Assumptions and worked example of repeated-visits, single-visit, and integrated occupancy models

Valentin Lauret

26/03/2021

Contents

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

Modeling assumptions

In this section we aimed to list the modeling assumptions of the Repeated-Visits (RV), Single-Visit (SV), and Integrated occupancy models that we ran in the manuscript.

Repeated-visits occupancy models

Citing, (MacKenzie 2006), there are several critical assumptions for the standard occupancy model, i.e. RV occupancy.

1. Occupancy status at each site does not change over the survey season; that is, sites are "closed" to changes in occupancy
2. The probability of occupancy is constant across sites, or differences in occupancy probability are modeled using covariates
3. The probability of detection is constant across all sites and surveys or is a function of site-survey covariates
4. There is no unmodeled heterogeneity in detection probabilities
5. Detection of species and detection histories at each location are independent

Single-visit occupancy

Developing SV occupancy models, (Lele, Moreno, and Bayne 2012) underlined that SV occupancy relax the closure assumption of sampled sites between visits. Besides, the literature about SV provide some requirements and guidance to a valid application of SV occupancy models. We listed the elements below:

1. Occupancy probability and detection probability depend on covariates
2. At least two independant continuous covariates are used to estimate occupancy probability and detection probability
3. Ensure adequate numbers of occurrence. (Peach, Cohen, and Frair 2017), suggested that "estimates of occupancy probability remained unbiased across our scenarios, whereas colonization and extinction estimates became biased as occupancy probability approached extremes (i.e. 0.1 or 0.9)."
4. Shared covariates can result in biased estimates for regression coefficients, although estimates of occupancy
5. Prefer nonlinear detection model to provide accurate parameter estimates and to assume a more realistic relationship between detection and effort. A nonlinear relationship also eliminates the parameter non-identifiability problem described by (Knape and Korner-Nievergelt 2015) for single-visit occupancy models

Integrated occupancy models

Combining multiple datasets into occupancy models have been developed previously by (Nichols et al. 2008) in details by estimating occupancy at two the spatial scales. In our integration process, we extended the parametrization of a standard occupancy model detection process to include two different datasets with different detection probabilities. Doing it, we must assume that

- the two monitoring programs are independent, i.e. detection by program 1 does not affect detection probability of program 2.

Subsequently, integrated repeated-visits occupancy models have the modeling assumption of both RV occupancy and of integrated occupancy. Similarly, using integrated single-visit occupancy models cumulate both the assumptions of SV occupancy and of integrated occupancy.

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 for one site s. We aimed at clarifying the differences in the methods, as to how the information is used.

Notations

Let's consider a fictive site S, and y_s refers to the detection history made at 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, *i.e.* we coded $y = 1$ if the species is detected, and $y = 0$ is the species is not detected. To consider the monitoring period, we separated two situations and applied different occupancy models in response:

1. We divided monitoring period into 4 sampling occasions J and analyzed the data with a RV occupancy model.
2. We consider the entire monitoring period as a single sampling occasion and analyzed the data with a SV occupancy model.

Detections histories and associated likelihood

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

1. For 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-visits occupancy model. For each sampling occasion j, we calculated the likelihood of collected data $y_{s,j}^A$ as a Bernoulli draw, $y_{s,j}^A$ $Bernoulli(z_s p_{s,j}^A)$, 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 standard RV occupancy model, and for each j, the likelihood is $y_{s,j}^B$ $Bernoulli(z_s p_{s,j}^B)$

2. For 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 $Bernoulli(z_s p_s^A)$, 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 $Bernoulli(z_s p_s^B)$

3. For integrated RV occupancy models

When analyzing jointly both programs **A** & **B**, the detection/non-detection is coded differently, we coded $y = 0$ is 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 iste 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 Categorical draw, y_s^A $Multinomial(1, z_s \pi_{s,j})$, with

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

.

4. For integrated SV occupancy models

Subsequently, when analyzing jointly both programs **A** & **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 Categorical draw, y_s^A $Multinomial(1, z_s \pi_s)$, with

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

.

Literature cited

Knape, Jonas, and Fränzi Korner-Nievergelt. 2015. "Estimates from Non-Replicated Population Surveys Rely on Critical Assumptions." Edited by Nigel Yoccoz. *Methods in Ecology and Evolution* 6 (3): 298–306. https://doi.org/10.1111/2041-210X.12329.

Lele, S. R., M. Moreno, and E. Bayne. 2012. "Dealing with Detection Error in Site Occupancy Surveys: What Can We Do with a Single Survey?" *Journal of Plant Ecology* 5 (1): 22–31. https://doi.org/10.1093/jpe/rtr042.

MacKenzie, Darryl I., ed. 2006. *Occupancy Estimation and Modeling: Inferring Patterns and Dynamics of Species*. Amsterdam ; Boston: Elsevier.

Nichols, James D., Larissa L. Bailey, Allan F. O'Connell Jr., Neil W. Talancy, Evan H. Campbell Grant, Andrew T. Gilbert, Elizabeth M. Annand, Thomas P. Husband, and James E. Hines. 2008. "Multi-Scale Occupancy Estimation and Modelling Using Multiple Detection Methods." *Journal of Applied Ecology* 45 (5): 1321–29. https://doi.org/10.1111/j.1365-2664.2008.01509.x.

Peach, Michelle A., Jonathan B. Cohen, and Jacqueline L. Frair. 2017. "Single-Visit Dynamic Occupancy Models: An Approach to Account for Imperfect Detection with Atlas Data." *Journal of Applied Ecology* 54 (6): 2033–42. https://doi.org/10.1111/1365-2664.12925.