A Bayesian estimation of the number of Antarctic blue whales in the Austral Ocean using passive acoustic sensors

Valsero Blanco, María Cruz
Departamento de Estadística e Investigación Operativa
Universidad de Valladolid, C/Doctor Mergelina S/n
Valladolid (47005), Spain
E-mail: mcruz@eio.uva.es

Prieto González, Rocío
Departamento de Estadística e Investigación Operativa and
Centre de Neurosciences de Paris Sud, CNPS-CNRS UMR8195
E-mail: rocio.prieto-gonzalez@u-psud.fr

Samaran, Flore
Centre de Recherche des Mammiferes Marins
Université de La Rochelle
La Rochelle (17000), France
E-mail: flore.samaran@cebc.cnrs.fr

Adam, Olivier
Centre de Neurosciences de Paris Sud, CNPS-CNRS UMR8195
Université Paris 11, Orsay Cedex (91405), France
E-mail: olivier.adam@u-psud.fr

1. Summary

Knowing the size of the population of blue whales is very important from the point of view of marine biology. Blue whale is a species endangered since 1967 as its population has been dramatically reduced. Traditionally the estimation of the size of cetacean population has relied on visual surveys. However, visual observation is related to meteorological conditions and sometimes, in some areas, this is not possible. The addition of acoustic monitoring can help to asset the size of the population in an area. In this paper we analyzed the data of BMi calls (Antarctic blue whales) registered on one hydrophone located in the North of the Crozet islands which covers a recording surface of 47123 m^2. Data from May 2003 to April 2004 were studied and modified to obtain the number of calls per hour. The number of calls was modeled as a random process depending on the number of whales in the area and on the number of calls from a single whale in one hour. Supposing that the number of whales in the area is a discrete random variable with unknown distribution, a Poisson distribution was considered a priori and a model was developed to estimate the parameters by likelihood methods. Then, we calculated the distribution a posteriori and obtained a distribution of the number of whales in the area. R language was used to implement these methods.

2. Introduction

On this paper, a method of population assessment based on acoustic techniques is applied to the study of the abundance of Antarctic Blue whales using a hydrophone located off the Crozet

Archipelago, in the southern Indian Ocean. An area where so few blue whales have been seen that their density has not been previously estimated. F. Samaran showed in her PhD that Antarctic blue whales were present around Crozet island during the whole year. The question is how many individuals are there, and how the population size is changing from one year to another. This information is crucial to know if this species will disappear and in such a case when it would hapen. Cetaceans live at low density over large areas of ocean. All species of large cetaceans in the open ocean are currently considered endangered and are protected by law. Knowing the population size is fundamental from the point of view of marine biology/ecology because blue whale is a specie endangered since 1967. It is estimated that its population has been reduced to 0.15% of the initial population and nowadays there are about 3000 individuals in the Austral Ocean. However, cetacean species are hard to survey, since they spend almost all of their time underwater. Currently, the main method for obtaining estimates of density are based on Distance Sampling theory (DS) (Buckland et al. 2001). This method is based on visual line transect surveys, and recently has been applied to their vocalizations collected from fixed passive acoustic sensors. DS gives a punctual estimation of the density based on a probability of detection,that is calculated using the distances at which, one whale is seen or it is recorded. This method needs to know the value of the sensitive parameter, call rate. In the case of blue whale, it is very hard to know the call rate with precision. Maybe this call rate is changing a little bit from one individual to another and it is possible that this call rate is different from one month to another. Marques et al.(2009) used acoustic tag to know exactly the call rate of a single beaked whale individual. This approach it is not possible for Antarctic Blue whale because until now,no one has been able to find one individual to put one acoustic tag on his back. The object of this paper is to increase the repertoire of tools available for estimating species abundance. To date,it has been tried another two approach. The first tries to find an acoustic individual signature, but nobody knows how we can do that, the second is DS, mentioned above. We present a model for estimating the number of whales in the area based on the number of the recorded calls in the area.We suppose that the number of whales in the area is a discrete random variable with unknown distribution, and that the number of calls from a single whale is a Poisson variable. We supposed independence between calls of every whale in the area. To estimate the distribution of the number of whales, we introduced a Poisson distribution as a non-informative a priori distribution. Based on this distribution, a posteriori distribution was calculated and used to estimate the abundance of Antarctic blue whales. So, we present a new framework for estimating cetacean density from data obtained by fixed passive acoustic detectors. We think that it might be used under considerably different scenarios, with appropriate modifications that are also discussed.

