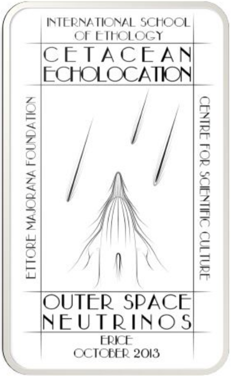
Cetacean density estimation from passive acoustic data

(*or avoiding the indigestion after a delicious spaghetti meal*)

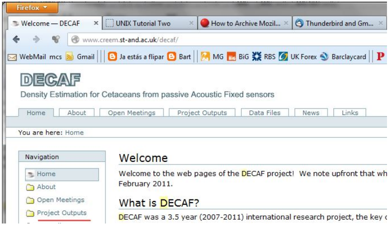
Tiago A. Marques, Danielle Harris & Len Thomas

Centre for Research into Ecological and Environmental Modeling, University of St Andrews

21-10-2013



Acknowledgments & DECAF



- Gianni Pavan for the invitation
- All DECAF team & steering committee, John Hildebrand's team, SAMBAH team and Luis Matias for discussions and collaborations
- DECAF funding under the National Oceanographic Partnership Program. Partners funding our project were the Ocean Acoustics Program of the US National Marine Fisheries Service Office of Protected Resources and the Joint Industry Program (via the National Fish and Wildlife Foundation). Cheap DECAF funded by ONR.

Physics and biology

Cetaceans and neutrinos

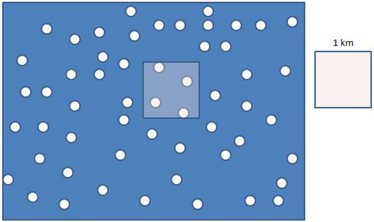
Acoustic surveys differences and similarities

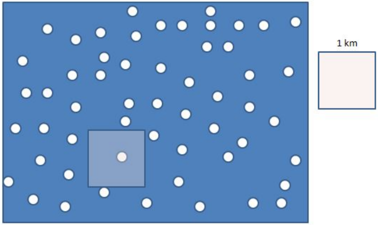
- as for neutrinos, we do not detect cetaceans but an outcome of their interaction with the environment
- availability bias is extreme for neutrinos: "many" cetaceans vs. "few" neutrinos potentially detectable
- false positives and missed events must be considered: but many more trials for cetaceans, so missing/misclassification of some not an issue
- from data to models and from models to data: relying on theoretical models to define the characteristics of the signal to detect

What is density?

Number of animals per unit area: a quantity that changes over time and space

Key: defining time and space we are interested in making inferences for





Location is fundamental

The problem at hand

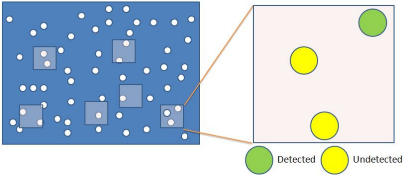
Objective

Based on acoustic data estimate average density (D) and/or abundance ($N = D \times A$) over a study area of size A and time period T

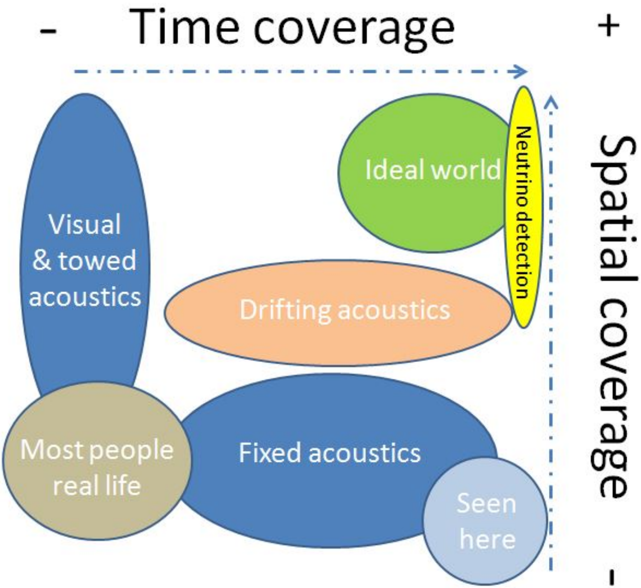
Approach

Given that a census is hardly ever an option, we must rely on some sort of sampling:

- sample of areas to cover within region of interest - sampling, survey design, design and model based estimates
- sample of animals (or cues) within covered areas - key interest here



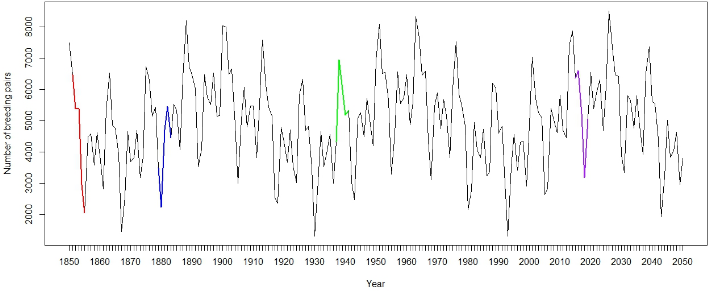
Temporal and spatial survey scales



Sampling over time is easier with acoustics

Assessing impacts on animal populations

Understanding natural fluctuations in wild populations is fundamental



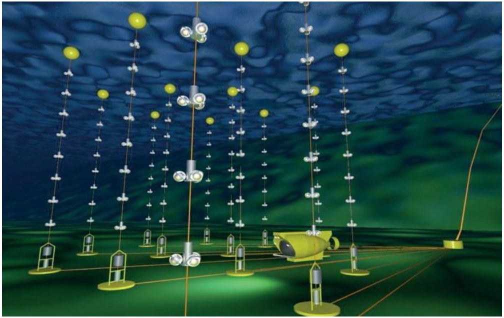
Time scale

Methods which can be used to collect data on density / abundance over long time periods are ideal

Sampling over space is key

Most existing acoustic data are NOT randomly located in space

More often, these come from single sensor/location, but density required at wider spatial scales



Do not forget

Usually key interest is at the population or species level

Traditional approaches

...to abundance estimation either at sea or inland, have largely been based on two groups of (mostly visually based) methods

Distance sampling

- Line transects (e.g. aerial or boat based line transect surveys)
- Point transects (e.g. point transects for birds)
- Cue counting (e.g. surveying whales from their blows)

Mark Recapture

- physical trapping of individuals
- Photo-ID
- DNA-based methods

Visual vs. acoustic-based methods

Visual methods are:

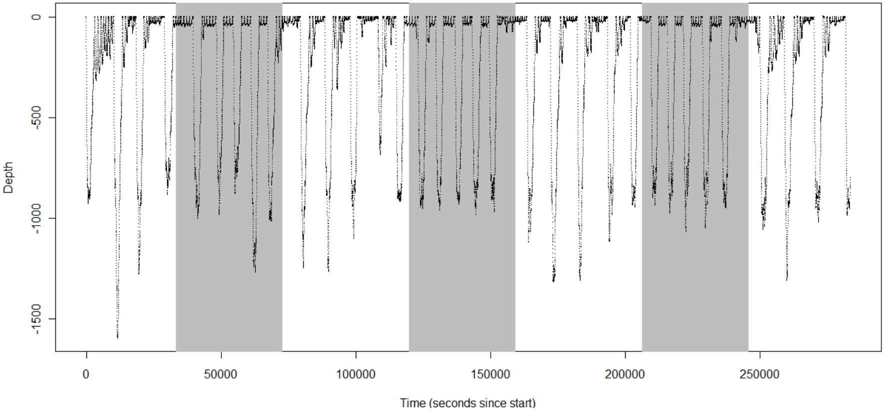
- Dependent on good weather conditions
- Restricted to daylight periods
- Problematic for animals which are submerged for long periods
- Temporarily restricted
- Expensive in the long term
- Dependent on a hard to model human search process

Why most acoustic work to date is on marine mammals?

All these, coupled with the fact that many marine mammal species produce sounds and are naturally expensive to survey has made them specially appealing to acoustic-based methods

To be or not to be (seen)? That is the question!

