## Estimating wolf occupancy with R

**Olivier Gimenez** 

### Lecture 1

#### Introduction

Estimating wolf occupancy with R

# Occupancy to map species distribution

#### **Occupancy:** proportion of an area occupied by a species

- Species range dynamics
- Habitat preferences

. . .

- Metapopulation dynamics



## Issue of detectability < 1



True occupancy = 25%

## lssue of detectability < 1</pre>





True occupancy = 25%

Species detected in 6 occupied sites

## Occupancy underestimation





True occupancy = 25%

Species detected in 6 occupied sites

Naive occupancy estimate = 6/40 = 15%

## Issue of detectability < 1



time

## Bias in occupancy trends



## Occupancy models



SECOND EDITION

#### OCCUPANCY ESTIMATION AND MODELING

INFERRING PATTERNS AND DYNAMICS OF SPECIES OCCURRENCE

Darryl I. MacKenzie, James D. Nichols, J. Andrew Royle, Kenneth H. Pollock, Larissa L. Bailey, James E. Hines



## ECOGRAPHY

A JOURNAL OF SPACE AND TIME IN ECOLOGY

Review & synthesis 🔂 Free Access

Modelling of species distributions, range dynamics and communities under imperfect detection: advances, challenges and opportunities

Gurutzeta Guillera-Arroita 💌

First published: 20 June 2016 | https://doi.org/10.1111/ecog.02445 | Citations: 134

## Occupancy protocol



Several sampling units surveyed

# Occupancy protocol



- Several sampling units surveyed
- Collection of detection/non-detection

# Occupancy protocol



- Several sampling units surveyed
- Collection of detection/non-detection
- Replicate surveys in each unit









#### Data structure

- Sampling units = *sites*
- We do repeated *observations* at each site

	visit 1	visit 2	visit 3	
site 1	1	0	1	
site 2	0	1	1	
site 3	0	0	0	

#### Lecture 2

#### Static aka single-season occupancy models

Estimating wolf occupancy with R

Assuming closure, and independence of surveys: Pr(1001) = ?

Assuming closure, and independence of surveys:

$$Pr(1001) = \psi_1 p (1-p) (1-p) p$$

Assuming closure, and independence of surveys: Pr(0000) = ?

Assuming closure, and independence of surveys:  $Pr(0000) = \psi_1 (1 - p) (1 - p) (1 - p) (1 - p) + (1 - \psi_1)$ 

#### Single-season occupancy model



 $\psi_1$  = occupancy, p = detection

### Single-season occupancy model



Markov model

O = occupied; U = unoccupied  $\psi_1$  = occupancy,  $\rho$  = detection

### Single-season occupancy model



O = occupied; U = unoccupied  $\psi_1$  = occupancy, p = detection

#### Data structure

• We do repeated *observations* at each site

	visit 1	visit 2	visit 3	
site 1	1	0	1	
site 2	0	1	1	
site 10	0	0	0	





• Site-level covariates (e.g. % forest cover)

	visit 1	visit 2	visit 3	habitat
site 1	1	0	1	good
site 2	0	1	1	bad
site 10	0	0	0	bad



• Observation-level covariates (e.g. temperature)

	visit 1	visit 2	visit 3	date1	date2	date3
site 1	1	0	1	2	5	0
site 2	0	1	1	-4	8	2
	••••	•••		•••		
site 10	0	0	0	-1	2	-3



- Allow occupancy and detection to be a function of covariates
- When dealing with probabilities between 0 and 1, we need a link function (as in GLMs) to force estimates to remain in range
- We usually use the logit function  $logit(\theta) = log\left(\frac{\theta}{1-\theta}\right)$

#### Covariates

- Allow occupancy and detection to be a function of covariates
- When dealing with probabilities between 0 and 1, we need a link function (as in GLMs) to force estimates to remain in range
- We usually use the logit function  $logit(\theta) = log\left(\frac{\theta}{1-\theta}\right)$
- E.g. for a site-level covariate % forest cover measured at site *i* :

