Supplementary material

Using integrated multispecies occupancy models to map co-occurrence between bottlenose dolphins and fisheries in the Gulf of Lion, French Mediterranean Sea.

Data overview and integrated multispecies occupancy model description

Valentin Lauret[1], Hélène Labach[2], Léa David[3], Matthieu Authier[4,5], Olivier Gimenez[1]

(1) CEFE, Univ Montpellier, CNRS, EPHE, IRD, Montpellier, France
(2) MIRACETI, Connaissance et conservation des cétacés, Place des traceurs de pierres, 13500 La Couronne, France
(3) EcoOcéan Institut, 18 rue des Hospices, 34090 Montpellier
(4) Observatoire Pelagis, UAR 3462, CNRS-La Rochelle Université, 17000 La Rochelle, France
(5) Adera, 162 Avenue Albert Schweitzer. CS 60040 – 33608 Pessac CEDEX, France

---

Data and codes are available on Github.

Data

We displayed transects and detection of each monitoring programs in Figure 1, and spatial variation of our environmental covariate depth in Figure 2.

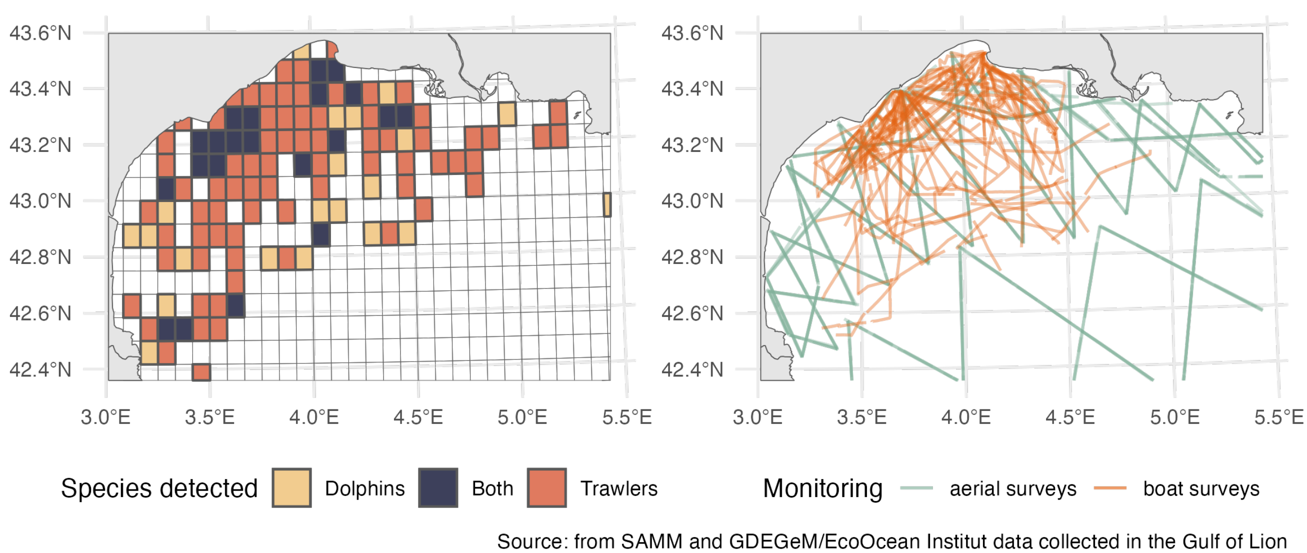

Figure 1: Gulf of Lion detections of bottlenose dolphins and trawlers by aerial surveys (SAMM) and boat surveys (GDEGeM and EcoOcean Institut) along with the sampling effort for each monitoring program. We plotted data on 397 5' × 5' contiguous Mardsen grid-cells (WGS 84)