3. Materials and methods

The data set was recorded from January to March 2004 at a station located in the South-western Indian Ocean (Crozet Islands - 4651'S-5153'E)(fig.1). The station is moored in the International Monitoring System (IMS) and support the Comprehensive Nuclear Test-Ban Treaty (CTBT). The station was designed to control nuclear tests in the ocean, but other sounds were found in the recordings, sounds that were analyzed to found that they were produced by large whales. Large whales spend much of their lives in pelagic environments, so it is not easy to estimate the abundance by traditional means. The Austral Ocean does not offer good climatic conditions to go and register the number of blue whales seen. So although the main objective was not register whales sounds, Blue whales calls were detected as a valious secondary product. There were two arrays of three instruments, each one located in the Northern and Southern coasts of Possession Island. The instruments were deployed on the seafloor at a depth between 1100 and 1500 meters in a triangular configuration (triad) with approximately 2 km spacing. The two arrays were located on opposite sides of the island and spaced 60 km apart. The hydrophones were suspended near the sound channel axis (SOFAR) at a depth of approximately 300 m. In this area of the ocean, the speed of propagation of sound waves is minimal, making that the ocean behaves at these depths as a wave guide. The hydrophones monitored sound continuously, 24 hours a day, 7 days a week. The acoustic data from each hydrophone were analyzed to check for the presence of calls typically associated with Antarctic and Pygmy Blue whales. An automatic detection method for both call types was designed (Samaran et al. 2008). This method was used to detect both blue whale calls with a Signal to- Noise Ratio (SNR) up to -15 dB in the bandwidth 17-30 Hz for BMi calls and bandwidth 17-50 Hz for BMb calls. We analyzed BMi calls (Antarctic blue whales) of only one hydrophone located in the North of the island which covers a recording surface of 47123 m^2. Recording began in may 2003 and finalized in April 2004. Although the hydrophone registered continuously, we have a gap of data due to technical incidents. We have used the R language to aggregate data to obtain the number of calls in every hour. After we have applied the model to these data to estimate the density of Antarctic blue whales in the area.

4. The model

Assessing the size of cetacean populations in the open ocean has traditionally relied on visual surveys alone. The addition of acoustic monitoring can complement these surveys if reliable protocols can be formulated and calibrated with visual techniques. Nowadays we do not known how to assign an individual acoustic signature to every whale. Distance Sampling is the only method for estimating the size of cetacean populations that has been adapted to the acoustic (Thomas et al. 2002) . This method use a punctual estimation of the density adding a probability of detection based on the distribution of observed detection distances.It require additional information like the probability of detecting calls and the rate at which animals produce it. To obtain estimation of this rate, (Marques et al. 2009) proposed to use specific tag including sensors attached on the back of the beaked whales. However, it is not possible to generalize this approach for different cases, especially when the individuals are not reachable. In the case presented, visual observation was not possible and the knowledge of these parameters neither, so a model is presented to estimate Antarctic blue whale population based on near-continuous recording from a single hydrophone as a new alternative method.

Let $N(t, s)$ be the number of calls in the surface s in a interval of time t the observed process
$B(s)$ the number of whales in the area s
$C(t)$ the number of registered calls for whale in the time interval t
We suppose that the unobserved variable $B(s)$ follows a discrete distribution with values from 0 to b with probabilities P_i and the unobserved process $C(t)$ is a Poisson process of intensity μ. The distribution of $N(t, s)$ is calculated by conventional methods.

$$P\big(N(t,s) = k\big) = \sum_{i=0}^{b} e^{-\mu it}\,\frac{(\mu it)^k}{k!}\,P_i$$

Our goal is to estimate the distribution of $B(s)$, by bayesian methods.
So, we introduce a Poisson distribution of intensity λ as a priori distribution of $B(s)$. And we develop

the distribution of the number of calls $N(t,s)$ in two etapes.

$$P\Big(N(t,s) = k/B(s) = i\Big) = P\big(C_1(t) + \ldots + C_i(t)\big) = e^{-\mu i t}\frac{(\mu i t)^k}{k!}$$

$$P\Big(N(t,s) = k\Big) = \sum_{i=0}^{\infty} e^{-\lambda s}\frac{(\lambda s)^i}{i!} e^{-\mu i t}\frac{(\mu i t)^k}{k!}$$

