Supplementary material

Food talk: 40-Hz fin whale calls are associated with prey biomass

Miriam Romagosa*[1], Sergi Pérez-Jorge[1], Irma Cascão[1], Helena Mouriño[2], Patrick Lehodey[3], Andreia Pereira[4], Tiago A. Marques[5,6], Luís Matias[4] and Mónica A. Silva[1]

1. Okeanos – Instituto de Investigação em Ciências do Mar, Universidade dos Açores & IMAR – Instituto do Mar, Horta, Portugal
2. Centro de Matemática, Aplicações Fundamentais e Investigação Operacional, Faculdade de Ciências, Universidade de Lisboa, Lisboa, Portugal
3. Collecte Localisation Satellite (CLS), Ramonville St Agne, France
4. Instituto Dom Luiz (IDL), Universidade de Lisboa, Lisboa, Portugal;
5. Centro de Estatística e Aplicações, Departamento de Biologia, Faculdade de Ciências, Universidade de Lisboa, Lisboa, Portugal
6. Centre for Research into Ecological and Environmental Modelling, University of St. Andrews, St. Andrews, UK

1. METHODS

(a) Acoustic data collection and analyses

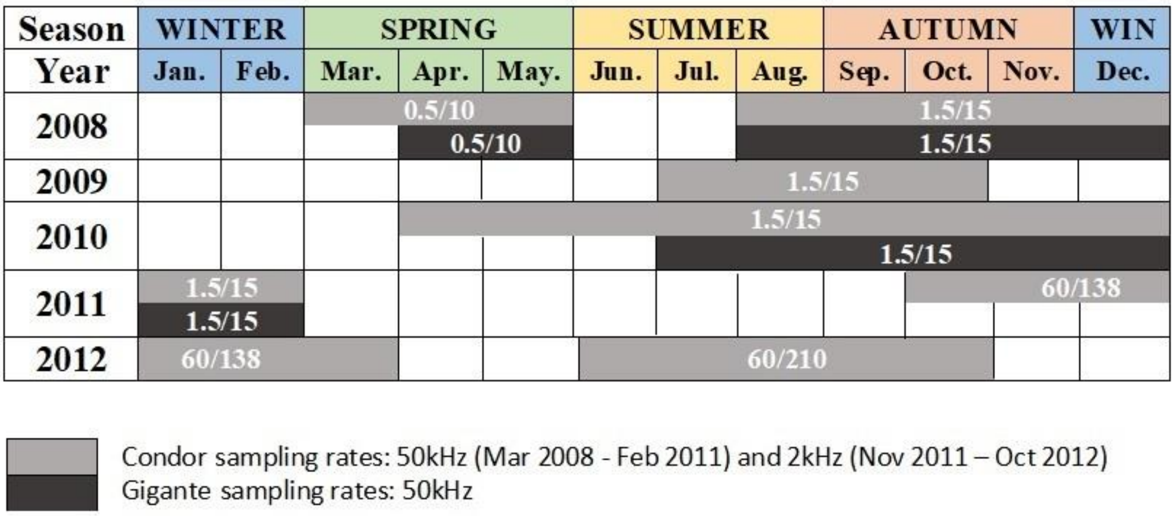

Figure S1. Deployment times, duty cycles and sampling rates for each season, month and location. Duty cycles for each deployment are in white numbers and indicate recording periods (min on/every x min). Gaps in the time series were caused by maintenance duties and equipment failure.

(b) Spatial scale of data integration

<u>Detection range estimation</u>

The maximum range at which the two types of fin whale vocalizations could be detected was estimated theoretically using the sonar equation and other mathematical models. For passive sonar, signal-to-noise ratio SNR (dB) of a signal is defined as [1]:

$$SNR = SL - TL - NL + 10\log_{10} BW \qquad \text{Eq. (1)}$$

where SL is the transmitted source level (dB rms re 1 µPa at 1 m), TL is one-way transmission loss (dB), NL is the ambient noise level at the receiver (dB rms re 1 µPa), and BW is the processing bandwidth (Hz). Knowing SL, NL, BW and SNR, TL can be calculated using Eq. (1). By matching the calculated TL in the TL model obtained for the study area, we can then obtain a theoretical maximum range at which a fin whale calls could be detected.