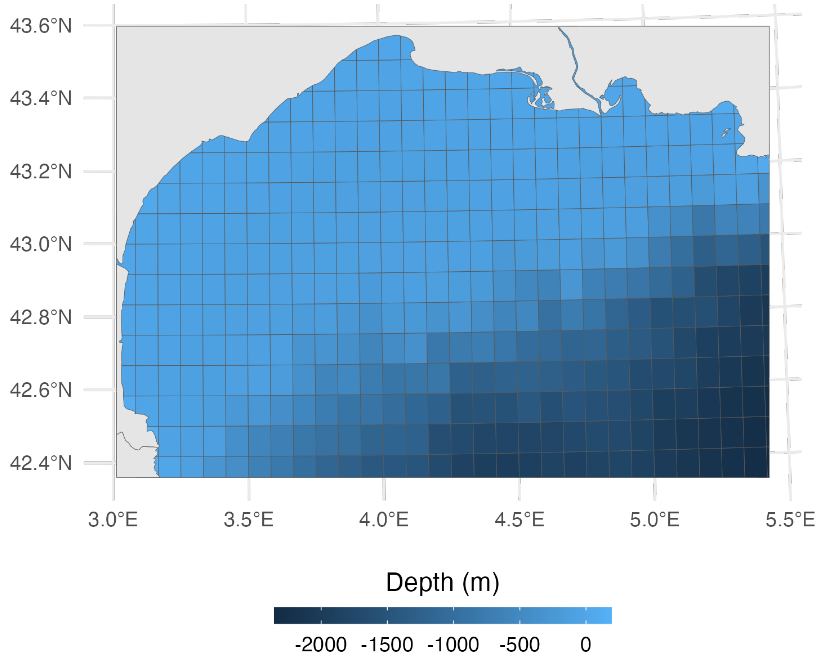

Figure 2: Spatial variation in depth over the Gulf of Lion study area

Multispecies occupancy models

We go through the methodology of multispecies occupancy modeling (Rota et al. 2016). Several assumptions need to be valid to safely apply multispecies occupancy models (similar assumptions to those of the single-species occupancy) : i) geographic and demographic closure, ii) independence of the detections over space and time, iii) accurate identification (i.e. no misidentification). In our case study, dolphins and trawlers obviously moved in and out grid-cells during the sampling period making the geographic closure unlikely to be respected. Thus, we interpret the occupancy as "space-use", that is the probability that the species uses the grid-cell given it is present in the study area, hence reflecting the usage a species makes of the different components of the study area.

Latent ecological process

We have four ecological states. Ignoring the site index, we use the following notation for four the occupancy probabilities:

- ψ_{11} is the prob. that species D and species T are both present;
- ψ_{10} is the prob. that species D is present and species T is absent;
- ψ_{01} is the prob. that species D is absent and species T is present;
- ψ_{00} is the prob. that species D and species T are both absent,

with $\psi_{11} + \psi_{10} + \psi_{01} + \psi_{00} = 1$.

The occupancy state of each grid-cell z is modeled as a multinomial draw in vector $\pi = (\psi_0, \psi_1, \psi_2, \psi_3)$ as:

$$z \sim Multinomial(1, \pi)$$

Then, the marginal probabilities of occupancy are:

- $\Pr(z_A = 1) = \Pr(\text{species D is present}) = \psi_{10} + \psi_{11}$
- $\Pr(z_B = 1) = \Pr(\text{species T is present}) = \psi_{01} + \psi_{11}$
- $\Pr(z_A = 0) = \Pr(\text{species D is absent}) = \psi_{01} + \psi_{00}$
- $\Pr(z_B = 0) = \Pr(\text{species T is absent}) = \psi_{10} + \psi_{00}$

And the conditional probabilities (reminder: $\Pr(\text{A}|\text{B}) = \Pr(\text{A and B}) / \Pr(\text{B})$):

- $\Pr(z_A = 1 | z_B = 0) = \psi_{10}/(\psi_{10} + \psi_{00}) = \Pr(\text{species D is present given species T is absent})$;
- $\Pr(z_A = 1 | z_B = 1) = \psi_{11}/(\psi_{11} + \psi_{01}) = \Pr(\text{species D is present given species T is present})$;
- $\Pr(z_B = 1 | z_A = 0) = \psi_{01}/(\psi_{01} + \psi_{00}) = \Pr(\text{species T is present given species D is absent})$;
- $\Pr(z_B = 1 | z_A = 1) = \psi_{11}/(\psi_{11} + \psi_{10}) = \Pr(\text{species T is present given species D is present})$.

To include covariates to the occupancy probabilities, we used the natural parameters formulation. For $k \in \{1, 2, 3\}$

$$\delta^k = \alpha_0^k + \alpha_1^k \mathrm{X}^k$$

where X^k is a covariate affecting δ^k. Quantities α_0^k and α_1^k were to be estimated. δ^1, δ^2 are called first-order parameters estimating log odds of species occurrence, conditional on absence of the other species, while δ^3 is a second-order parameter estimating change in log odds when both species are co-occurring. We obtained ψ:

$$\begin{aligned}
\psi^{00} &= \frac{1}{1 + \exp(\delta^1) + \exp(\delta^2) + \exp(\delta^1 + \delta^2 + \delta^3)} \\
\psi^{10} &= \frac{\exp(\delta^1)}{1 + \exp(\delta^1) + \exp(\delta^2) + \exp(\delta^1 + \delta^2 + \delta^3)} \\
\psi^{01} &= \frac{\exp(\delta^2)}{1 + \exp(\delta^1) + \exp(\delta^2) + \exp(\delta^1 + \delta^2 + \delta^3)} \\
\psi^{11} &= \frac{\exp(\delta^1 + \delta^1 + \delta^3)}{1 + \exp(\delta^1) + \exp(\delta^2) + \exp(\delta^1 + \delta^2 + \delta^3)}
\end{aligned}$$

