

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
- ...

0	0	1	0	0
0	1	0	0	0
0	1	1	0	0
0	0	1	0	0
1	0	1	0	0
0	1	0	0	0
0	0	0	0	1
0	0	0	0	1

$$10/40 = 0.25$$

Issue of detectability < 1

0	0	1	0	0
0	1	0	0	0
0	1	1	0	0
0	0	1	0	0
1	0	1	0	0
0	1	0	0	0
0	0	0	0	1
0	0	0	0	1

True occupancy = 25%

Issue of detectability < 1

0	0	1	0	0
0	1	0	0	0
0	1	1	0	0
0	0	1	0	0
1	0	1	0	0
0	1	0	0	0
0	0	0	0	1
0	0	0	0	1

True occupancy = 25%

0	0	1	0	0
0	1	0	0	0
0	1	1	0	0
0	0	1	0	0
1	0	1	0	0
0	1	0	0	0
0	0	0	0	1
0	0	0	0	1

Species detected in 6 occupied sites

Occupancy underestimation

0	0	1	0	0
0	1	0	0	0
0	1	1	0	0
0	0	1	0	0
1	0	1	0	0
0	1	0	0	0
0	0	0	0	1
0	0	0	0	1

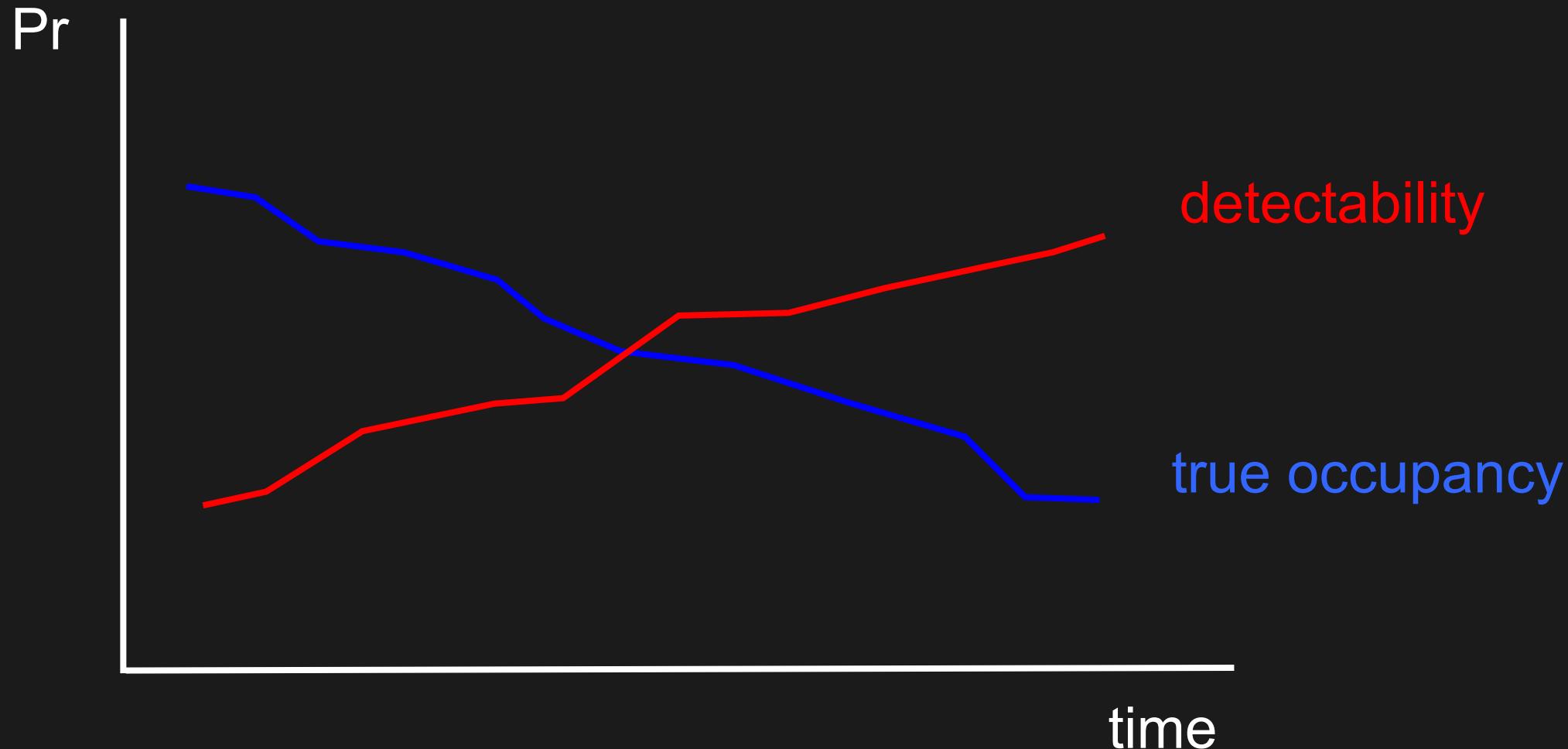