The Range-dependent Acoustic Model (RAM)[2,3] is a parabolic equation model that was used to model the propagation loss estimates along the range-depth plane from source to receiver. We developed propagation loss models based on the frequency of each fin whale vocalization along 12 bathymetric transects from each EAR position, obtained every 30º measured from North. Since seasonal changes can cause differences in the water column and affect propagation conditions, we used sound speed profiles for winter and summer months (January and June for Condor and January and September for Gigante). In total, we developed 96 propagation loss models up to 180 km, with a 50 m range step, 24 for each fin whale call in each area. Additional information and parameters for the propagation loss modelling are described below:

Frequency and depth of the vocalizing whale: We used the centre frequency of a sample of fin whale calls recorded around the Azores area. The centre frequency was estimated to be of 25 Hz for the 20-Hz call and 62 Hz for the 40-Hz call. The depth of the vocalizing whale was estimated to be 50 m and was obtained from the literature [4].

Depth of the hydrophones: The two EARS were deployed in Condor and Gigante at similar depths (190 m).

Sound speed profile: The sound speed profiles for January, June and September were extracted from the Levitus climatological database [5]. The profiles were obtained for one point between Condor and Gigante.

Ocean bottom composition: Since there were no direct measurements of sediment properties in the two areas, we used an average sound speed in the sediments of 1700 m/s and a seabed density of 1500 kg/m^3.

Bathymetry along the transmission path: The bathymetric relief of transects around the EARS were obtained from the European Marine Observation and Data Network (EMODnet) (https://portal.emodnet-bathymetry.eu/).

For the TL calculation using Eq. (1) we also used the following data:

Source level: We calculated average source level estimates for a sample of fin whale 20-Hz (n = 139) and 40-Hz (n = 42) calls recorded in 3 EARS deployed off Faial-Pico Island, close to the study area, and located using Time-of-Arrival-Differences (TOADs) [6]. Source level estimates were calculated using the "inband power" feature in RAVEN PRO 2.0 (Bioacoustics Research Program) and root-mean-squared (RMS) received levels were extracted in the measured bandwidth of the vocalizations. Then we used the passive sonar equation by adding the received levels of the transmitted signal in the EARS and the associated transmission loss. The average estimated source levels for the 20-Hz call was 147.4 dB rms re 1 µPa at 1 m (± 15.5) and for the 40-Hz call was 144.3 dB rms re 1 µPa at 1 m (± 3.6).

Ambient noise level: Received ambient noise levels for the noisiest and quietest month, previously identified in this dataset[7], were calculated in 1/3 octave bands centred at the target frequencies for each fin whale call: centred at 20 Hz (14.15 - 28.3 Hz) and 62 Hz (44 - 88 Hz). Measurements were made using PAMGuide [8] by entering the manufacturer's specifications for the end-to-end sensitivity of the instrument (−193.14/-194.17 re 1 V/µPa; depending on deployment), a gain of 47.5 dB and a 0-peak voltage of 1.25 V.

Signal-to-noise ratio (SNR): Since the automatic detection process for the 20-Hz fin whale call was based on a detection threshold of 10 dB, we used this value to indicate the detectability of this call type and conservatively assumed the same SNR for manually detected 40-Hz calls.

(c) Statistical analyses

Model information

Let X_{it} be a random variable that represents the *i-th* call type (corresponding to 20-Hz song call and 40-Hz call, respectively) at week t, $t = 1, \ldots, 144.$ Assuming that $E\left(X_{it}\right) = \mu_{it}$, the basic log link function that describes the relationship between the mean and the explanatory variables takes the form:

$$\log(\mu_{it}) = \alpha_{0i} + \alpha_{1i}\ \text{zoo\_biomass}_{it} + \alpha_{2i}\ \text{season\_aut}_{it} + \alpha_{3i}\ \text{zoo\_biomass}_{it} \times \text{season\_aut}_{it} + $$
$$+\alpha_{4i}\ \text{zoo\_biomass}_{it} \times \text{season\_spr}_{it} + \sum_{j=1}^{4} \gamma_{ji}\ year_{ijt}, \quad i = 1, 2;\ t = 1, \ldots, 144.$$