Mesoplodon densirostris, Bahamas (± 80 hours)



Data from Robin Baird

Intuitive canonical estimator

Given n detected animals (or cues) abundance can be estimated by

$$\hat{D} = \frac{n}{\hat{C}}$$

C includes multipliers to account for many factors:

- Detection probability or effective survey area
- Sound production rate
- False positives
- Observation effort (line length, recording time, number of sensor used, etc)

Formulation

so density is

$$D = \frac{N}{A}$$

- D - density
- N - number of animals
- A - area

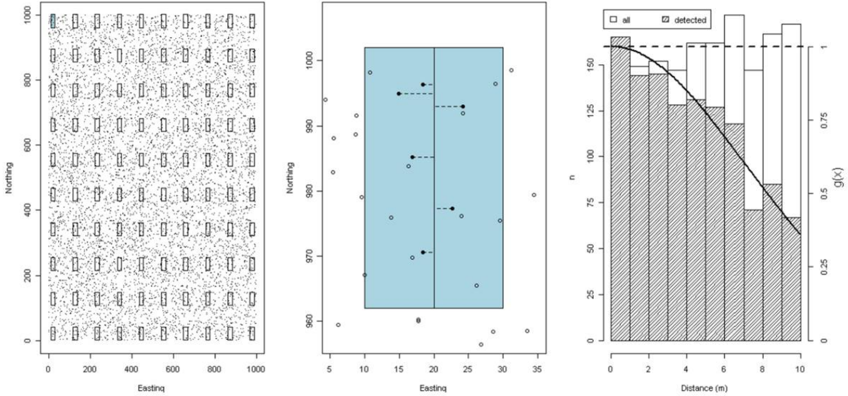
if we can't count all animals in the covered area:

$$D = \frac{n}{ap}$$

- n - number of detected animals
- a - covered area
- p - detection probability

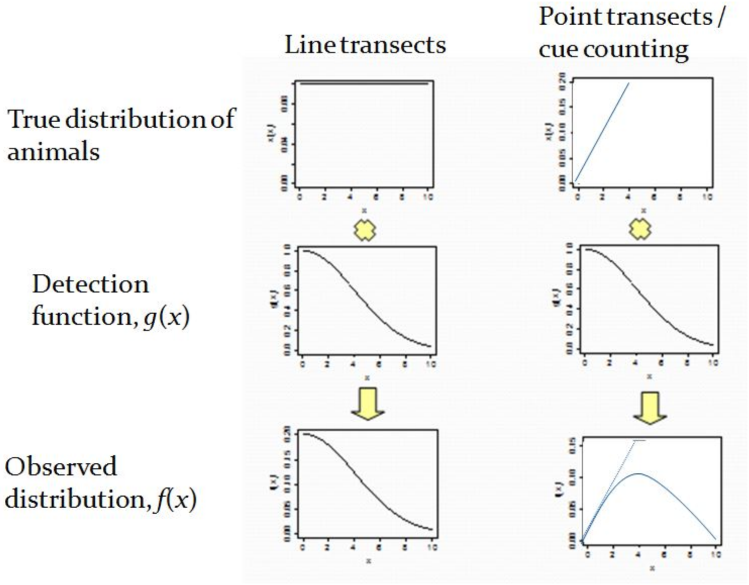
Probability of detection or effective sampling area

Distance sampling concepts: p is the proportion of the area under the detection function with respect to the total area



"essentially, all models are wrong, but some are useful", George Box

Probability of detection or effective sampling area



Acoustics-based density estimation

Many terrestrial surveys are already mixed or essentially acoustic, as the animals are only heard and rarely ever seen (e.g. birds in closed canopy forests), with humans acting as microphones

Some taxa's densities have already been estimated based on acoustic data

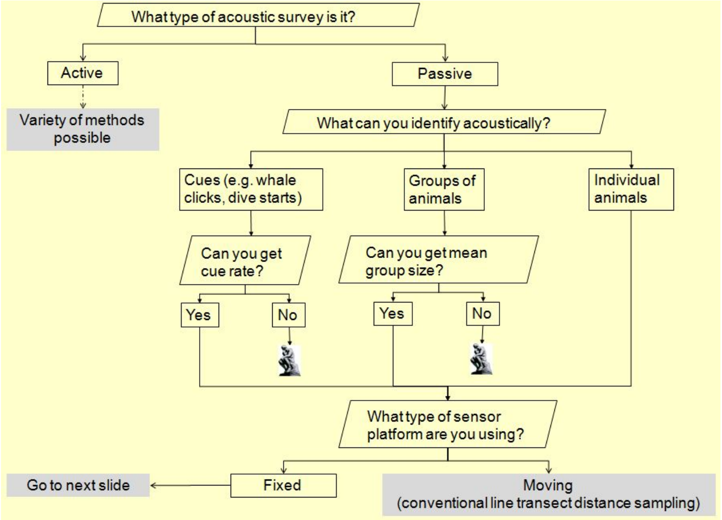
- Birds using Spatially Explicit Capture Recapture (SECR, Efford and Dawson 2009)
- Elephants using regression models for detection function (Thompson *et al.* 2009, 2010)
- several cetacean species (examples later!)