Observation process

We consider dataset A (e.g SAMM aerial line transects), and dataset B (e.g. photo-id boat search-encounter program). Both monitoring programs collected detection / non-detection data about species D (i.e. bottlenose dolphin) and T (i.e. trawlers). Each species has a different detection probability depending on the monitoring program considered. For example, p_D^A is the probability of detecting deolphins by aerial monitoring program. Then, 16 observation 'events' can occur. We coded them as follow:

- `1` for none species detected neither by B nor A
- `2` for species D detected by B, nothing by A
- `3` for species T detected by B, nothing by A
- `4` for both species detected by B, nothing by A
- `5` for none species detected neither by B, species D detected by A
- `6` for species D detected by B, species D detected by A
- `7` for species T detected by B, species D detected by A
- `8` for both species detected by B, species D detected by A
- `9` for none species detected neither by B, species T detected by A

- `10` for species D detected by B, species T detected by A
- `11` for species T detected by B, species T detected by A
- `12` for both species detected by B, species T detected by A
- `13` for none species detected neither by B, both species detected by A
- `14` for species D detected by B, both species detected by A
- `15` for species T detected by B, both species detected by A
- `16` for both species detected by B, both species detected by A

From the 4 ecological states (in rows) and the 16 observation events (in columns), we get the observation process with the following (transposed) 4x16 matrix.

$$t(\theta) = \begin{bmatrix}
1 & (1-p_D^B)(1-p_D^A) & (1-p_T^B)(1-p_T^A) & (1-p_T^B)(1-p_T^A)(1-p_D^A)(1-p_D^B) \\
0 & p_D^B(1-p_D^A) & 0 & (1-p_T^A)(1-p_D^A)p_D^B(1-p_T^B) \\
0 & 0 & p_T^B(1-p_T^A) & (1-p_T^A)(1-p_D^A)p_T^B(1-p_D^B) \\
0 & 0 & 0 & (1-p_T^A)(1-p_D^A)p_D^B p_T^B \\
0 & p_D^A(1-p_D^B) & 0 & p_D^A(1-p_T^A)(1-p_D^B)(1-p_T^B) \\
0 & p_D^B p_D^A & 0 & p_D^A(1-p_T^A)p_D^B(1-p_T^B) \\
0 & 0 & 0 & p_D^A(1-p_T^A)p_T^B(1-p_D^B) \\
0 & 0 & 0 & p_D^A(1-p_T^A)p_D^B p_T^B \\
0 & 0 & 0 & p_T^A(1-p_D^A)(1-p_D^B)(1-p_T^B) \\
0 & 0 & p_T^B p_T^A & p_T^A(1-p_D^A)p_D^B(1-p_T^B) \\
0 & 0 & 0 & p_T^A(1-p_D^A)p_T^B(1-p_D^B) \\
0 & 0 & 0 & p_T^A(1-p_D^A)p_T^B p_D^B \\
0 & 0 & 0 & p_D^A p_T^A(1-p_T^B)(1-p_D^B) \\
0 & 0 & 0 & p_D^A p_T^A p_D^B(1-p_T^B) \\
0 & 0 & 0 & p_D^A p_T^A p_T^B(1-p_D^B) \\
0 & 0 & 0 & p_T^B p_T^A p_D^B p_D^A
\end{bmatrix}$$

Each observation y is linked to the ecological state z via a Categorical distribution in vector of length 16, $\theta_z = (Pr(y=1), Pr(y=2), ..., Pr(y=16))$, where θ_z also corresponds in the θ matrix to the 16 columns of the row refering to the ecological state z.

$$y|z \sim \text{Categorical}(\theta_z)$$

References

Rota CT, Wikle CK, Kays RW, Forrester TD, McShea WJ, Parsons AW, Millspaugh JJ (2016) A Two-Species Occupancy Model Accommodating Simultaneous Spatial and Interspecific Dependence. Ecology 97:48–53.