where μ_{it} represents the mean of the Poisson distribution for the *i-th* call type ($i = 1, 2,$ corresponding to 20-Hz and 40-Hz calls, respectively) at week t, $t = 1, \ldots, 144.$ The variable $zoo\_biomass_{it}$ describes zooplankton biomass at week t, for the *i-th* call type; the variables $season\_aut_{it}$ and $season\_spr_{it}$ are the season indicators, for the *i-th* call type: $season\_aut_{it}$ is equal to 1 if week t belongs to autumn, and 0 otherwise; $season\_spr_{it}$ is equal to 1 if week t belongs to spring, and 0 otherwise; the reference category is the winter season. The interaction terms between zooplankton biomass and the season of the year were also included in the model. The dummy variable $year_{ijt}$ is the year indicator, equal to 1 if week t belongs to year j ($j = 1, 2, 3, 4,$ corresponding to year 2008, 2009, 2010, 2011, respectively); the reference category is the year 2012. The vector of the parameters for the *i*-th model is given by

$\boldsymbol{\theta} = \left(\alpha_{0i} \quad \alpha_{1i} \quad \alpha_{2i} \quad \alpha_{3i} \quad \alpha_{4i} \quad \gamma_{1i} \quad \gamma_{2i} \quad \gamma_{3i} \quad \gamma_{4i}\right)^T$, where α_{0i} is the intercept, $i = 1, 2.$

For the quasi-Poisson model, it is assumed that the variance is equal to the mean multiplied by a dispersion parameter (denoted by ϕ, $\phi \neq 1$), that is, $Var\left(X_{it}\right) = \phi\ \mu_{it}$. In this work, we considered $\phi > 1$ because the datasets were overdispersed. The parameters of the model are obtained by maximum quasi-likelihood estimation, where only the relationship between the mean value and the variance is specified. Thus, there is no need to establish the form of the underlying probability distribution.

Model fitting relies upon the quasi-Poisson models, which means that we are working in the quasi-likelihood framework. Therefore, we used the Quasi-Akaike's Information Criterion (QAIC) instead of the well-known AIC, which is given by

$$\mathrm{QAIC} = 2k - \frac{2\ln\left(\hat{L}\right)}{\hat{\phi}},$$

where k is the number of parameters in the model; $\hat{L}$ represents the quasi-likelihood function evaluated at the maximum quasi-likelihood estimators; $\hat{\phi}$ is the estimate of the variance inflation factor that accommodates overdispersion.

2. RESULTS

(a) Detection range and zooplankton biomass spatial scale

Table S1. Summary table showing averaged source levels, ambient noise levels, transmission loss and average detection ranges for each vocalization type and noise conditions. SL –Source levels, BW – Bandwidth, NL-Noise levels and TL-Transmission loss.

<table>
<tr><th>Type</th><th>SL
(dB re 1 μPa at 1 m)</th><th>BW</th><th>Location</th><th>Month</th><th>Description</th><th>NL (RMS)
(dB re 1 μPa)</th><th>TL (dB)</th><th>Average detection range (km)</th></tr>
<tr><td rowspan="4">20-Hz note</td><td rowspan="4">147.3</td><td rowspan="4">14</td><td rowspan="2">Condor</td><td>Jan-12</td><td>Quietest</td><td>69.7</td><td>79.1</td><td>70</td></tr>
<tr><td>Jun-12</td><td>Noisiest</td><td>71</td><td>77.8</td><td>58</td></tr>
<tr><td rowspan="2">Gigante</td><td>Jan-11</td><td>Quietest</td><td>62.2</td><td>86.6</td><td>149</td></tr>
<tr><td>Sep-10</td><td>Noisiest</td><td>77.4</td><td>71.4</td><td>20</td></tr>
<tr><td rowspan="4">40-Hz call</td><td rowspan="4">144.3</td><td rowspan="4">44</td><td rowspan="2">Condor</td><td>Jan-12</td><td>Quietest</td><td>78.7</td><td>72</td><td>26</td></tr>
<tr><td>Jun-12</td><td>Noisiest</td><td>82.4</td><td>68.3</td><td>11</td></tr>
<tr><td rowspan="2">Gigante</td><td>Jan-11</td><td>Quietest</td><td>77.7</td><td>73</td><td>34</td></tr>
<tr><td>Sep-10</td><td>Noisiest</td><td>83.5</td><td>67.2</td><td>11</td></tr>
</table>

