Worked example of the likelihood functions for repeated-visit, single-visit, and integrated occupancy models

Valentin Lauret, Hélène Labach, Matthieu Authier, Olivier Gimenez

Contents

Supporting information of the article *Using single visits into integrated occupancy models to make the most of existing monitoring programs* in *Ecology*.

A worked example

In this section, we provide a worked example of the detection histories and the likelihood functions for SV, RV and Integrated occupancy models for the same hypothetical data. We aimed at clarifying the differences in the methods, as to how the information is used.

Notation

Let's consider a fictive site s, and y_s refers to the detection history made for this site s. Two monitoring programs **A** & **B** collect data at site s during one year. Then, y_s^A and y_s^B refer to the detection histories collected at site s by respectively monitoring program **A** and **B**.

Each monitoring program collected binary data during one year at site s, with $y = 1$ if the species is detected, and $y = 0$ otherwise. For the monitoring period, we considered two situations and applied different occupancy models accordingly:

1. We divided the monitoring period into 4 sampling occasions $j = \{1, ..., 4\}$ with $J = 4$ and analyzed the data with a RV occupancy model.
2. We considered the entire monitoring period as a single sampling occasion and analyzed the data with a SV occupancy model.

Detections histories and associated likelihood

We now present the detection histories we obtain from each of the two approaches presented above to deal with the collected data. Then, we name the relevant occupancy model to analyze the detection history and we display the likelihood to link the detection history and latent occupancy state (*i.e.* site s occupied by the species is $z_s = 1$, site s unoccupied by the species $z_s = 0$).

1. RV occupancy

When dividing the monitoring period into 4 sampling occasion, program **A** detected the species during occasion $j = 2$, and $j = 4$. The species remained undetected during sampling occasions $j = 1$, and $j = 3$.

$$y_s^A = \{0, 1, 0, 1\}$$

To analyze data collected by program **A** when having 4 sampling occasions, we used a standard repeated-visit occupancy model. For each sampling occasion j, we calculated the likelihood of collected data $y_{s,j}^A$ as a Bernoulli draw, $y_{s,j}^A \sim Bernoulli(z_s p_{s,j}^A)$, with $p_{s,j}^A$ the probability of detecting the species with program **A** at site s during sampling occasion j.

Similarly, for program **B**, the detection history is $y_s^B = \{1, 0, 0, 1\}$. We used the same RV occupancy model, and for each j, the likelihood is $y_{s,j}^B \sim Bernoulli(z_s p_{s,j}^B)$

2. SV occupancy

When considering the entire monitoring period as a single sampling occasion, both program **A** and **B** detected the species at site s:

$y_s^A = \{1\}$

$y_s^B = \{1\}$

To analyze data collected by program **A** when considering one single sampling occasion, we used a single-visit occupancy model. We calculated the likelihood of collected data y_s^A as a Bernoulli draw, $y_s^A \sim Bernoulli(z_s p_s^A)$, with p_s^A the probability of detecting the species with program **A** at site s.

Similarly, for program **B**, the detection history is $y_s^B = \{1\}$. We used the same SV occupancy model, and the likelihood is $y_s^B \sim Bernoulli(z_s p_s^B)$

3. Integrated RV occupancy models

When analyzing jointly both programs **A** & **B**, the detection/non-detection is coded differently, we coded $y = 0$ if the species is not detected by program **A** nor by program **B**, $y = 1$ if the species is detected only by program **A**, $y = 2$ if the species is detected only by program **B**, and $y = 3$ if the species is detected by both programs **A** & **B**.

When dividing the monitoring period into 4 sampling occasions, we saw above that binary detection histories of both program at site s are $y_s^A = \{0, 1, 0, 1\}$ and $y_s^B = \{1, 0, 0, 1\}$.

Then, when analyzing jointly both programs with a RV detection process, the detection history is:

$$y_s^{AB} = \{2, 1, 0, 3\}$$

To analyze data collected by both programs **A** and **B** when considering 4 sampling occasions, we used an integrated RV occupancy model. We calculated the likelihood of collected data $y_{s,j}^{AB}$ as a multinomial draw, $y_s^A \sim Multinomial(1, z_s \pi_{s,j})$, with

$$\pi_{s,j} = \{P(y_{s,j}^{AB} = 0), P(y_{s,j}^{AB} = 1), P(y_{s,j}^{AB} = 2), P(y_{s,j}^{AB} = 3)\}$$
$$\pi_{s,j} = \{1 - p_{s,j}^A - p_{s,j}^B + p_{s,j}^A p_{s,j}^B, p_{s,j}^A(1 - p_{s,j}^B), p_{s,j}^B(1 - p_{s,j}^A), p_{s,j}^A p_{s,j}^B\}$$

This likelihood formulation requires that the two detection processes are independent.

4. Integrated SV occupancy models

Subsequently, when analyzing jointly both programs **A** and **B** with a SV detection process, the detection history is

$$y_s^{AB} = \{3\}$$

Then, to analyze data collected by both programs **A** and **B** when considering a single sampling occasion, we used an integrated **SV** occupancy model. We calculated the likelihood of collected data y_s^{AB} as a multinomial draw, $y_s^A \sim Multinomial(1, z_s \pi_s)$, with

$$\pi_s = \{P(y_s^{AB} = 0), P(y_s^{AB} = 1), P(y_s^{AB} = 2), P(y_s^{AB} = 3)\}$$
$$\pi_s = \{1 - p_s^A - p_s^B + p_s^A p_s^B, \quad p_s^A(1 - p_s^B), \quad p_s^B(1 - p_s^A), \quad p_s^A p_s^B\}$$