 $logit(\psi_i) = a + b \text{ forest}_i$ 

• Where paramaters a and b are intercept and slope to be estimated

#### Covariates

- Allow occupancy and detection to be a function of covariates
- When dealing with probabilities between 0 and 1, we need a link function (as in GLMs) to force estimates to remain in range
- We usually use the logit function  $logit(\theta) = log\left(\frac{\theta}{1-\theta}\right)$
- E.g. for an observation-level covariate temperature at site *i* in visit *j* :

 $logit(p_{ij}) = a + b temperature_{ij}$ 

• Where paramaters *a* and *b* are intercept and slope to be estimated

## Key occupancy model assumptions

- 1. Sites are closed (occupation does not change)
- 2. Independent detections
- 3. No unmodelled heterogeneity
- 4. No false positives

### Key occupancy model assumptions

#### 1. Sites are not closed (occupation does change)

- Occupancy should be interpreted as 'use'.

- Relax assumption, see Lecture 3.

### Key occupancy model assumptions

#### 2. Dependent detections

- Species easier/more difficult to detect at a site where it has already been detected, or sampling close in time.

- Adapt sampling design; account for dependence in model.
## Key occupancy model assumptions

#### 3. Heterogeneity in detection

- Occupancy is lower than it should be.

- Account for heterogeneity in model (random effects [package *ubms*], finite mixtures, Royle-Nichols model if heterogeneity due to variation in abundance [function occuRN() in *unmarked*]).

## Key occupancy model assumptions

4. False positives

- See Lecture 4.

#### Live demo

Estimating wolf occupancy with R

# ECOGRAPHY

A JOURNAL OF SPACE AND TIME IN ECOLOGY

Research 🔂 Free Access

# Mapping and explaining wolf recolonization in France using dynamic occupancy models and opportunistic data

Julie Louvrier 📉, Christophe Duchamp, Valentin Lauret, Eric Marboutin, Sarah Cubaynes, Rémi Choquet , Christian Miquel, Olivier Gimenez

### Lecture 3

#### Dynamic aka multiple-season occupancy models

Estimating wolf occupancy with R

## Single-season model assumptions

- 1. Sites are closed (occupation does not change)
- 2. Independent detections
- 3. No unmodelled heterogeneity
- 4. No false positives

## Single-season model assumptions

## 1. Sites are not closed (occupation does change)

## 2. Independent detections

### 3. No unmodelled heterogeneity

4. No false positives

• We do repeated observations at each site

	season 1						
	visit 1	visit 2	visit 3				
site 1	1	0	1				
site 2	0	1	1				
		•••					
site 10	0	0	0				

• We do repeated observations at each site within season (or year)

	Ç	season 1		season 2			
	visit 1 visit 2 visit		visit 3	visit 1	visit 2	visit 3	
site 1	1	0	1	0	0	1	
site 2	0	1	1	0	1	1	
site 10	0	0	0	0	1	0	

• We do repeated observations at each site within season (or year)

	season 1			ļ	season 2	)	season 3		
	visit 1	visit 2	visit 3	visit 1	visit 2	visit 3	visit 1	visit 2	visit 3
site 1	1	0	1	0	0	1	1	1	1
site 2	0	1	1	0	1	1	0	1	1
site 10	0	0	0	0	1	0	1	0	0

 A sequence of single-season studies conducted over several seasons (or years) at same sites

	season 1			season 2			season 3		
	visit 1	visit 2	visit 3	visit 1	visit 2	visit 3	visit 1	visit 2	visit 3
site 1	1	0	1	0	0	1	1	1	1
site 2	0	1	1	0	1	1	0	1	1
site 10	0	0	0	0	1	0	1	0	0

Sites are closed *within* season, but occupancy may change *across* seasons due to colonisation/extinction events

	season 1				season 2	2	season 3		
	visit 1	visit 2	visit 3	visit 1	visit 2	visit 3	visit 1	visit 2	visit 3
site 1	1	0	1	0	0	1	1	1	1
site 2	0	1	1	0	1	1	0	1	1
site 10	0	0	0	0	1	0	1	0	0