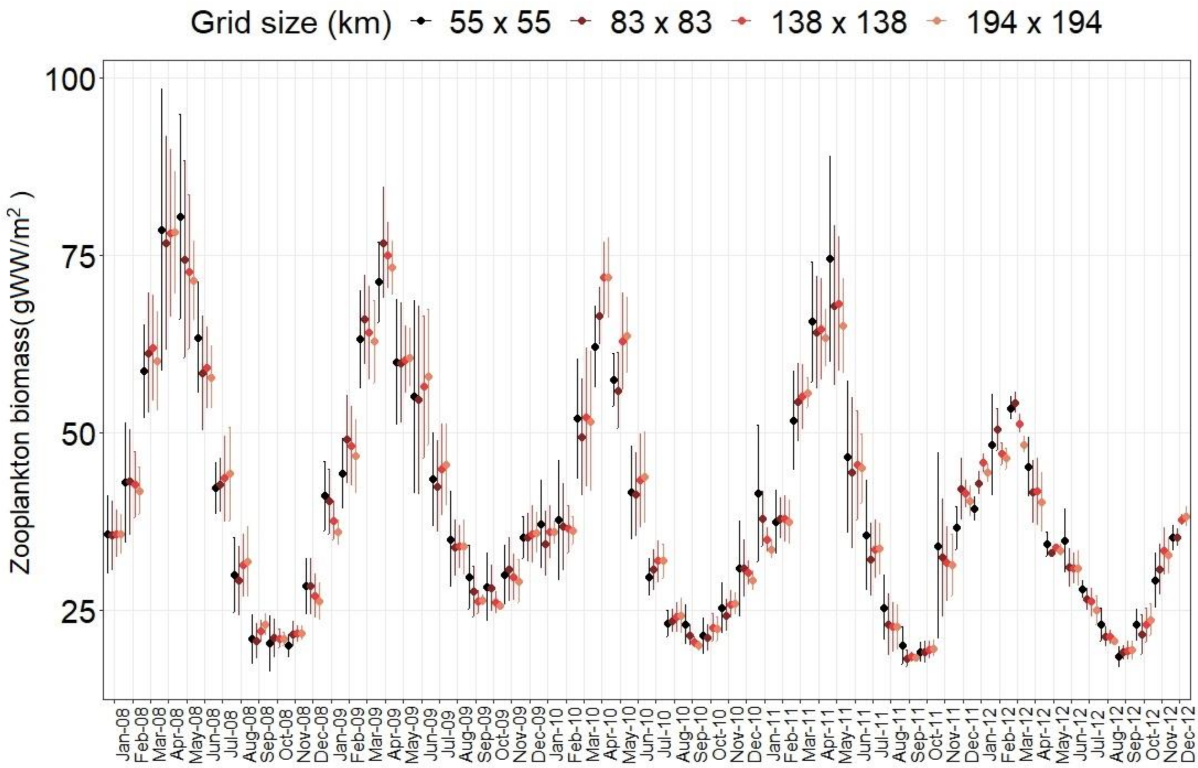

Figure S2. Monthly averaged modelled zooplankton biomass by grid size. Points represent averaged values and error bars represent standard deviations.

(a) Models of the 20-Hz and 40-Hz call

Table S2. Model selection results for the 20-Hz and 40-Hz call ordered by lowest QAIC. Best model is shown in bold.

ID	Model	Number of parameters	QAIC	Δ(QAIC)	Weight (QAIC)
20-Hz call					
1	**season + year**	**7**	**186.8**	**0**	**0.69**
2	zoo+season+year	8	188.8	1.9	0.25
3	zoo+season+year+zoo×season	10	191.8	5.0	0.05
4	zoo+season	4	224.3	37.4	0
5	zoo+season+zoo × season	6	226.3	39.5	0
6	season	3	230.0	43.1	0
7	zoo+year	6	265.0	78.1	0
8	year	5	283.1	96.3	0
9	zoo	2	288.9	102.1	0
40-Hz call					
1	**zoo**	**2**	**176.0**	**0**	**0.57**
2	zoo+year	6	178.0	2	0.21
3	zoo+season	4	179.5	3.5	0.1
4	zoo+season+year	8	181	4.9	0.05

5	zoo+season+zoo×season	6	182.3	6.3	0.02
6	season+year	7	182.4	6.4	0.02
7	zoo+season+zoo×season	10	183.1	7.2	0.02
8	season	3	184.8	8.8	0.01
9	year	5	193.2	17.3	0