True occupancy = 25%

0	0	1	0	0
0	1	0	0	0
0	1	1	0	0
0	0	1	0	0
1	0	1	0	0
0	1	0	0	0
0	0	0	0	1
0	0	0	0	1

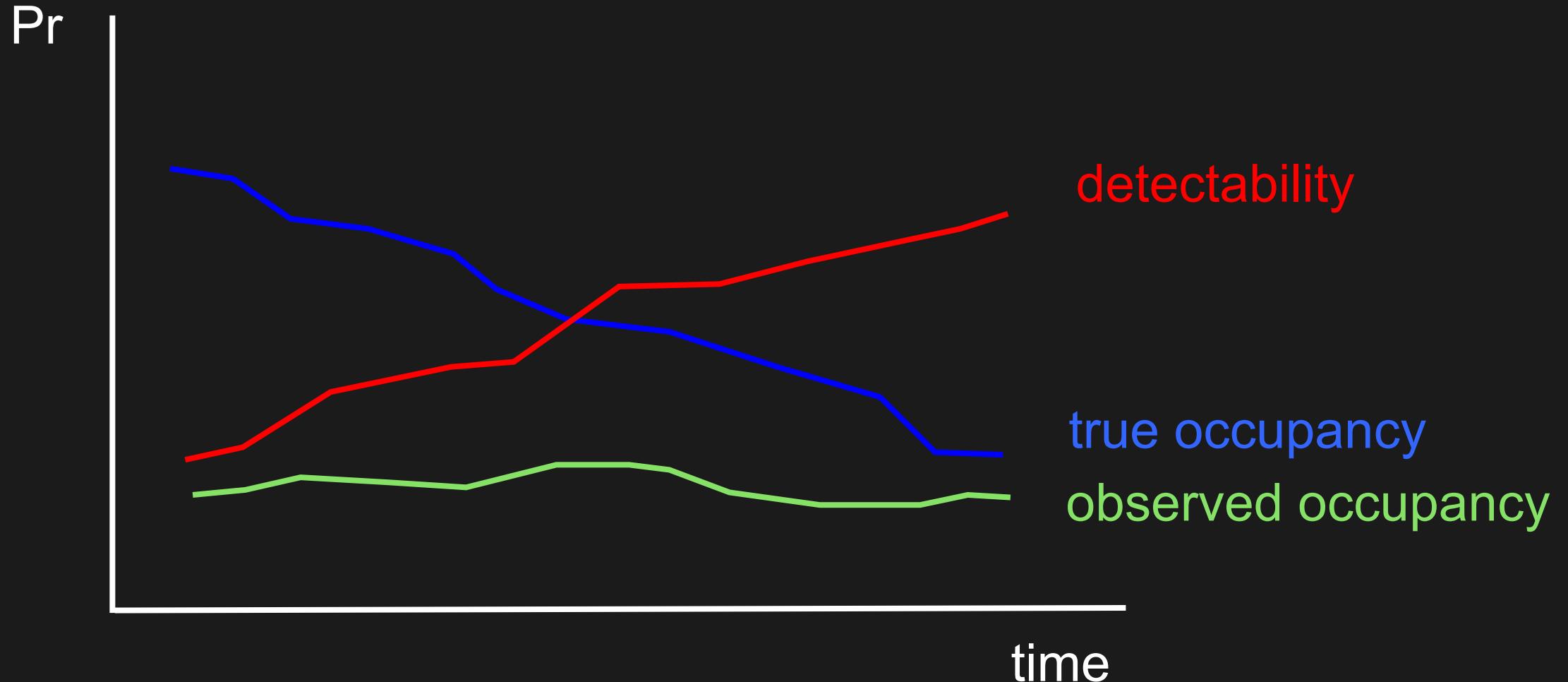
Species detected in 6 occupied sites

Naive occupancy estimate = 6/40 = 15%

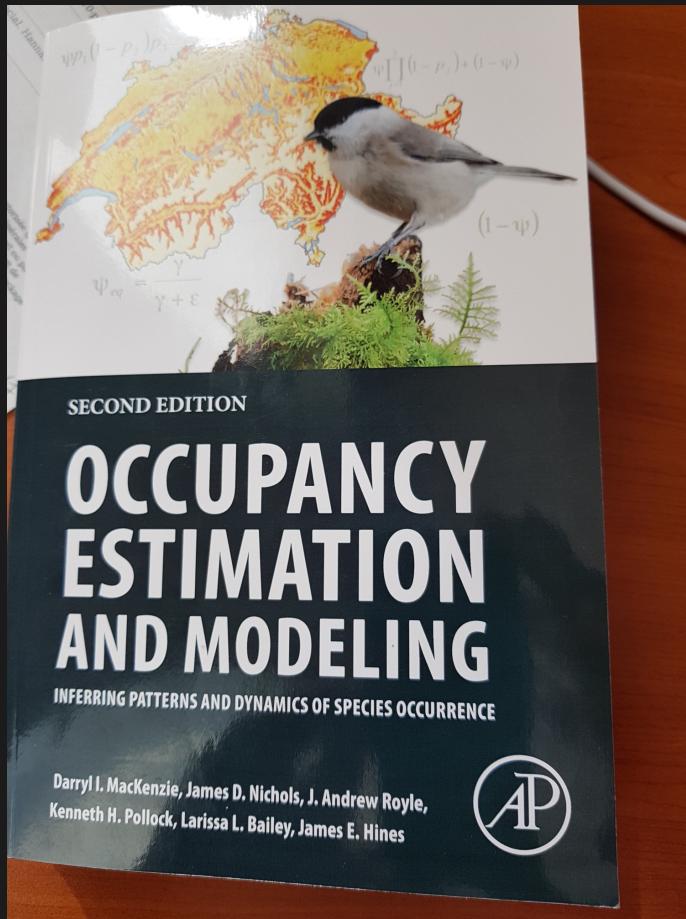
Issue of detectability < 1



Bias in occupancy trends



Occupancy models



ECOGRAPHY

A JOURNAL OF SPACE AND TIME IN ECOLOGY

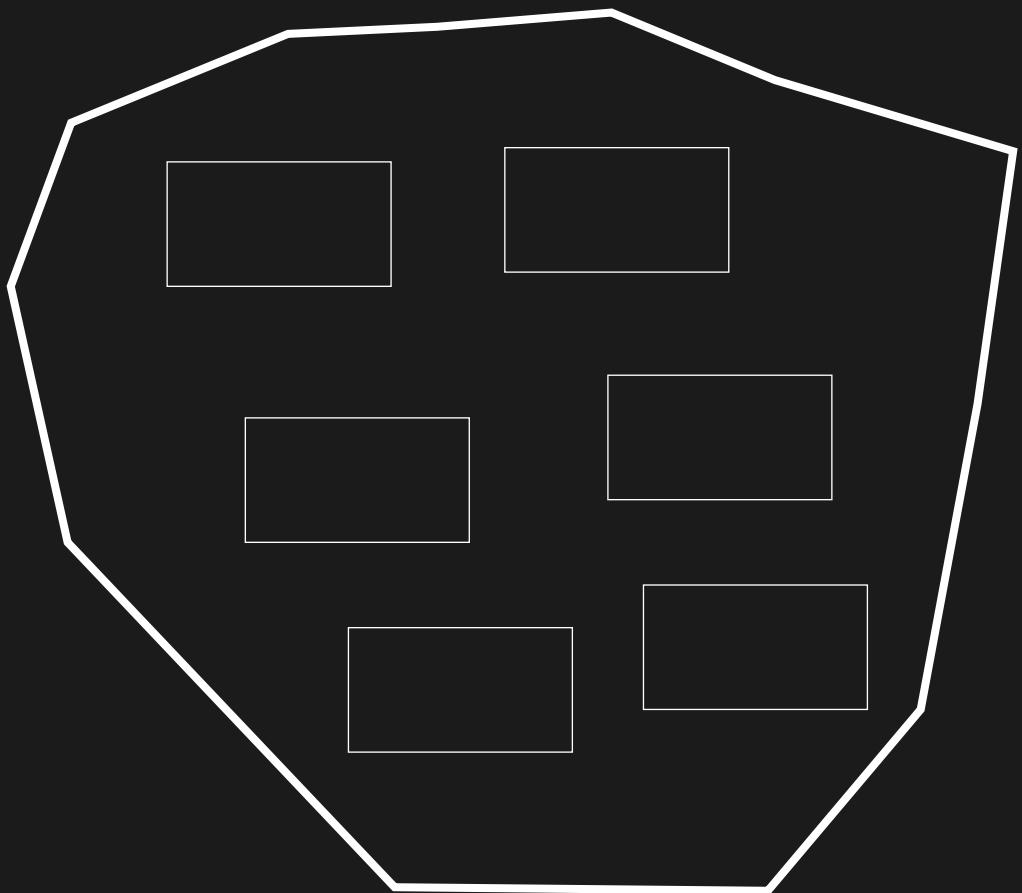
Review & synthesis | Free Access |

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

Gurutzeta Guillera-Arroita

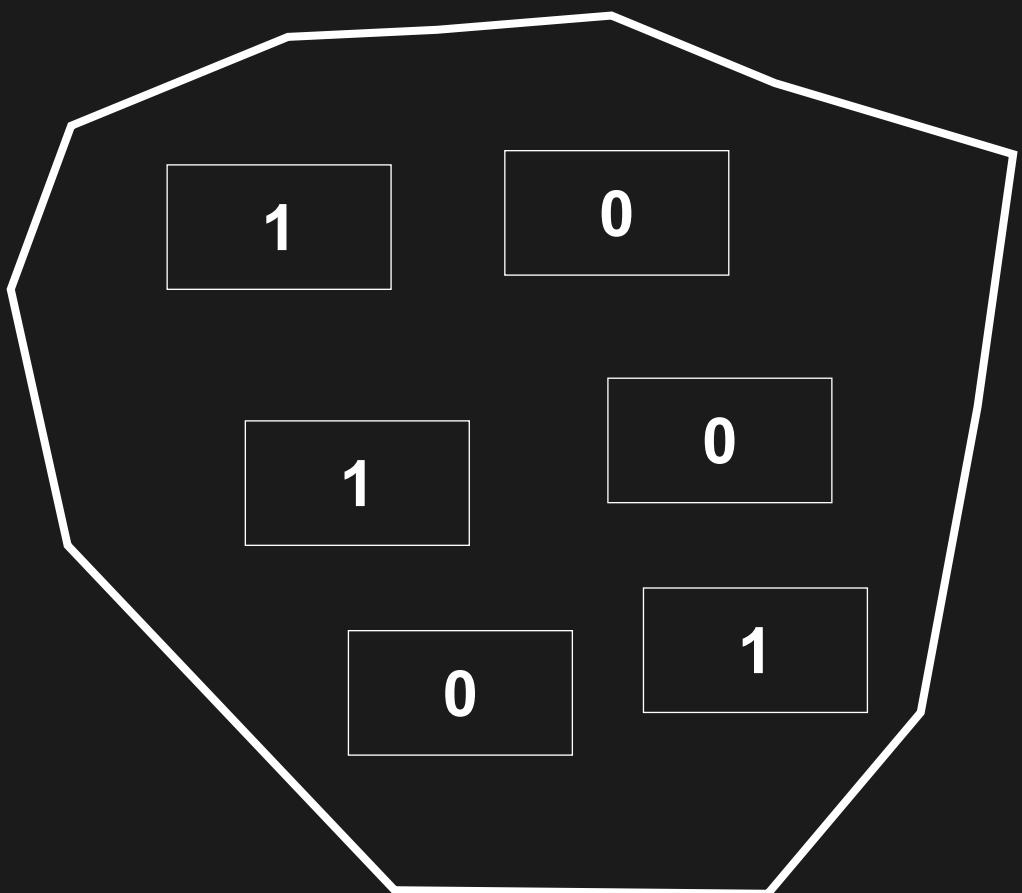
First published: 20 June 2016 | <https://doi.org/10.1111/ecog.02445> | Citations: 134