Work in progress on

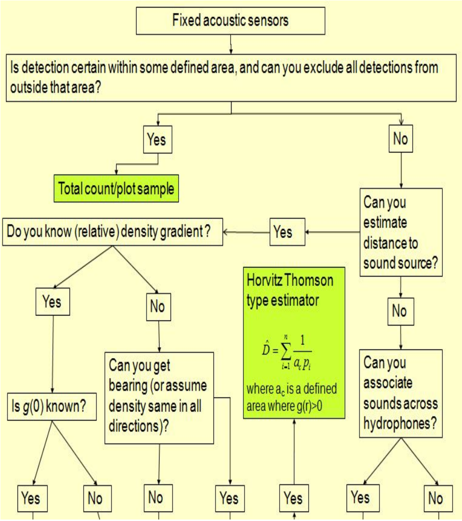
other taxa just waiting to get densities estimated using acoustics

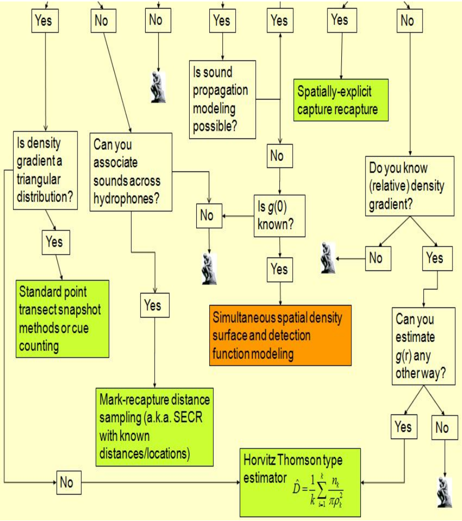
- fish (marine and freshwater)
- amphibians
- primates
- insects
- bats
- others? (...like neutrinos :)

A framework for estimating density from acoustic data



A framework for estimating density from acoustic data





A framework for estimating density from acoustic data

BIOLOGICAL REVIEWS

Cambridge Philosophical Society

Biol. Rev. (2013), 88, pp. 287–309.
doi: 10.1111/brv.12001

Estimating animal population density using passive acoustics

Tiago A. Marques[1,2,*], Len Thomas[1], Stephen W. Martin[3], David K. Mellinger[4], Jessica A. Ward[5], David J. Moretti[5], Danielle Harris[1] and Peter L. Tyack[6]

[1] *Centre for Research into Ecological and Environmental Modelling, University of St Andrews, The Observatory, Buchanan Gardens, Fife, KY16 9LZ, UK*
[2] *Centro de Estatística e Aplicações da Universidade de Lisboa, Bloco C6, Piso 4, 1749-016, Lisboa, Portugal*
[3] *Space and Naval Warfare Systems Center Pacific, 53560 Hull Street, San Diego, CA 92152 USA*
[4] *Cooperative Institute for Marine Resources Studies, Oregon State University and NOAA Pacific Marine Environmental Laboratory, 2030 SE Marine Science Drive, Newport, OR 97365, USA*
[5] *Naval Undersea Warfare Center, 1176 Howell Street, Newport, RI 02841, USA*
[6] *Sea Mammal Research Unit, Scottish Oceans Institute, University of St. Andrews, Fife, KY16 8LB, UK*

ABSTRACT

Reliable estimation of the size or density of wild animal populations is very important for effective wildlife management, conservation and ecology. Currently, the most widely used methods for obtaining such estimates involve either sighting animals from transect lines or some form of capture-recapture on marked or uniquely identifiable individuals. However, many species are difficult to sight, and cannot be easily marked or recaptured. Some of these species produce readily identifiable sounds, providing an opportunity to use passive acoustic data to estimate animal density. In addition, even for species for which other visually based methods are feasible, passive acoustic methods offer the potential for greater detection ranges in some environments (e.g. underwater or in dense forest), and hence potentially better precision. Automated data collection means that surveys can take place at times and in places where it would be too expensive or dangerous to send human observers.