QAIC: Quasi-Akaike Information Criteria

Δ(QAIC) = QAIC of the current model – QAIC of the best model (i.e., the model with the lowest QAIC)

Weight(QAIC): is the relative likelihood of the current model, when compared to the other models under consideration, and can be obtained by

$$\mathrm{Weight}(\mathrm{QAIC}_i) = \frac{\exp(-0.5\Delta(\mathrm{QAIC}_i))}{\sum_{j=1}^{k}\exp(-0.5\Delta(\mathrm{QAIC}_k))}$$

, where QAIC_i is the QAIC of model *i* ($i = 1, \ldots, k$), *k* is the number of models under analysis.

Table S3. Estimation results from the quasi-Poisson best selected models for the 20-Hz and 40-Hz call: point and interval estimates for each parameter and respective estimate of the standard error; test statistic and *p*-value from the Wald test.

	Estimate	**Std. error**	**t-value**	**P-value**	**95% CI**
20-Hz call					
Intercept	-0615	0.531	-1.159	0.248	(-1.748, 0.348)
Season_spr	Reference				
Season_aut	1.471	0.362	4.057	**<0.001**	(0.816, 2.254)
Season_win	2.379	0.381	6.237	**<0.001**	(1.680, 3.192)
Year2008	1.816	0.424	4.281	**<0.001**	(1.056, 2.745)
Year2009	0.770	0.659	1.168	0.24	(-0.637, 2.033)
Year2010	1.323	0.422	3.132	**<0.01**	(1.056, 2.745)
Year2011	1.250	0.415	3.009	**<0.01**	(0.567, 2.250)
Year 2012	Reference				
20rate_lag_1	0.012	0.002	4.834	**<0.001**	(0.008, 0.018)
Dispersion parameter	10.11	0.156*	—	—	(6.549,13.817)
40-Hz call					
Intercept	-2.076	0.322	-6.441	**<0.001**	(-2.735, -1.467)
Zooplankton	0.029	0.005	5.641	**<0.001**	(0.019, 0.039)
Dispersion parameter	1.527	0.024	—	—	(1.052, 2.160)

* Results obtained by parametric bootstrap based on 1000 replications

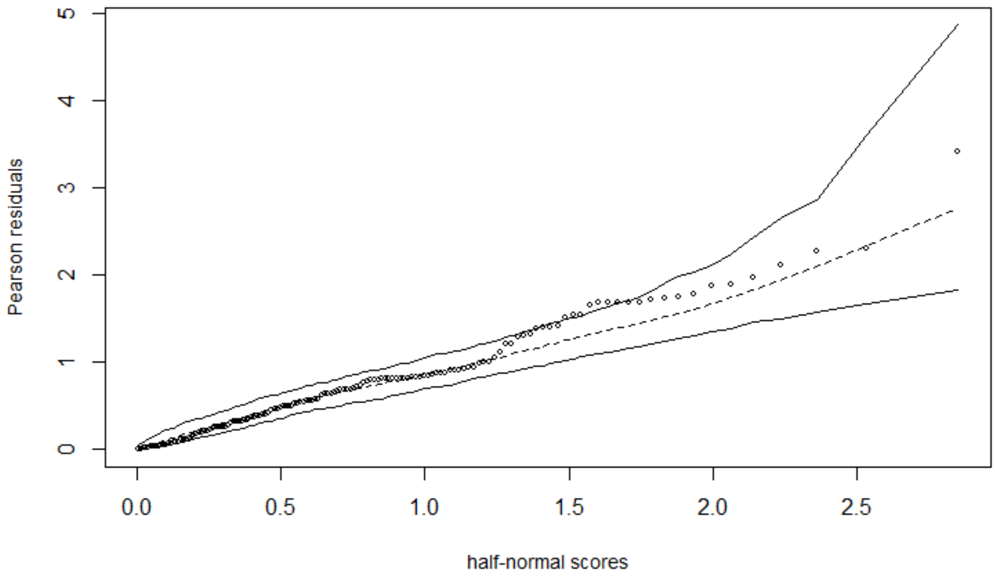

Figure S3. Quasi-Poisson fitted to the 20-Hz call rate: half-normal plot of the Pearson residuals, with simulation envelope based on 1000 runs. The scatter points (represented by circles) correspond to the ordered absolute values of the Pearson residuals *versus* the expected order statistics of the half-normal distribution. Solid lines indicate the 99% limits of the simulated envelope.