1 = species detected; 0 = species undetected



1 = species detected; 0 = species undetected



1 = species detected; 0 = species undetected



- $\psi_1$  = prob. a site is occupied **occupancy**
- *p* = prob. species is detected (given presence) **detection**
- $\gamma$  = prob. unoccupied site becomes occupied **colonisation**
- $\epsilon$  = prob. occupied site becomes unoccupied **extinction**

 $\psi_1$  = occupancy

p = detection

 $\gamma$  = colonisation

 $\epsilon$  = extinction

#### $Pr(110\ 000) = ?$

 $\psi_1$  = occupancy

p = detection

 $\gamma$  = colonisation

 $\varepsilon = extinction$ 

 $Pr(110\ 000) =$ 

Three replicated surveys or secondary occasions Closure assumption

 $\psi_1$  = occupancy

p = detection

 $\gamma$  = colonisation

 $\epsilon$  = extinction

#### $Pr(110\ 000) = \psi_1 \ p \ p \ (1-p)$

Three replicated surveys or secondary occasions Closure assumption

 $\psi_1$  = occupancy

p = detection

 $\gamma$  = colonisation

 $\epsilon$  = extinction

Pr(110 000) =  $\psi_1 p p (1-p) [ε + (1-ε) ...]$ 

 $\psi_1$  = occupancy

p = detection

 $\gamma$  = colonisation

 $\epsilon$  = extinction

 $Pr(110\ 000) = \psi_1 p p (1-p) [\epsilon + (1-\epsilon) (1-p) (1-p) (1-p)]$ 

 $\psi_1$  = occupancy

p = detection

 $\gamma$  = colonisation

 $\epsilon$  = extinction

#### $Pr(000\ 010) = ?$

 $\psi_1$  = occupancy

p = detection

 $\gamma$  = colonisation

 $\epsilon$  = extinction

 $Pr(000\ 010) = [\psi_1\ (1-p)\ (1-p)\ (1-\epsilon) + (1-\psi_1)\ \gamma]$ 

 $\psi_1$  = occupancy

p = detection

 $\gamma$  = colonisation

 $\epsilon$  = extinction

$$Pr(000 \ 010) = [\psi_1 \ (1-p) \ (1-p) \ (1-p) \ (1-\epsilon) + (1-\psi_1) \ \gamma] \\ \times (1-p) \ p \ (1-p)$$

## **Derived parameters**

 $\psi_1$  = occupancy

p = detection

 $\gamma$  = colonisation

 $\epsilon$  = extinction

#### - Season-specific occupancy:

$$\psi_{t+1} = \psi_t \left(1 - \varepsilon_t\right) + \left(1 - \psi_t\right) \gamma_t$$

## - Rate of change in occupancy:

$$\lambda_t = \psi_{t+1} / \psi_t$$

## Dynamic occupancy model



Markov model

hidden

## Single-season is a particular case of multi-season

#### No colonization ( $\gamma = 0$ ) and no extinction ( $\epsilon = 0$ )



#### Live demo

Estimating wolf occupancy with R

# ECOGRAPHY

A JOURNAL OF SPACE AND TIME IN ECOLOGY

Research 🔂 Free Access

# Mapping and explaining wolf recolonization in France using dynamic occupancy models and opportunistic data

Julie Louvrier 📉, Christophe Duchamp, Valentin Lauret, Eric Marboutin, Sarah Cubaynes, Rémi Choquet , Christian Miquel, Olivier Gimenez

#### Lecture 4

#### Occupancy models with species misidentification

Estimating wolf occupancy with R
## Species misidentification

OPEN OACCESS Freely available online 2013



#### Determining Occurrence Dynamics when False Positives Occur: Estimating the Range Dynamics of Wolves from Public Survey Data

David A. W. Miller<sup>1,2</sup>\*, James D. Nichols<sup>1</sup>, Justin A. Gude<sup>3</sup>, Lindsey N. Rich<sup>4</sup>, Kevin M. Podruzny<sup>3</sup>, James E. Hines<sup>1</sup>, Michael S. Mitchell<sup>4</sup>

How to account for false
 positives due to species
 misidentification?