Here, we present an overview of animal density estimation using passive acoustic data, a relatively new and fast-developing field. We review the types of data and methodological approaches currently available to researchers and we provide a framework for acoustics-based density estimation, illustrated with examples from real-world case studies. We mention moving sensor platforms (e.g. towed acoustics), but then focus on methods involving sensors at fixed locations, particularly hydrophones to survey marine mammals, as acoustic-based density estimation research to date has been concentrated in this area. Primary among these are methods based on distance sampling and spatially explicit capture-recapture. The methods are also applicable to other aquatic and terrestrial sound-producing taxa.

We conclude that, despite being in its infancy, density estimation based on passive acoustic data likely will become an important method for surveying a number of diverse taxa, such as sea mammals, fish, birds, amphibians, and insects, especially in situations where inferences are required over long periods of time. There is considerable work ahead, with several potentially fruitful research areas, including the development of (i) hardware and software for data acquisition, (ii) efficient, calibrated, automated detection and classification systems, and (iii) statistical approaches optimized for this application. Further, survey design will need to be developed, and research is needed on the acoustic behaviour of target species. Fundamental research on vocalization rates and group sizes, and the relation between these and other factors such as season or behaviour state, is critical. Evaluation of the methods under known density scenarios will be important for empirically validating the approaches presented here.

Key words: acoustic surveys, bioacoustics, density estimation, distance sampling, passive acoustic monitoring, spatially explicit capture-recapture, fixed sensors, hydrophones.

* Address for correspondence (Tel: 00441334461801; Fax: 00441334461842; E-mail: tiago@mcs.st-and.ac.uk).

Marques *et al.* (2013). Estimating animal population density using passive acoustics. *Biological Reviews*. **88**: 287-309. DOI: 10.1111/brv.12001

Thoughts from a Nobel prize winner

Ernest Rutherford
(1871-1937)

"If your experiment needs statistics, you ought to have done a better experiment"

Thoughts from a Nobel prize winner

Ernest Rutherford
(1871-1937)

"If your experiment needs statistics, you ought to have done a better experiment"

"All science is either physics or stamp collecting"

Thoughts from a Nobel prize winner

Ernest Rutherford
(1871-1937)

“If your experiment needs statistics, you ought to have done a better experiment”

“All science is either physics or stamp collecting”

”We haven’t the money, so we’ve got to think.”

Thoughts from a Nobel prize winner

Ernest Rutherford (1871-1937)

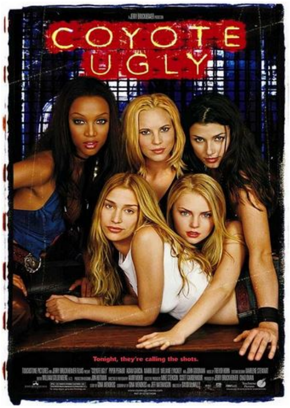
“If your experiment needs statistics, you ought to have done a better experiment”

“All science is either physics or stamp collecting”

"We haven't the money, so we've got to think."

"An alleged scientific discovery has no merit unless it can be explained to a barmaid."

Thoughts from a Nobel prize winner



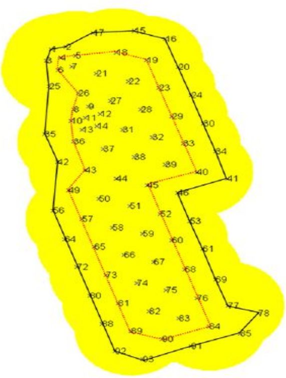
“If your experiment needs statistics, you ought to have done a better experiment”

“All science is either physics or stamp collecting”

”We haven't the money, so we've got to think.”

”An alleged scientific discovery has no merit unless it can be explained to a barmaid.”