Occupancy protocol



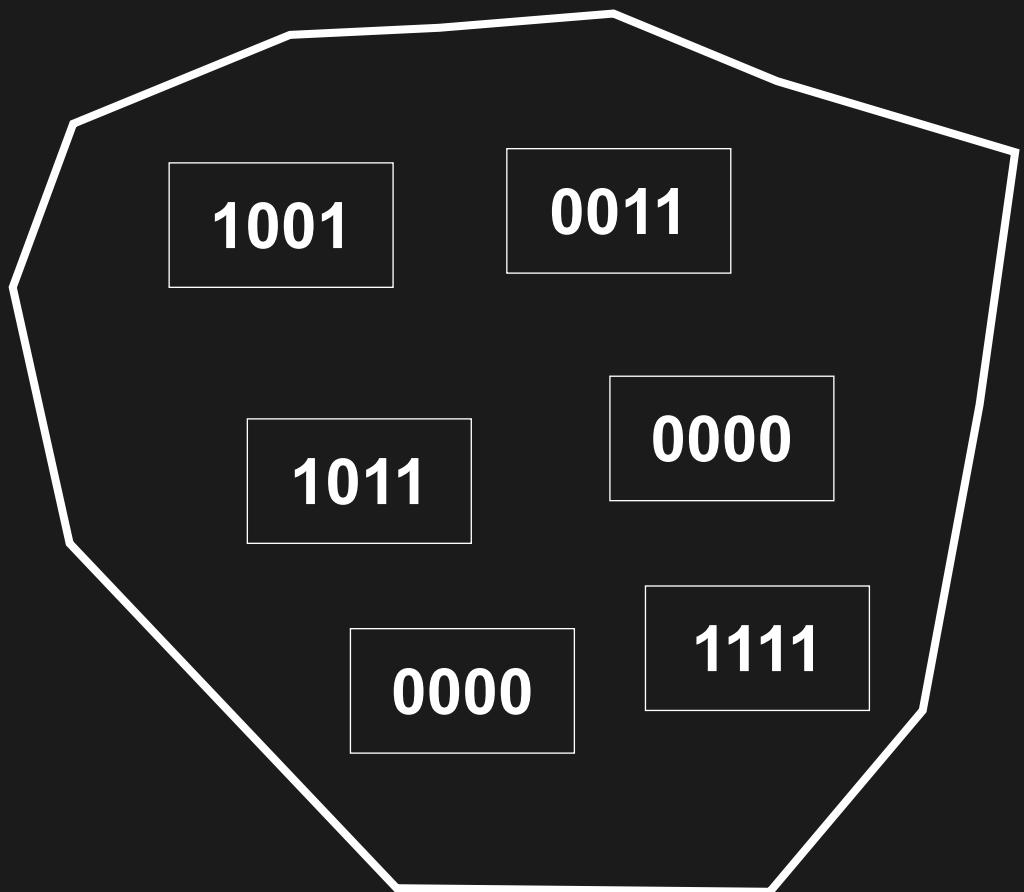
- Several sampling units surveyed

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

True or false absence?

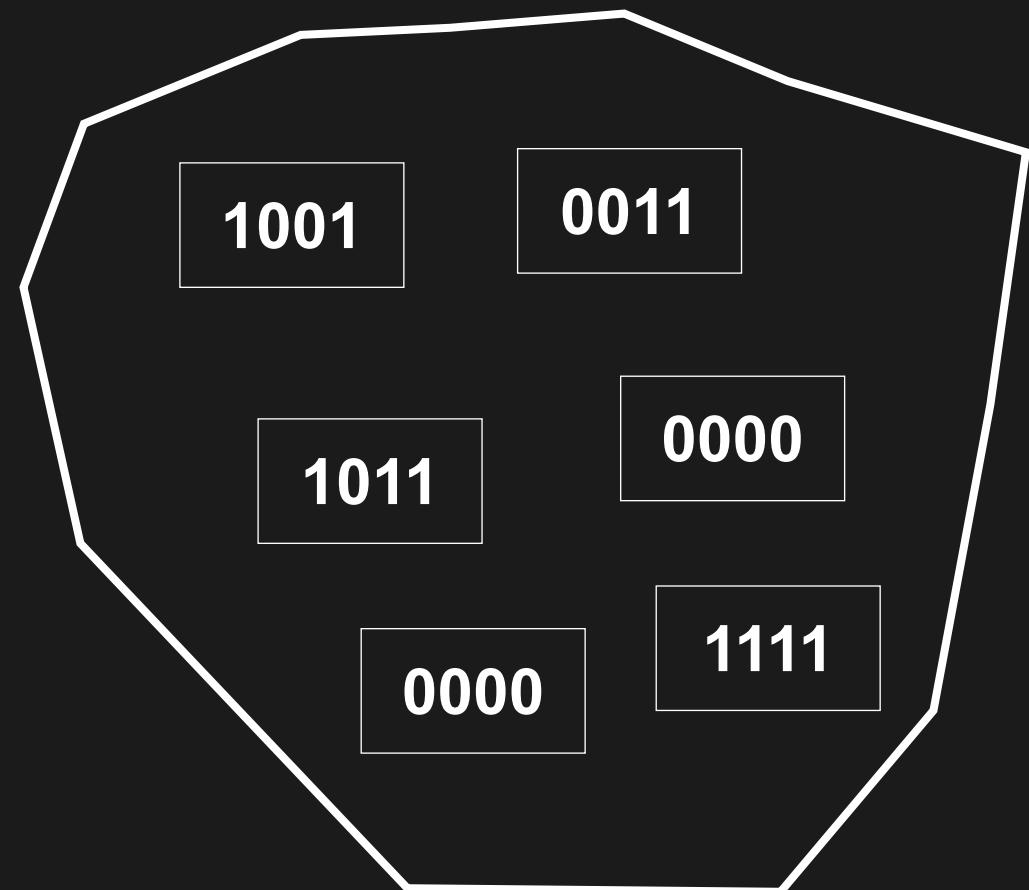


reality

True or false absence?



reality



observation

True or false absence?



reality



observation

True or false absence?



reality



observation

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

Single-season occupancy model likelihood

ψ_1 = prob. a site is initially occupied - **occupancy**

p = prob. species is detected (given presence) – **detection**

Single-season occupancy model likelihood

ψ_1 = prob. a site is initially occupied - **occupancy**

p = prob. species is detected (given presence) – **detection**

Assuming closure, and independence of surveys:

$\Pr(1001) = ?$

Single-season occupancy model likelihood

ψ_1 = prob. a site is initially occupied - **occupancy**

p = prob. species is detected (given presence) – **detection**

Assuming closure, and independence of surveys:

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

Single-season occupancy model likelihood

ψ_1 = prob. a site is initially occupied - **occupancy**

p = prob. species is detected (given presence) – **detection**

Assuming closure, and independence of surveys:

$\Pr(0000) = ?$

Single-season occupancy model likelihood

ψ_1 = prob. a site is initially occupied - **occupancy**

p = prob. species is detected (given presence) – **detection**

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

Initial states

$$\begin{matrix} U & O \\ (1 - \psi_1) & \psi_1 \end{matrix}$$

O = occupied; U = unoccupied

ψ_1 = occupancy, p = detection

Single-season occupancy model

Initial states

$$\begin{matrix} U & O \\ (1 - \psi_1) & \psi_1 \end{matrix}$$

State process

$$t \begin{matrix} U & O \\ O & U \end{matrix} \left(\begin{matrix} 1 & 0 \\ 0 & 1 \end{matrix} \right)^{t+1}$$

Markov model

O = occupied; U = unoccupied

ψ_1 = occupancy, p = detection

Single-season occupancy model

Initial states

$$\begin{matrix} U & O \\ (1 - \psi_1) & \psi_1 \end{matrix}$$

State process

$$t \begin{matrix} U & O \\ O & U \end{matrix} \left(\begin{matrix} 1 & 0 \\ 0 & 1 \end{matrix} \right)^{t+1}$$

Markov model

Observation process

$$t \begin{matrix} U & O \\ O & U \end{matrix} \left(\begin{matrix} 1 & 0 \\ 1 - p & p \end{matrix} \right)$$

hidden

O = occupied; U = unoccupied

ψ_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 3	0	0	0

Covariates

Covariates

- 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 3	0	0	0	bad

Covariates

- 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 3	0	0	0		-1	2	-3

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 $\text{logit}(\theta) = \log\left(\frac{\theta}{1-\theta}\right)$

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 $\text{logit}(\theta) = \log\left(\frac{\theta}{1-\theta}\right)$
- E.g. for a site-level covariate % forest cover measured at site i :

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

- Where parameters 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 $\text{logit}(\theta) = \log\left(\frac{\theta}{1-\theta}\right)$
- E.g. for an observation-level covariate temperature at site i in visit j :

$$\text{logit}(p_{ij}) = a + b \text{ temperature}_{ij}$$

- Where parameters 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 Louvier , 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

Data structure

- 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 3	0	0	0

Data structure

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

	season 1			season 2		
	visit 1	visit 2	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 3	0	0	0	0	1	0

