAs in window, number of TOEs in window, number of matches and number of similar matches (i.e. 2, 3 or more different TDOA's could all produce the same number of matches; in particular, the code defaults to use the first one if that is the case).

3.2 TDOA plots

These present conspicuously to the naked eye a set of "TDOA lines", which would be horizontal lines if the animal was stationary, but have slopes depending on how fast the animal is approaching or moving away from the hydrophone involved in a particular TDOA.

At this stage the plots are cluttered with TDOA's, many (if not most) of which are necessarily spurious, and hence a cleaning step is implemented, as described next.

3.3 Cleaning TDOA data

While we tend to refer to this as cleaning TDOA plots, strictly, one is cleaning the data, not the plot.

Spurious correlations, resulting from a multitude of reasons, will contribute with considerable noise to the TDOA plots. As mentioned before, there are no checks for the quality of the called TDOA, so the largest correlation wins, but actually that could correspond in the limit to a single click being in the right time bin, which could be caused by false positives (say false association of clicks).

For each hydrophone, the TDOA data is filtered by thresholding the histogram of TDOAs in sequential blocks of time and then manually selecting whether to include the selected points in the final TDOA output. The procedure involved is:

For each 1 minute period in the DTAG:

- create an histogram from the minimum TDOA to the maximimum TDOA, with bins with a width of 0.01 seconds
- find the bin with the maximum number of TDOA's (the "maximum class")
- select all the TDOA's within 0.05 seconds of that bin and discard all others
- (a manual check of whether this was a reasonable thing to do, using a plot with different colors for the removed and kept TDOA's, was routinely implemented)

The rationale behind this procedure is that given the hydrophones are fixed and the animal moves slowly, TDOA cannot change abruptly over a short time period. This step depends, to some extent, on an arbitrary threshold, namely the bin width used and the range of TDOA's considered, and these values where obtained based on a trial and error empirical process, ending up with values that lead to reasonable tracks.

An additional check that might be (but was not) implemented would be to check that successive "maximum class" values were close, as these should not vary abruptly either.

3.4 Interpolation

Due to the clicks narrow beam pattern, usually each click is at best only detected at one or two, maximum 3, hydrophones. Using interpolation on can obtain a TDOA for the clicks which were not detected at a given hydrophone, provided clicks (not much) before and/or after were detected. Thes corresponds to interpolating the clearly visible TDOA lines in the plots for times of clicks not detected. Therefore, after the cleaning up step, the resulting TDOA is interpolated using a piecewise method that preserves the original data. A binary matrix also records which TDOAs are real versus interpolated.

For the conspecific analysis, the TDOA data was interpolated in sections, so that a pause of 20 s or greater with no data would result in a new section being created for interpolation. No interpolation was attempted during those longer gaps with no measurements. This was not what was done for the georeferentiation used in Marques *et al.* (2009), when the start and end times for interpolation were manually chosen.

3.5 Localization of clicks

Finally, the (hydrophone-DTAG) TDOA sets are converted to hydrophone TOAs by adding the DTAG TOE to the TDOA vector. The TOA vectors are independently solved by the localization algorithm using the depth of the tagged whale for each click.

These position estimation problems typically utilize an hyperbolic model[1]. The model generates a system of non-linear equations which must be solved to estimate the object's position. Additionally, a number of localization algorithm "flavours" are possible: as an example, one might choose between considering the propagation velocity sensor dependent versus constant across sensors. We do not dwell into the details of the actual algorithm here. The reader is referred to Vincent II (2001) for further details on localization algorithms and many additional references on the subject.

The localization algorithm not only produces an estimate for the animal position when the click was emitted, as it also produces an estimate of the TOE of the click. Checking this TOE against what it was supposed to be given the initial time alignment was always done, as it allows an additional check for problems.

Note localization works even if the DTAG and the hydrophones are not perfectly synchronous, because what gets used are the between hydrophone TDOA's, and any asynchronous behavior cancels out when making the differences between the TOA's for pairs of hydrophones, as hydrophones are perfectly synchronous.

At this point there is a choice to be made about the quality of the TOA's used for localization. One could consider localizing the animal only for:

- all clicks for which there are at least 3 interpolated values (this actually means you have sections of the localizations considered which are really smooth, because they are based on smoothed TDOA's), or
- for clicks for which there are at least two actual measured TDOA's + 1 interpolated TDOA (this was what actually was used for the georeferencing prior to the analysis reported by Marques *et al.* (2009), and leads to more noisy localizations).

The TOA vectors are independently solved by the localization algorithm for each click. It is important to state explicitly that the actual depth, as measured from the DTAG (remember the pressure sensors from which depth is derived are very accurate) is used as an input in the localization algorithm, hence significantly improving the ability to localize in the (x, y) plane. Note this also means you only need 3 TDOA's (for an exactly determined solution, from 4 onwards on gets an over-determined solution), rather than 4 required if depth was not used as an input, for the localization algorithm to work.

[1] This is a consequence of the fact that the TOE is unknown. If it were known, we could use a spherical model. Note that had the DTAG and hydrophones been perfectly synchronized, that would be the case.

Because localization is done independently on a click by click basis, it might be inefficient, but on the other hand, it actually means once we plot all the localizations in 3D space if they do not look like a track, something went wrong: so there is some internal consistency check that is automatically implemented. On the other hand, if the process was not independently done click by click, the output might look like a reasonable track just because the way it was built successively using prior results.

4 Conclusions and potential work ahead

It is our hope that this report can serve as a record of what was done to obtain localizations based on data collected at AUTEC hydrophones combined with the DTAG data.

A number of areas that could benefit from further work include:

1. look closer to the procedure that occurs prior to obtaining the pitch and heading from the DTAG accelerometer and magnetometer data. While this data was not used here, because it is then combined with the locations obtained here, it could have a major impact in the final output;

2. quantify uncertainty/precision for the x, y coordinates obtained through localization. These might presumably be obtained as a by-product from the optimization algorithm (this is possible especially when one has more than 3 TDOA's for the localization algorithm);

3. Understand the sensitivity of the procedure to a number of ad hoc settings, including:

   - bin width for the comb sieve
   - adjustment for bin width for a match to be called
   - time over which correlation is computed (in both DTAG and hydrophone records)
   - histogram settings for TDOA clean up

4. develop a one-stop-shop procedure that takes as inputs the heading, pitch, depth and speed (strictly required to get pseudo-tracks) and the DTAG TOE's and hydrophone TOA's. This would allow at once to obtain the localizations that both best fit the acoustic data and the DTag data, and that included error in the hydrophone TDOA's in an appropriate way.

References

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

Marques, T. A. and Thomas, L. (2012a). Obtaining pseudo-tracks and georeferenced locations from DTAG data. Technical report, LATTE WORKING DOCUMENT.

Marques, T. A. and Thomas, L. (2012b). Using dedicated kalman filter r packages to fit the georeferencing ssm. Technical report, LATTE WORKING DOCUMENT.

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. *The Journal of the Acoustical Society of America*, **125**, 1982–1994.

Morrissey, R. P., Ward, J., DiMarzio, N., Jarvis, S., and Moretti, D. J. (2006). Passive acoustic detection and localization of sperm whales (*Physeter macrocephalus*) in the tongue of the ocean. *Applied Acoustics*, **67**, 1091–1105.

Vincent II, H. T. (2001). *Models, algorithms, and measurements for underwater acoustic positioning*. Ph.D. thesis, University Of Rhode Island.