Group counting of beaked whales



screenshot “stolen” from Moretti *et al.* 2010

- assume a group diving within the Atlantic Undersea Test and Evaluation Center (AUTEC) cannot be missed
- count group dives (over some time period T)
- account for dive rate (from Dtag data)
- account for estimate of mean group size
- similar example with sperm whales in Ward *et al.* (2012)

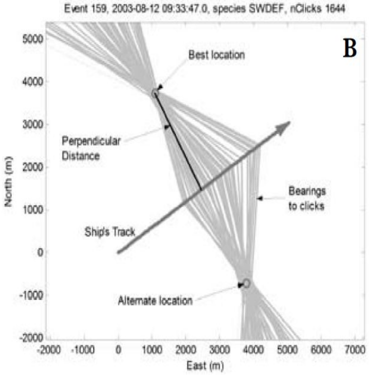
Census/total counts are great!

So if you can count all the animals/groups/sounds within a known area you are essentially there...

But really, who has access to a very dense array of sensors like AUTEC?

Let's go do some sampling shall we?

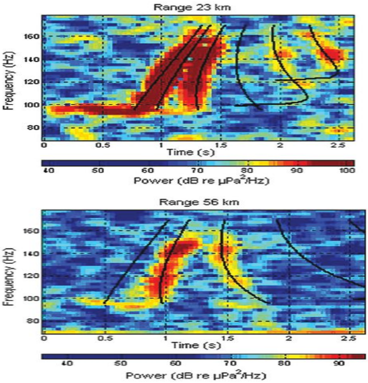
Conventional distance sampling with sperm whales



screenshot "stolen" from Lewis *et al.* 2007

- boat line transects, towed sensors, sperm whale detections
- multiple bearings intersection (+ depth) give you distance
- assumes individual animals or visual group size estimates
- conventional line transect distance sampling
- Hastie *et al.* 2003; Barlow & Taylor 2005;

Conventional distance sampling with right whales



screenshot "stolen" from Munger *et al.* 2011

- moored sensors
- detection of "up calls"
- cue counting, so needs cue rate
- normal mode separation to get distances from single sensor
- conventional point transect distance sampling
- terrible design (3 single sensors!)
- Marques *et al.* 2011;

Distances are good!

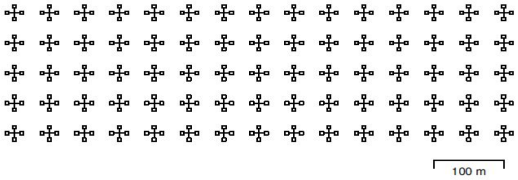
So if you can get distances to animals directly from your individual sensors (or closely spaced sets of sensors) you are really half way through to estimating density via conventional distance sampling methods.

Distances are good!

So if you can get distances to animals directly from your individual sensors (or closely spaced sets of sensors) you are really half way through to estimating density via conventional distance sampling methods.

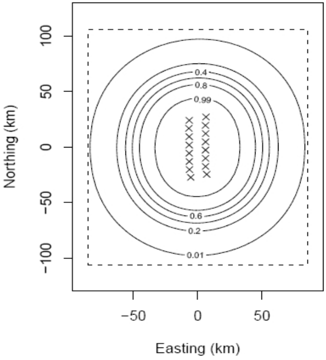
What about when you can not do that?

There are still some options out there...



screenshot "stolen" from Dawson & Efford 2009

SECR with minke whales



screenshot "stolen" from Marques *et al.* 2012

- minke boing detections at Pacific Missile Range Facility sensors
- manual validation = no false positives
- capture histories: association of detection of same sound over set of hydrophones
- spatially explicit capture recapture
- bypass explicit localization and use all data for density estimation
- no animal density (cue rates?)

Associations across hydrophones are useful!

So without distances, but if you can associate sounds across sensors, you can still get density estimates.

Associations across hydrophones are useful!

So without distances, but if you can associate sounds across sensors, you can still get density estimates.

But what if you don't even get that?

There are still some options out there, if you can estimate a detection function using auxiliary data or via sound propagation models

A cue counting approach

Animal density (D) can be estimated [1] based on the n_c detected clicks over time T and K hydrophones by (cue counting)

$$\hat{D} = \frac{n_c(1-\hat{c})}{K\pi\hat{P}w^2T\hat{r}}$$

where

- $\hat{r}$ - click rate
- $\hat{c}$ - false positive proportion
- w - maximum distance at which a click might be detected
- $\hat{P}$ - probability of detection of a click within w around the hydrophone

[1] $\hat{\theta}$ represents an estimator for θ

Intuitively...