Data structure

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

	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 3	0	0	0	0	1	0	1	0	0

Data structure

- 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 3	0	0	0	0	1	0	1	0	0

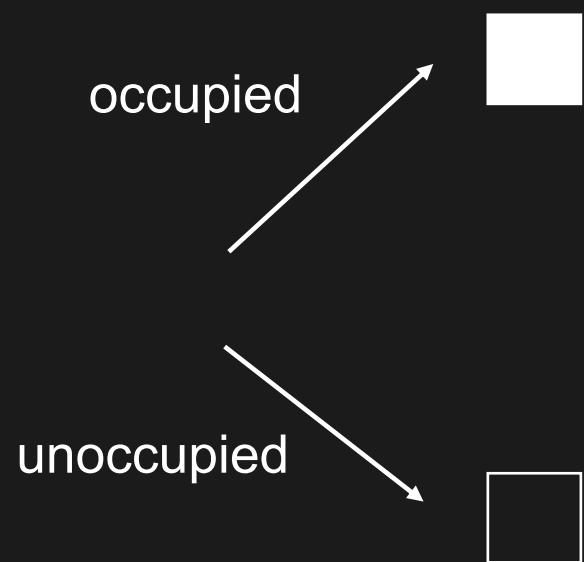
Data structure

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

	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 3	0	0	0	0	1	0	1	0	0

Dynamic of the states

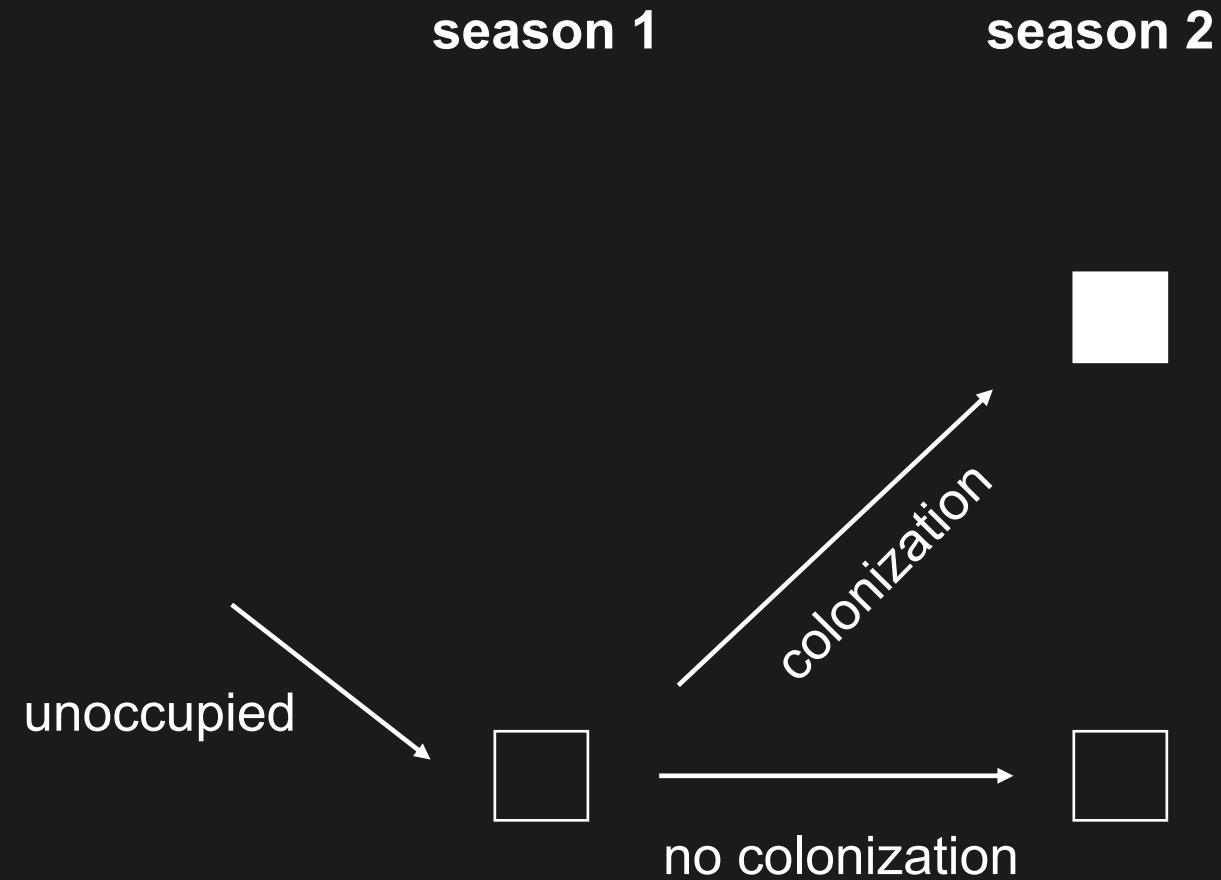
season 1



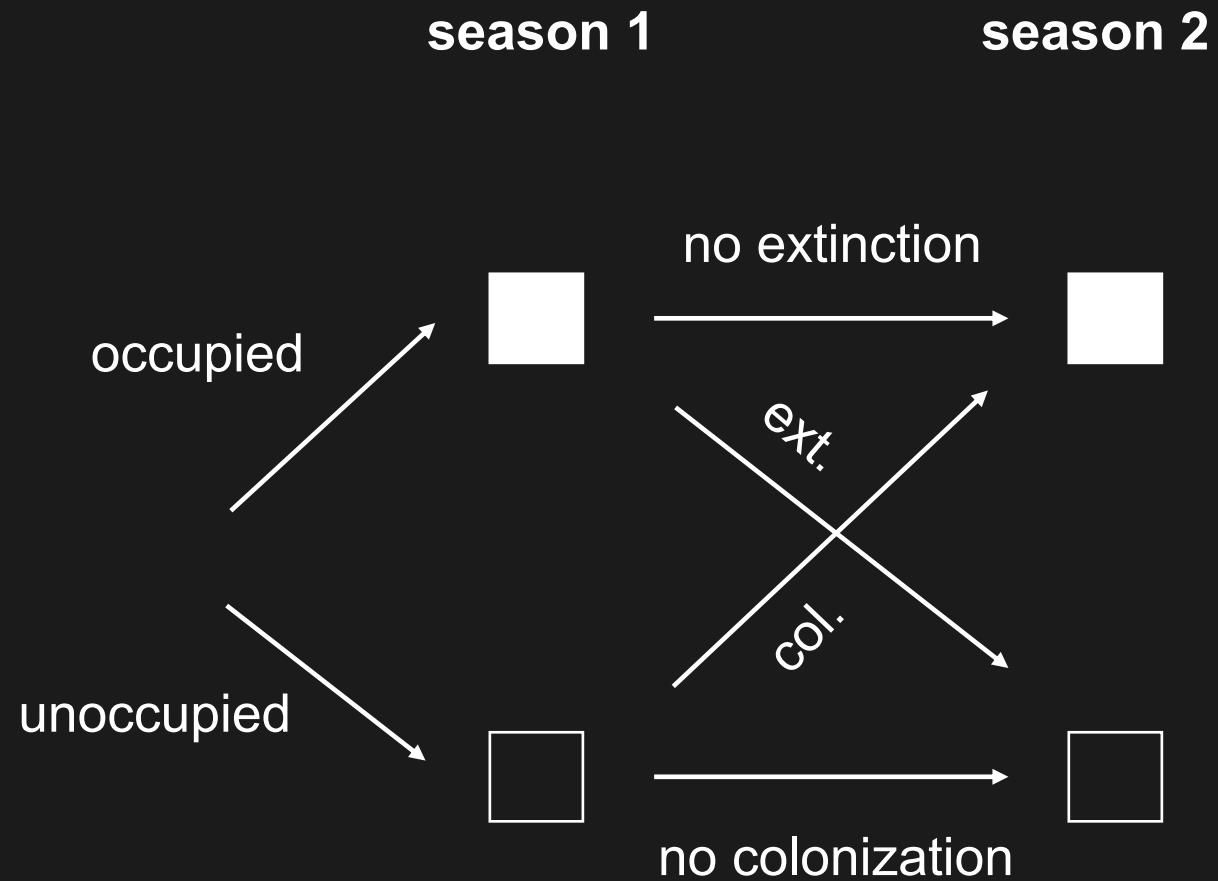
Dynamic of the states



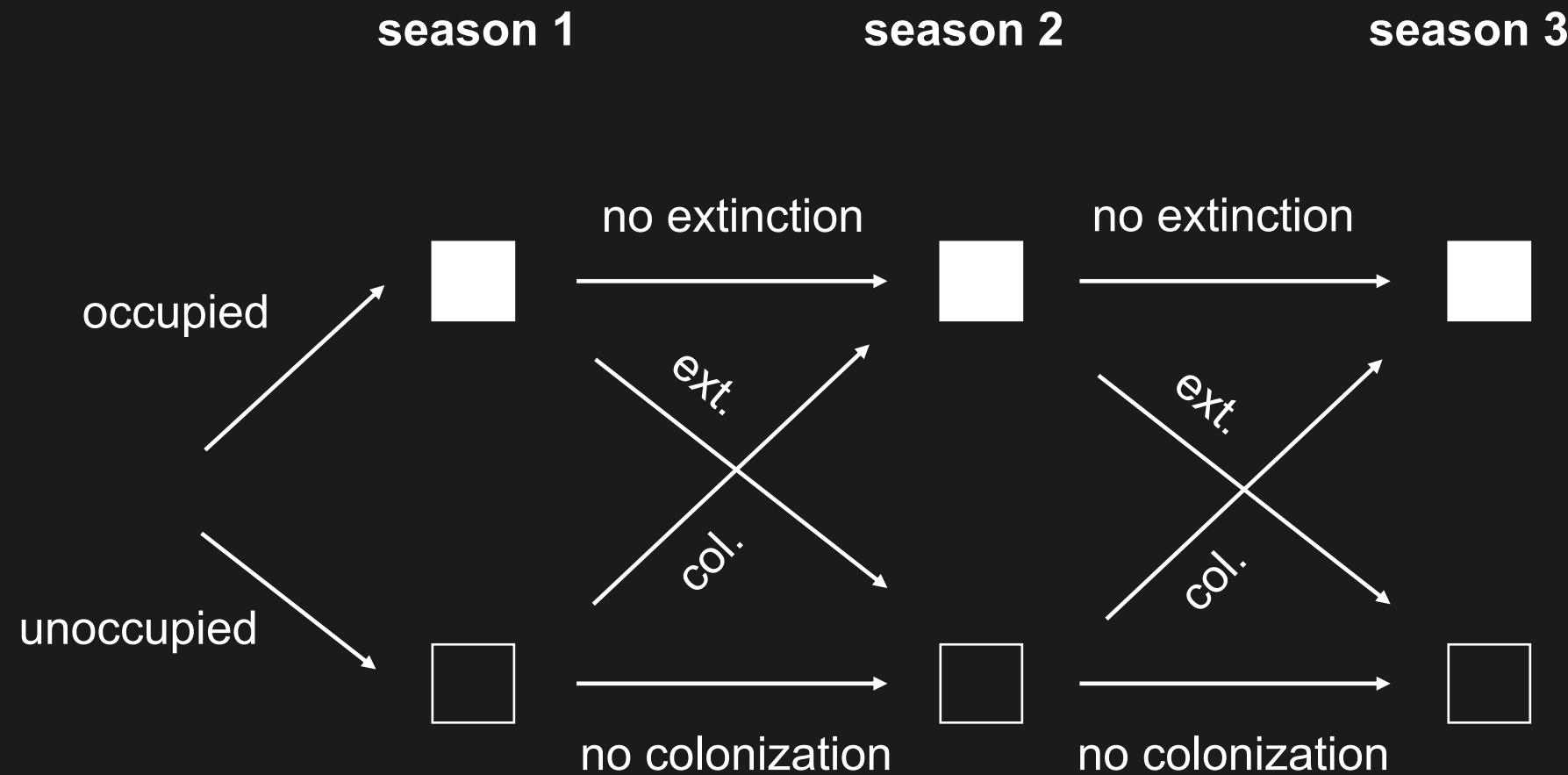
Dynamic of the states



Dynamic of the states



Dynamic of the states



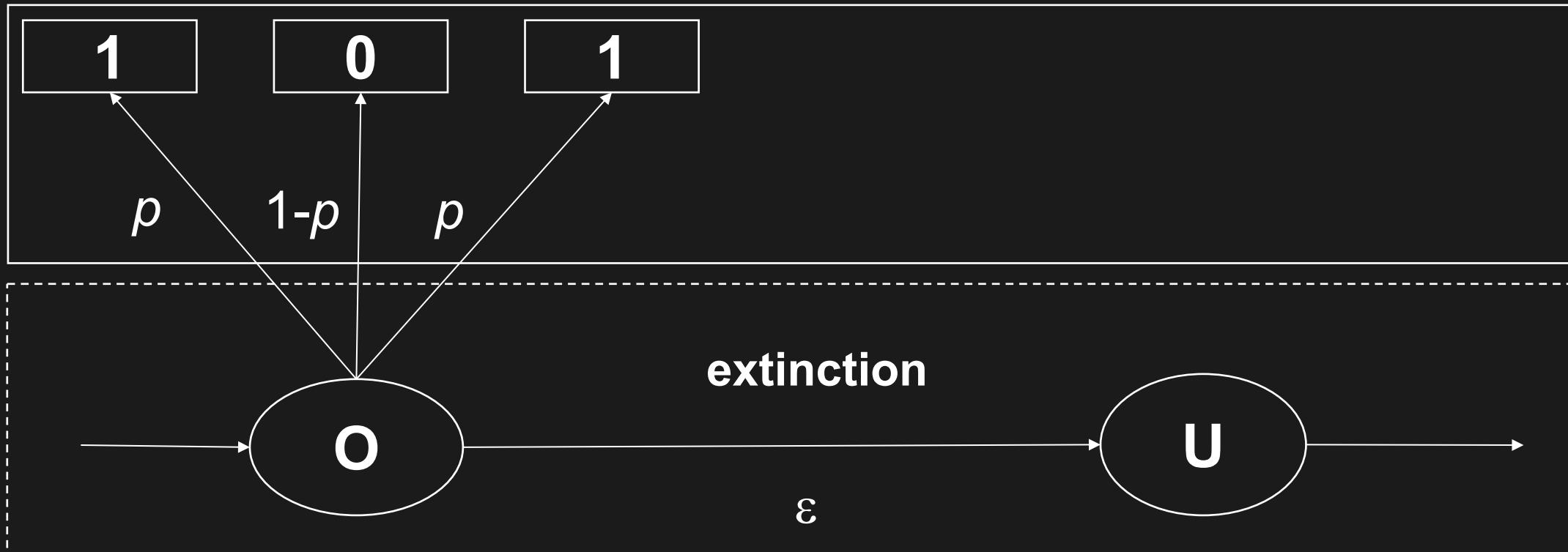
Dynamic (multi-season) occupancy models



O = occupied; U = unoccupied

Dynamic (multi-season) occupancy models

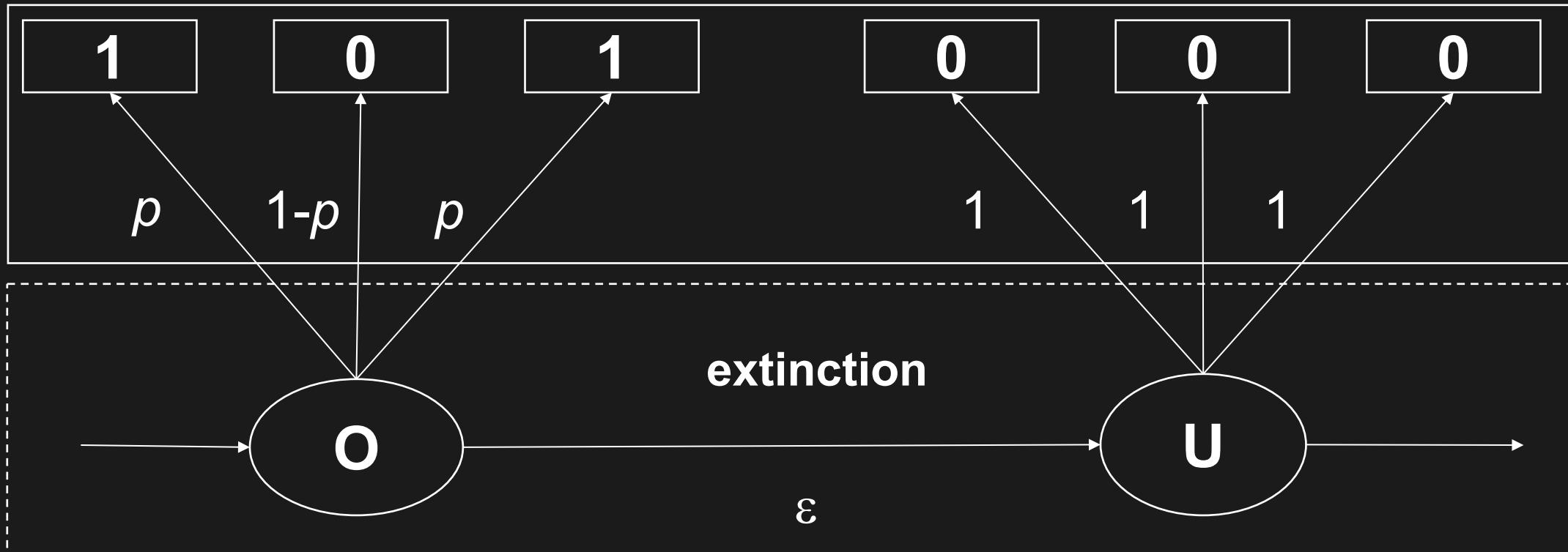
1 = species detected; 0 = species undetected



O = occupied; U = unoccupied

Dynamic (multi-season) occupancy models

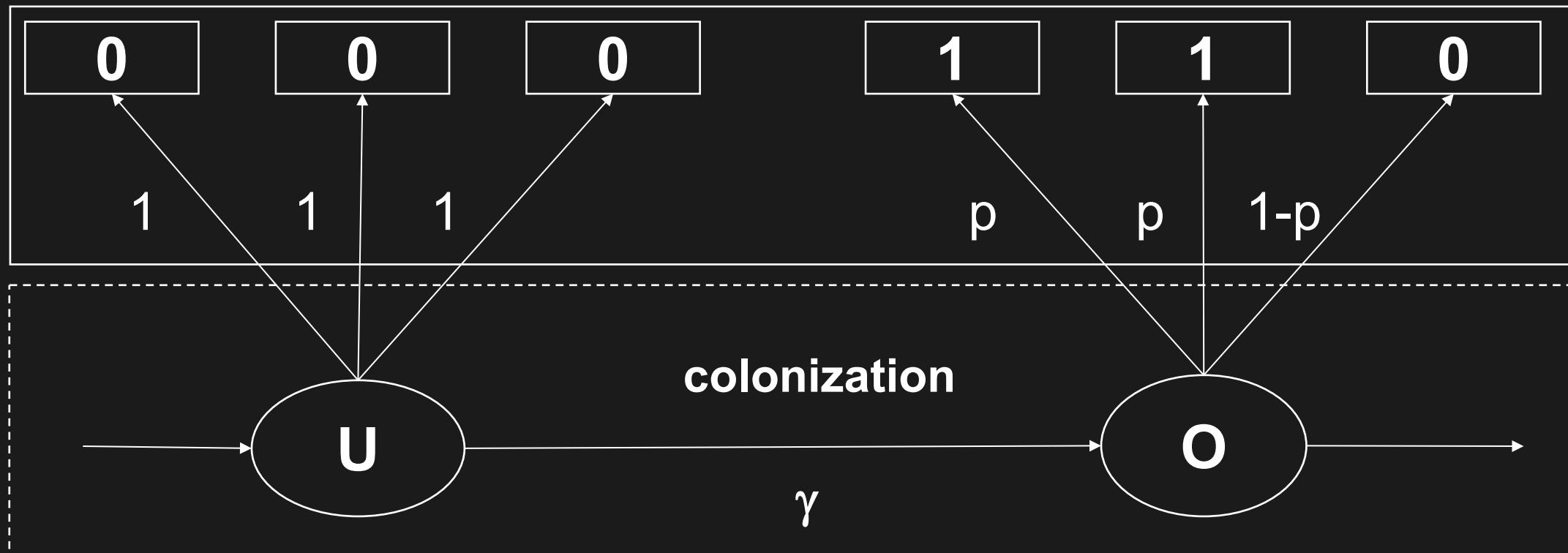
1 = species detected; 0 = species undetected



O = occupied; U = unoccupied

Dynamic (multi-season) occupancy models

1 = species detected; 0 = species undetected



O = occupied; U = unoccupied

Dynamic (multi-season) occupancy models

ψ_1 = prob. a site is occupied - **occupancy**

p = prob. species is detected (given presence) – **detection**

γ = prob. unoccupied site becomes occupied – **colonisation**

ε = prob. occupied site becomes unoccupied – **extinction**

Dynamic occupancy model likelihood

ψ_1 = **occupancy**

p = **detection**

γ = **colonisation**

ε = **extinction**

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

Dynamic occupancy model likelihood

ψ_1 = **occupancy**

p = **detection**

γ = **colonisation**

ε = **extinction**

$\Pr(110\ 000) =$

Three replicated surveys or secondary occasions

Closure assumption

Dynamic occupancy model likelihood

ψ_1 = **occupancy**

p = **detection**

γ = **colonisation**

ε = **extinction**

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

Three replicated surveys or secondary occasions

Closure assumption

Dynamic occupancy model likelihood

ψ_1 = **occupancy**

p = **detection**

γ = **colonisation**

ε = **extinction**

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

Dynamic occupancy model likelihood

ψ_1 = **occupancy**

p = **detection**

γ = **colonisation**

ε = **extinction**

$$\Pr(110 \text{ } 000) = \psi_1 p p (1 - p) [\varepsilon + (1 - \varepsilon) (1 - p) (1 - p) (1 - p)]$$