#### Data

#### Observations by hunters (phone interviews); uncertainty in species identification



#### Telemetry; *no doubt* about species



#### Habitat quality



#### Reminder: static occupancy model



## Model allowing for false positives

## **Observation process**



## Model allowing for false positives

## **Observation process**

 $p_{10}$  = probability of false positive detection

## Model allowing for false positives

## **Observation process**

$$\begin{array}{cccc} 0 & 1 & 2 \\ U \left( \begin{array}{ccc} 1 - p_{10} & p_{10} & 0 \\ 1 - p_{11} & (1 - b)p_{11} & bp_{11} \end{array} \right) \end{array}$$

 $p_{11}$  = probability of detection b = probability that a detection is classified as unambiguous

# Estimates of occupancy for gray wolves in northern Montana from 2007–2010



# Estimates of occupancy for gray wolves in northern Montana from 2007–2010



#### Live demo

Estimating wolf occupancy with R

#### Lecture 5

#### Estimating species co-occurrence

Estimating wolf occupancy with R

## Rationale

- Several (say 2) different species on a site
- Interactions affect occupancy probabilities
- Detection of a species affected by presence of another one: blurred interactions
- Examples: predation, mutualism, competition, ...

#### Questions you might want to ask

- Are the species interacting or not? (beware: co-occurrence is not necessary interaction)
- Do species interactions vary along an environmental gradient?
- What is marginal occupancy probability of some species (that is averaged on presence/absence of all other species)
- What is probability of some species conditional on presence or absence of other species
- What is the relative contribution of environmental vs species interactions in occupancy?

#### States

U = site unoccupied
A = site occupied by species A only
B = site occupied by species B only
AB = site occupied by both species

#### State process

#### $\psi^{A}$ = prob. a site is occupied by species A

#### $\psi^{B}$ = prob. a site is occupied by species B

#### $\psi^{AB}$ = prob. a site is occupied by species A and B

## **Conditional probabilities**

 $\psi^{A|B}$  = prob. a site is occupied by species A given presence of species B =  $\psi^{AB} / \psi^{B}$ 

 $\psi^{B|A}$  = prob. a site is occupied by species B given presence of species A =  $\psi^{AB} / \psi^{A}$ 

## Venn diagram



Site unoccupied with prob.:  $1-\psi^A-\psi^B$  +  $\psi^{AB}$ 

#### Events

- 0 = species undetected
- 1 = A detected
- 2 = B detected
- 3 = both species detected

#### **Observation process**

- $p^{A}$  = prob. detecting species A given only species A is present
- $p^{B}$  = prob. detecting species B given only species B is present
- $r^{AB}$  = prob. detecting both species A and B when both present
- r<sup>Ab</sup> = prob. detecting species A but not B when both present
- $r^{aB}$  = prob. detecting species B but not A when both present
- r<sup>ab</sup> = prob. detecting neither species when both present

#### Initial states

## Initial states



#### State process

## State process

# 

#### **Observation process**

## **Observation process**

## Quantifying interactions

- Interaction estimated by:  $\eta = \psi^{AB} / (\psi^A \psi^B)$ 
  - $\eta < 1 avoidance$  (less frequent than expected)
  - $\eta > 1 convergence$  (more frequent than expected)
  - $\eta = 1 \text{independence} (\psi^{AB} = \psi^A \psi^B)$

#### Live demo

Estimating wolf occupancy with R

#### Lecture 6

#### Conclusions

Estimating wolf occupancy with R

## Conclusions

1. We covered several occupancy models

- Single-season, dynamic models
- False-positives
- Species interactions

## Conclusions

1. We covered several occupancy models

- Single-season, dynamic models
- False-positives
- Species interactions

2. Implementation in R using package *unmarked* 

## We did not cover...

- How to choose sites? Occasions?
  - site selection
  - allocation of effort
  - design comparisons
  - survey timing
- Goodness-of-fit testing
- A few other models...