We estimate the value of the parameters λ et μ by the maximum likelihood method.
Given n observations, $N_1(t,s)$, $N_2(t,s)$, $\ldots$, $N_n(t,s)$

x_1, x_2, $\ldots$, x_n

$$\begin{aligned} L_N(\lambda,\mu) &= P\big(N_1(t,s) = x_1, \ldots, N_n(t,s) = x_n\big) \\ &= \prod_{j=1}^{n}\Big(\sum_{i=1}^{\infty} e^{-\lambda s}\frac{(\lambda s)^i}{i!} e^{-\mu i t}\frac{(\mu i t)^{x_j}}{x_j!}\Big) \end{aligned}$$

We find the estimators $\widehat{\lambda}_{MVE}$, $\widehat{\mu}_{MVE}$. In this way , we have a estimate distribution of the number of calls produced and we use it to estimate a posteriori distribution of the number the whales $B(s)$ $\forall$ t

$$P\Big(B(s) = i/N(t,s) = k\Big) = \frac{e^{-\mu i t}\frac{(\mu i t)^k}{k!}\frac{(\lambda s)^i}{i!}}{\sum \frac{(\lambda s)^j}{j!} e^{-\mu j t}\frac{(\mu j t)^k}{k!}} \propto e^{-\mu i t}\, i^k\, \frac{(\lambda s)^i}{i!}$$

Then, for every t observed,we take the i that maximize the posteriori probability distribution,

$$P\big(B(s) = i/N(t,s) = k\big) = \max_j \Big\{P\big(B(s) = j/N(t,s) = k\big)\Big\}$$

So, we have obtained a random sample of the unobserved variable $B(s)$, and we estimate its empirical distribution. $\widehat{P}_0, \widehat{P}_1, \ldots, \widehat{P}_b$.
The empirical distribution is used to find confidence interval for the distribution of $B(s)$.

5. Results and discussion

We have introduced a new model to estimate the abundance of Antarctic blue whales. The estimation found by the model are closed to the number expected by biologists. The model is a simple one and we hope that in future could be modified ,when we had more knowledge about the behavior of Antarctic blue whales. It is difficult to verify the hypothesis introduced. The independence of the number of calls produced by a whale is introduced by the observation of the behavior of other marines mammals. We will not find any publication to ensure the independence of the sounds of whales but neither one which denies it. The model introduced can be interesting for biologists,because until now, they do not have any knowledge about the way that this species emit its sounds.
Table 1 shows the confidence intervals for the number of whales.
Fig. 2 represents the histogram of the number of whales estimated.
Accurate estimation of global whale populations remains a difficult problem that must be addressed using complementary techniques to be successful. In this paper, we present for the first time the approach based on the distribution of the number of whales using passive acoustic methods. Other papers deals with the punctual estimation of the density (distance sampling). DS use advanced and expensive technical means,but some times it is not possible to use these technical means due to lack

of resources, the furtive behavior of the animals or geographical and meteorological conditions.
To date, no method has been proposed that might be used in general to estimate cetacean density from fixed passive acoustic devices, although several papers have addressed elements of the problem (e.g., Marques et Thomas, 2009; McDonald and Fox, 1999; Wiggins et al., 2004).
The mayor advantage of the model introduced is that it can be modified to get a good fit of the data. It can be used for estimating the density of other species, with appropriate modifications.