Dynamic occupancy model likelihood

ψ_1 = **occupancy**

p = **detection**

γ = **colonisation**

ε = **extinction**

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

Dynamic occupancy model likelihood

ψ_1 = **occupancy**

p = **detection**

γ = **colonisation**

ε = **extinction**

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

Dynamic occupancy model likelihood

ψ_1 = **occupancy**

p = **detection**

γ = **colonisation**

ε = **extinction**

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

Derived parameters

ψ_t = **occupancy**

- Season-specific occupancy:

p = **detection**

$$\Psi_{t+1} = \Psi_t (1 - \varepsilon_t) + (1 - \Psi_t) \gamma_t$$

γ = **colonisation**

ε = **extinction**

- Rate of change in occupancy:

$$\lambda_t = \Psi_{t+1} / \Psi_t$$

Dynamic occupancy model

Initial states

$$\begin{pmatrix} U & O \\ 1 - \psi_1 & \psi_1 \end{pmatrix}$$

State process

$$\begin{matrix} U & O \\ O & O \end{matrix} \begin{pmatrix} 1 - \gamma & \gamma \\ \epsilon & 1 - \epsilon \end{pmatrix}$$

Markov model

Observation process

$$\begin{matrix} 0 & 1 \\ O & O \end{matrix} \begin{pmatrix} 1 & 0 \\ 1 - p & p \end{pmatrix}$$

hidden

Single-season is a particular case of multi-season

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

Initial states	State process	Observation process
$\begin{matrix} U & O \\ (1 - \psi_1) & \psi_1 \end{matrix}$	$\begin{matrix} U & O \\ U & \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ O & \end{pmatrix}$	$\begin{matrix} 0 & 1 \\ U & \begin{pmatrix} 1 & 0 \\ 1 - p & p \end{pmatrix} \\ O & \end{pmatrix}$

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 Louvier , 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  ACCESS Freely available online 2013



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

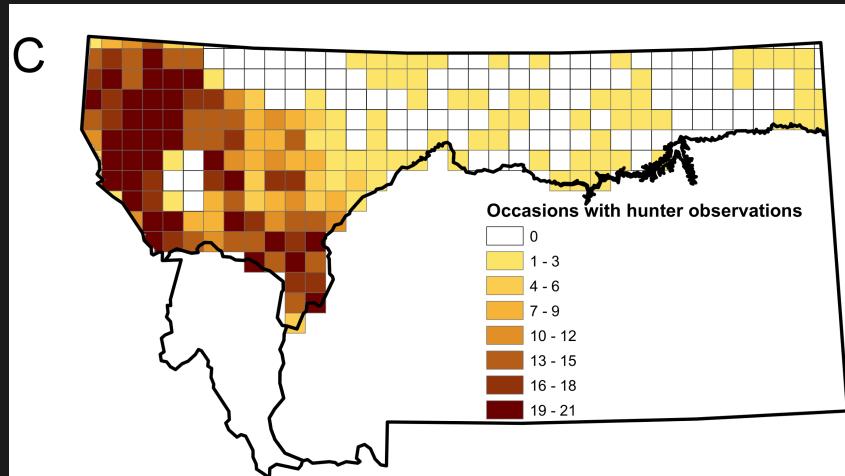
David A. W. Miller^{1,2*}, James D. Nichols¹, Justin A. Gude³, Lindsey N. Rich⁴, Kevin M. Podruzny³, James E. Hines¹, Michael S. Mitchell⁴

- How to account for false positives due to species misidentification?

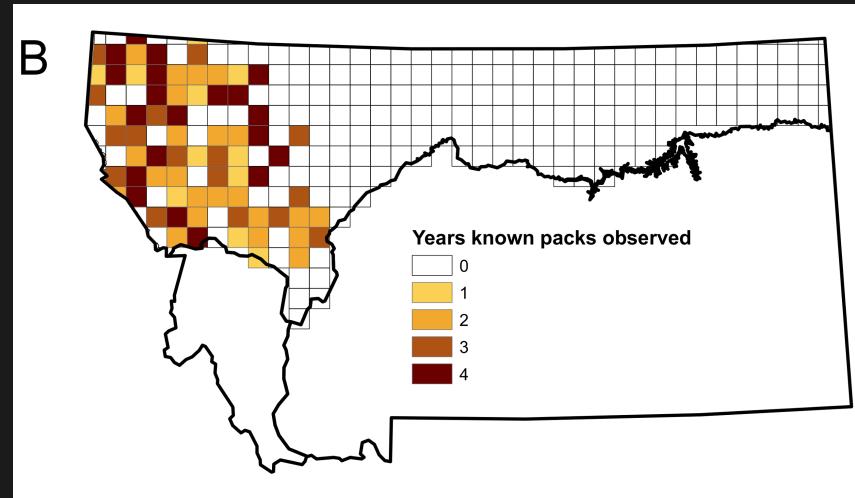


Data

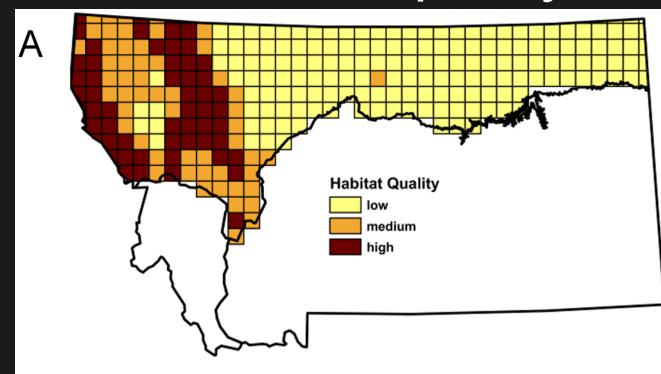
Observations by hunters (phone interviews);
uncertainty in species identification



Telemetry;
no doubt about species



Habitat quality



Reminder: static occupancy model

Initial states

$$\begin{pmatrix} U & O \\ (1 - \psi_1) & \psi_1 \end{pmatrix}$$

State process

$$\begin{matrix} U & O \\ O & \left(\begin{matrix} 1 & 0 \\ 0 & 1 \end{matrix} \right) \end{matrix}$$

Observation process

$$\begin{matrix} 0 & 1 \\ U & \left(\begin{matrix} 1 & 0 \\ 1 - p & p \end{matrix} \right) \end{matrix}$$