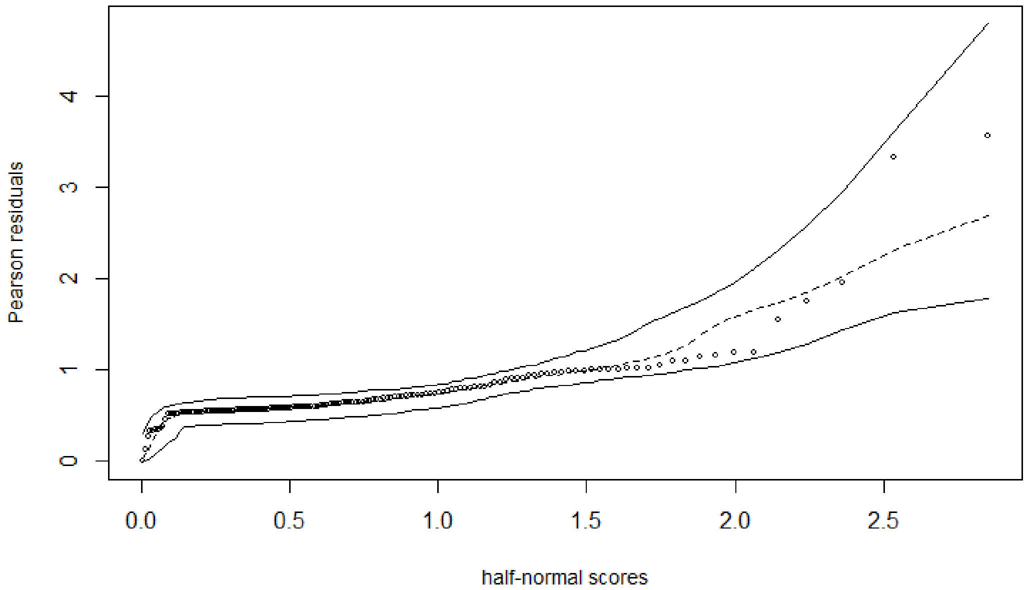

Figure S4. Quasi-Poisson fitted to the 40-Hz call rate: half-normal plot of the Pearson residuals, with simulation envelope based on 1000 runs. The scatter points (represented by circles) correspond to the ordered absolute values of the Pearson residuals *versus* the expected order statistics of the half-normal distribution. Solid lines indicate the 99% limits of the simulated envelope. Model residuals reveal that the model is adequate to describe the data and does not show the existence of outliers, with the residuals placed inside the envelope.

REFERENCES

1. Lurton X. 2002 *An Introduction to Underwater Acoustics: Principles and Applications*. Praxis publishing; UK.

2. Collins MD. 1993 A split-step Padé solution for the parabolic equation method. *J. Acoust. Soc. Am.* **93**, 1736–1742.

3. Collins MD. 1995 User's guide for RAM versions 1.0 and 1.0p.

4. Watkins WA. 1987 The 20-Hz signals of finback whales (Balaenoptera physalus). *J. Acoust. Soc. Am.* **82**, 1901–1912. (doi:https://doi.org/10.1121/1.395685)

5. Levitus S. 1983 Climatological Atlas of the World Ocean. *Eos, Trans. Am. Geophys. Union* **64**, 962–963. (doi:https://doi.org/10.1029/EO064i049p00962-02)

6. Baumgartner MF, Freitag L, Partan J, Ball keenan R, Prada KE. 2008 Tracking large marine predators in three dimensions: The real-time acoustic tracking system. *IEEE J. Ocean. Eng.* **33**, 146–157. (doi:10.1109/JOE.2007.912496)

7. Romagosa M, Cascão I, Merchant ND, Lammers MO, Giacomello E, Marques TA, Silva MA. 2017 Underwater Ambient Noise in a Baleen Whale Migratory Habitat Off the Azores. *Front. Mar. Sci.* **4**, 1–14. (doi:10.3389/fmars.2017.00109)

8. Merchant ND, Fristrup KM, Johnson MP, Tyack PL, Matthew J, Blondel P, Parks SE. 2015 Measuring acoustic habitats. *Methods Ecol. Evol.* (doi:10.1111/2041-210X.12330)