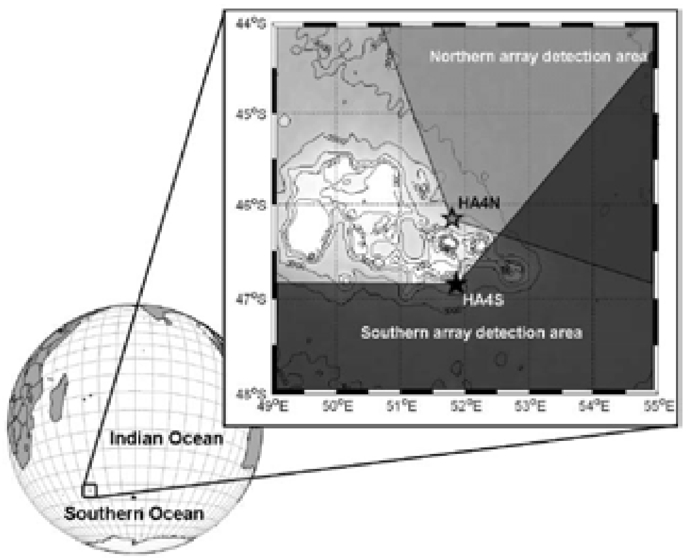

Figure 1: Location of hydrophones of IMS in Crozet Archipelago

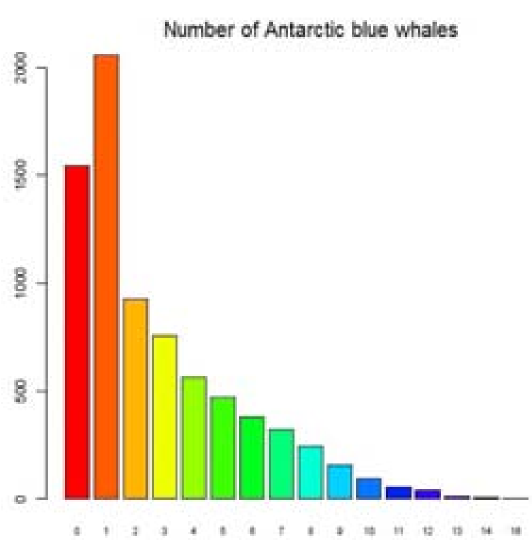

Figure 2: Number of Antarctic blue whales in the year

Table 1: $\text{CI}_{90\%}$ and $\text{CI}_{95\%}$ of the number of whales ($B(s)$)

CI	90%	95%
May	[1, 8]	[1, 9]
June	[1, 8]	[0, 9]
July	[0, 8]	[0, 10]
August	[0, 9]	[0, 10]
September	[0, 9]	[0, 9]
October	[0, 8]	[0, 8]
November	[0, 9]	[0, 10]
December	[0, 10]	[0, 11]
January	[0, 7]	[0, 9]
February	[0, 9]	[0, 9]
March	[0, 2]	[0, 3]
April	[0, 6]	[0, 7]
Year	[0, 9]	[0, 10]

References

[1] Thomas, L., Buckland, S.T., Burnham, K.P., Anderson, D.R., Laake, J.L., Borchers, D.L. & Samantha Strindberg (2002). *Distance Sampling*. Encyclopedia of Environmetrics (ISBN 0471 899976) Abdel H. El-Shaarawi and Walter W. Piegorsch 544 - 552

[2] Buckland, S.T., Anderson, D.R., Burnham, K.P. & Laake, J.L. (1993). *Distance Sampling: Estimating Abundance of Biological Populations*. Chapman & Hall, London, reprinted (1999) by Research Unit for Wildlife Population Assessment, St Andrews.

[3] Buckland, S.T., Anderson, D.R., Burnham, K.P., Laake, J.L., Borchers, D.L. & Thomas, L. (2001). *Introduction to Distance Sampling*. Oxford University Press, Oxford.

[4] Samaran F., Adam O., Motsch J.-F. & C. Guinet, (2008). *Definition of the Antarctic and Pygmy Blue whale call templates. Application to fast automatic detection*. Canadian Acoustic, 36, 93 - 102.

[5] McDonald M.A., Mesnick S.L. & Hildebrand J.A. (2006). *Biogeographic characterization of blue whale song worldwide: using song to identify populations*. Press, Oxford. JCRM. 8, 55 - 65.

[6] Sirovic? A., Hildebrand J.A., Wiggins S.M., McDonald M.A., Moore S.E. & Thiele D. (2004). *Seasonality of blue and fin whale calls and the influence of sea ice in the Western Antarctic Peninsula*. DSR II, 51, 2327 - 2344

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

[8] M.A. McDonald, C.G. Fox (1999). *Passive acoustic methods applied to fin whale population density estimation*. J. Acoust. Soc. Am. Volume 105, 2643-2651

[9] H.M. Taylor, S. Karlin (1998). *An introduction to Stochastics Modeling*. Academic Press, third ediction.

[10] N.L. Johnson, S. Kotz & N. Balakrishnan (1997). *Discrete Multivariate Distributions*. Wiley-Interscience.