Model allowing for false positives

Observation process

0

1

2

Uncertain
detection

Detection with no doubt
about the species

Model allowing for false positives

Observation process

$$U \begin{pmatrix} 1 - p_{10} & p_{10} & 0 \end{pmatrix}$$

p_{10} = probability of false positive detection

Model allowing for false positives

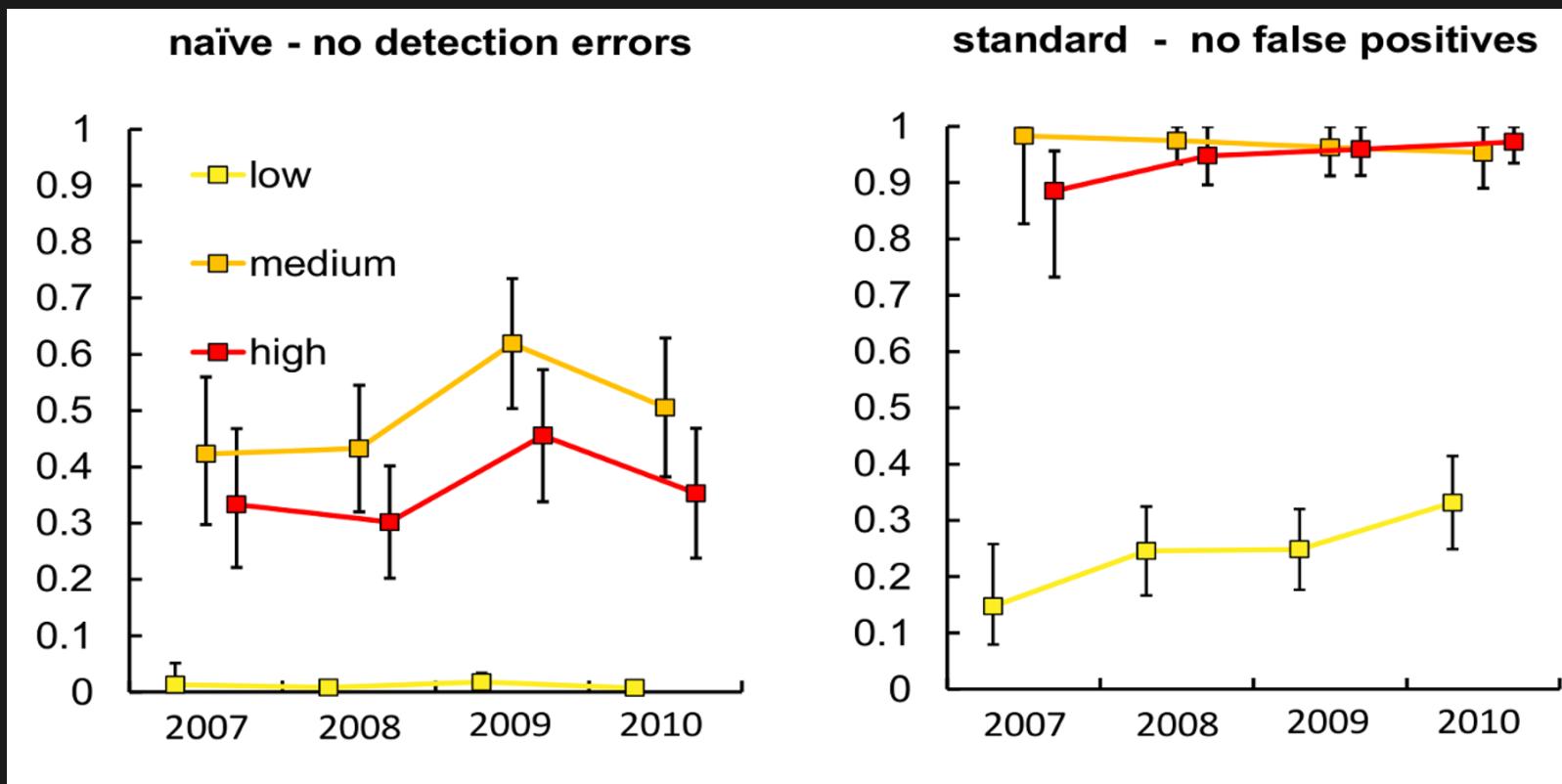
Observation process

$$U \begin{pmatrix} 0 & 1 & 2 \\ 1 - p_{10} & p_{10} & 0 \\ 1 - p_{11} & (1 - b)p_{11} & bp_{11} \end{pmatrix}$$

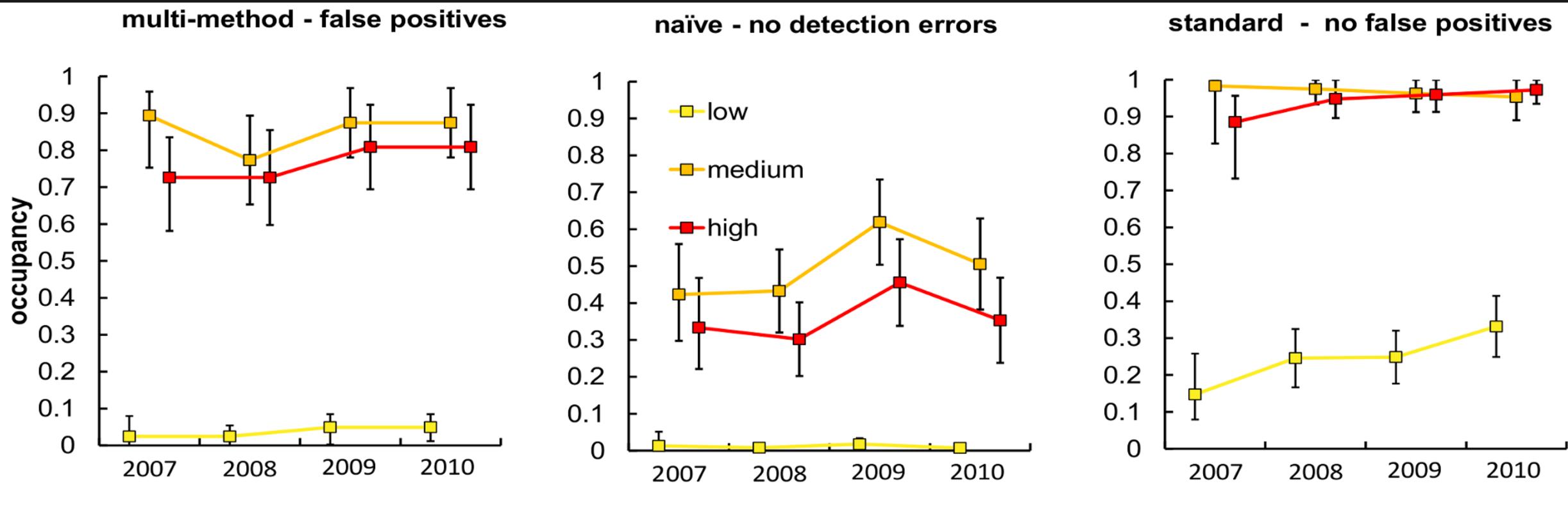
p_{11} = probability of detection

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/presence 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

ψ^A = prob. a site is occupied by species A

ψ^B = prob. a site is occupied by species B

ψ^{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 = ψ^{AB} / ψ^B

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

Venn diagram



Occupied by
species A only

$$\psi^A - \psi^{AB}$$



Occupied by
both species

$$\psi^{AB}$$



Occupied by
species B only

$$\psi^B - \psi^{AB}$$

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^{Ab} = prob. detecting species A but not B when both present

r^{aB} = prob. detecting species B but not A when both present

r^{ab} = prob. detecting neither species when both present

Initial states

Initial states

$$\begin{matrix} \text{U} & \text{A} & \text{B} & \text{AB} \\ \left[\begin{array}{cccc} 1 - \Sigma & \psi^A - \psi^{AB} & \psi^B - \psi^{AB} & \psi^{AB} \end{array} \right] \end{matrix}$$

State process

State process

$$\begin{array}{ccccc} & U & A & B & AB \\ U & \left[\begin{array}{cccc} 1 & 0 & 0 & 0 \end{array} \right] \\ A & \left[\begin{array}{cccc} 0 & 1 & 0 & 0 \end{array} \right] \\ B & \left[\begin{array}{cccc} 0 & 0 & 1 & 0 \end{array} \right] \\ AB & \left[\begin{array}{cccc} 0 & 0 & 0 & 1 \end{array} \right] \end{array}$$

Observation process

Observation process

$$\begin{array}{ccccc} & 0 & 1 & 2 & 3 \\ \text{U} & 1 & 0 & 0 & 0 \\ \text{A} & 1 - p^A & p^A & 0 & 0 \\ \text{B} & 1 - p^B & 0 & p^B & 0 \\ \text{AB} & 1 - \Sigma & r^{Ab} & r^{aB} & r^{AB} \end{array}$$

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)

$\eta = 1$ – 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...

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

*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†

**Proteus Wildlife Research Consultants, PO Box 5193, Dunedin, New Zealand;* †*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

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.
- Louvier, 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.
- 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.