$$D = \frac{n_c(1-c)}{k\pi w^2 P T r}$$

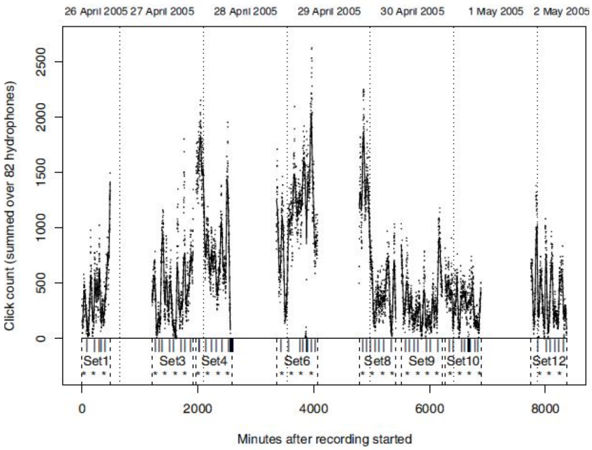
Feynman on Dirac: "to understand a physical problem means to be able to see the answer without solving equations"

$$\frac{\text{really beaked whale sounds}}{(\text{effective detection area}) \times (\text{clicks by animal during } T)}$$

Note that:

- $K \times \pi \times w^2 \times P = K \times \pi \times \rho^2$ is the effective detection area

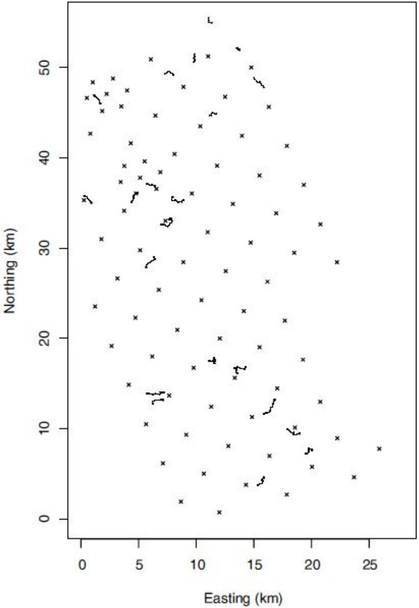
Detection function estimation with DTag beaked whales



screenshots "stolen" from Marques *et al.* 2009

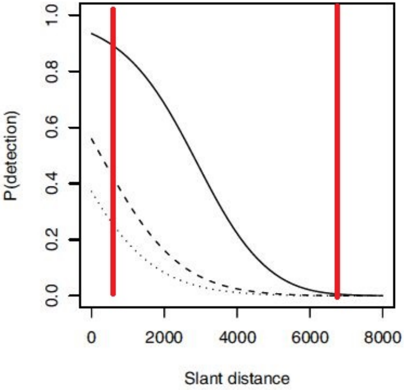
- Blainville's beaked whale click counts over AUTEC hydrophones
- false positives: manual validation over random sample of periods

Detection function estimation with DTag beaked whales

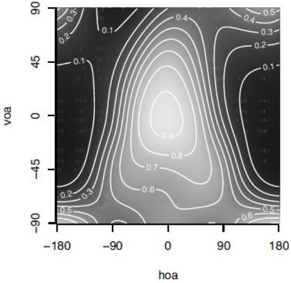


- DTag data with animals known location and orientation correlated with AUTEC hydrophone data

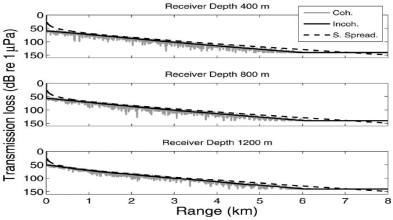
Detection function estimation with DTag beaked whales

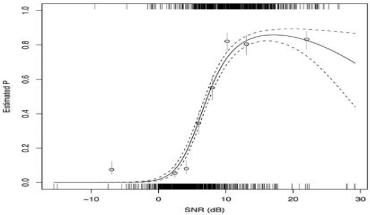


- regression model of detection as a function of covariates (distance, vertical and horizontal angle) leads to an estimate of probability of detection
- cue rate from DTag's



Modelling sound source attenuation with beaked whales





screenshot "stolen" from Küsel *et al.* 2009

- Blainville's beaked whale click counts over AUTEC hydrophones
- false positives from a manual validation exercise over random sample of periods
- model combines sound source level, ambient noise, detector characterization, sound propagation loss to estimate probability of detecting a click
- single hydrophone proof of concept approach
- Harris *et al.* in prep with blue whales

Finally...