*Journal of Applied Ecology* 2005 **42**, 1105–1114

#### METHODOLOGICAL INSIGHTS Designing occupancy studies: general advice and allocating survey effort

DARRYL I. MACKENZIE\* and J. ANDREW ROYLE<sup>†</sup> \*Proteus Wildlife Research Consultants, PO Box 5193, Dunedin, New Zealand; <sup>†</sup>US Geological Survey, Patuxent Wildlife Research Center, 12100 Beech Forest Road, Laurel, MD 20708–4017, USA

#### Assessing the Fit of Site-Occupancy Models

Darryl I. MACKENZIE and Larissa L. BAILEY

©2004 American Statistical Association and the International Biometric Society Journal of Agricultural, Biological, and Environmental Statistics, Volume 9, Number 3, Pages 300–318 DOI: 10.1198/108571104X3361

habitat and species occurrence dynamics, multistate, heterogeneity, ...

#### References – methods & applications

- Guillera-Arroita, G. (2017), Modelling of species distributions, range dynamics and communities under imperfect detection: advances, challenges and opportunities. *Ecography*, 40: 281-295.
- Louvrier, J., C. Duchamp, V. Lauret, E. Marboutin, S. Cubaynes, R. Choquet, C. Miquel, O. Gimenez (2017). Mapping and explaining wolf recolonization in France using dynamic occupancy models and opportunistic data. *Ecography*. 41: 647-660.
- MacKenzie, D. I., Nichols, J. D., Royle, J. A., Pollock, K. H., Bailey, L., & Hines, J. E. (2018). Occupancy estimation and modeling: Inferring patterns and dynamics of species occurrence, 2nd edition. San Diego, CA: Academic Press.
- Miller, D. A. W., Nichols, J. D., Gude, J. A., Rich, L. N., Podruzny, K. M., Hines, J. E., & Mitchell, M. S. (2013). Determining occurrence dynamics when false positives occur: Estimating the range dynamics of wolves from public survey data. *PLoS ONE*, 8, e65808.
- Rota, C. T., Fletcher Jr, R. J., Dorazio, R. M., & Betts, M. G. (2009). Occupancy estimation and the closure assumption. *Journal of Applied Ecology*, 46, 1173–1181.

#### References – software

- Doser, J. W., Finley, A. O., Kéry, M., & Zipkin, E. F. (2022). spOccupancy: An R package for single-species, multi-species, and integrated spatial occupancy models. *Methods in Ecology* and Evolution, 13, 1670–1678.
- Fiske, I., & Chandler, R. (2011). **unmarked**: An R Package for Fitting Hierarchical Models of Wildlife Occurrence and Abundance. *Journal of Statistical Software*, *43*(10), 1–23.
- Kellner, K.F., A.D. Smith, J.A. Royle, M. Kéry, J.L. Belant, R. B. Chandler (2023). The unmarked R package: Twelve years of advances in occurrence and abundance modelling in ecology. *Methods in Ecology and Evolution*. In Press.
- Gimenez, O., Blanc, L., Besnard, A., Pradel, R., Doherty, P.F., Jr, Marboutin, E. and Choquet, R. (2014), Fitting occupancy models with E-SURGE: hidden Markov modelling of presence– absence data. *Methods in Ecology and Evolution*, 5: 592-597.
- Kellner, K. F., Fowler, N. L., Petroelje, T. R., Kautz, T. M., Beyer, D. E., & Belant, J. L. (2022).
   ubms: An R package for fitting hierarchical occupancy and N-mixture abundance models in a Bayesian framework. *Methods in Ecology and Evolution*, 13, 577–584.