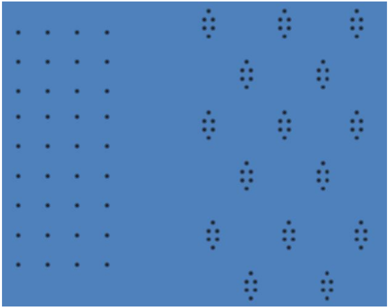
And so what?

Are there any general “take-home” messages out of all this?

So for density estimation really we *just* need...

A network of randomly or **systematically** distributed sensors capable of localizing or **ranging** animals, covering the space and time we are interested in estimating density for.

Lee’s neutrino detection sensor field



Movement and fixed passive acoustic density estimation

Wide temporal coverage and fixed sensors

Animal movement becomes a key issue

Individual or animal based methods

How to deal with movement? Essentially pretend it does not exist by considering snapshots... but?

Cue based methods

How to obtain reliable cue rates?

What is the question you are trying to answer?

If you set out to find something about everything, chances are you'll learn nothing about everything...

Take home messages

- **passive acoustics density estimation: young and yet promising field**

Take home messages

- passive acoustics density estimation: young and yet promising field
- **ideal for long term monitoring and detecting trends**

Take home messages

- passive acoustics density estimation: young and yet promising field
- ideal for long term monitoring and detecting trends
- **and (as for visual-based methods) distance sampling and capture recapture based methods look most promising**

Take home messages

- passive acoustics density estimation: young and yet promising field
- ideal for long term monitoring and detecting trends
- and (as for visual-based methods) distance sampling and capture recapture based methods look most promising
- **detection/classification is a fruitful research area. Nonetheless, given the reliable characterization of a system's performance, density estimation is possible using non-optimal systems**

Take home messages

- passive acoustics density estimation: young and yet promising field
- ideal for long term monitoring and detecting trends
- and (as for visual-based methods) distance sampling and capture recapture based methods look most promising
- detection/classification is a fruitful research area. Nonetheless, given the reliable characterization of a system's performance, density estimation is possible using non-optimal systems
- **required multipliers (often unavailable) should be obtained at the time and place of the main survey**

Take home messages

- passive acoustics density estimation: young and yet promising field
- ideal for long term monitoring and detecting trends
- and (as for visual-based methods) distance sampling and capture recapture based methods look most promising
- detection/classification is a fruitful research area. Nonetheless, given the reliable characterization of a system's performance, density estimation is possible using non-optimal systems
- required multipliers (often unavailable) should be obtained at the time and place of the main survey
- **development of hardware (cheap, reliable and ranging-capable sensors) and methods/software (e.g. ranging clicks from single sensors?) optimized for density estimation**

Take home messages

- passive acoustics density estimation: young and yet promising field
- ideal for long term monitoring and detecting trends
- and (as for visual-based methods) distance sampling and capture recapture based methods look most promising
- detection/classification is a fruitful research area. Nonetheless, given the reliable characterization of a system's performance, density estimation is possible using non-optimal systems
- required multipliers (often unavailable) should be obtained at the time and place of the main survey
- development of hardware (cheap, reliable and ranging-capable sensors) and methods/software (e.g. ranging clicks from single sensors?) optimized for density estimation
- yeah it's obvious yet... **methods have assumptions: plausibility and implications of their failure must be discussed**

Take home messages

- passive acoustics density estimation: young and yet promising field
- ideal for long term monitoring and detecting trends
- and (as for visual-based methods) distance sampling and capture recapture based methods look most promising
- detection/classification is a fruitful research area. Nonetheless, given the reliable characterization of a system's performance, density estimation is possible using non-optimal systems
- required multipliers (often unavailable) should be obtained at the time and place of the main survey
- development of hardware (cheap, reliable and ranging-capable sensors) and methods/software (e.g. ranging clicks from single sensors?) optimized for density estimation
- yeah it's obvious yet... methods have assumptions: plausibility and implications of their failure must be discussed
- **is there anything sexier than cetaceans OR neutrinos? maybe cetaceans AND neutrinos! How can collaborations grow?**

Questions?

Welcome now, or later via

tiago@mcs.st-and.ac.uk

Last, but not least

Thank You !