

OCCUPANCY MODELING

**INTRODUCTION TO THEIR USE AND
APPLICATION IN ECOLOGY AND
EPIDEMIOLOGY**

Fidisoa Rasambainarivo, DVM, MSc
University of Missouri St Louis

UMSL



f.rasambainarivo@gmail.com



[@Fidydm](https://twitter.com/Fidydm)



www.fidyas.wordpress.com

MY 1-SLIDE, 1 MIN PRESENTATION

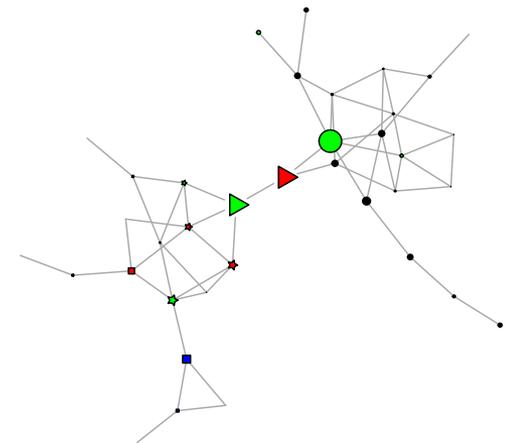
Interactions and disease transmission between carnivores in Betampona, Madagascar



Short communication

The global impacts of domestic dogs on threatened vertebrates

Tim S. Doherty^{a,*}, Chris R. Dickman^b, Alistair S. Glen^c, Thomas M. Newsome^{d,e,f,g}, Dale G. Nimmo^h, Euan G. Ritchie^{i,j}, Abi T. Vanak^{k,l}, Aaron J. Wirsing^g



OCCUPANCY MODEL: LECTURE AND TUTORIAL OBJECTIVES



- **Understand what are occupancy models**
- **Why would we use occupancy models**
- **How to design a study using occupancy modeling**
- **Working example:**
 - Prevalence of a parasite in a population with imperfect test.
- **Overview of other use of occupancy models**

WHY?

Abundance and distribution in ecology

- Used to assess/re-assess population trends

WHY?

Abundance and distribution in ecology

- Used to assess/re-assess population trends
- Often very difficult to know abundance
 - Capture Mark Recapture...

WHY?

Abundance and distribution in ecology

- Used to assess/re-assess population trends
- Often very difficult to know abundance
 - Expensive (CMR...)
- Sometimes presence/absence is just what we need...
 - Distribution and range
 - Animal invasions
 - Disease dynamics

WHY?

Abundance and distribution in ecology

- Used to assess/re-assess population trends
- Often very difficult to know abundance
 - Expensive (CMR...)
- Sometimes presence/absence is just what we need...
 - Distribution and range
 - Animal invasions
 - Disease dynamics
- Occupancy modeling: a *tool* to estimate true populations and monitor changes
 - But what do I mean by occupancy?

WHAT IS OCCUPANCY?

Probability that focal object (taxon) occupies or uses a sample unit during a specified period of time.

$$\psi = x/s$$

x = number of occupied sites

s = total number of sites

WHAT IS OCCUPANCY?

Probability that focal object (taxon) occupies or uses a sample unit during a specified period of time.

- Within disease ecology/epidemiology framework : what is the probability that a parasite is found in an individual
ie. what is the prevalence of the disease in a population?

Prevalence = Number of infected/population

WHAT IS OCCUPANCY?

Probability that focal object (taxon) occupies or uses a sample unit.

Taxon is said to be present if:

- Taxon is present

AND

- Taxon is detected

WHAT IS OCCUPANCY?

Probability that focal object (taxon) occupies or uses a sample unit.

Taxon is said to be present if:

- Taxon is present

AND

- Taxon is detected

What does absence data tell us?

- Taxon is absent

OR

- Taxon was not detected by investigator

ESTIMATE DISEASE PREVALENCE

For example:

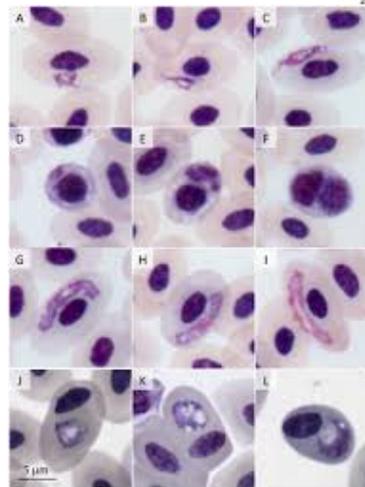
Researchers want to know the prevalence of *Plasmodium* sp. in birds of Galapagos?

Maricruz Jaramillo, Sage Rohrer and Patricia G. Parker University of Missouri – St. Louis, Department of Biology



Zenaidura macroura, *Setophaga ruticilla*, *Mimus parvulus*, *Geospiza fortis*

Photo credit: Jeisson Zamudio



ESTIMATE DISEASE PREVALENCE

Example:

Researchers want to know the prevalence of *Plasmodium* in sampled birds of Galapagos?

Collect a biological sample from each individual in a sample population and test it once to determine the presence of the parasite in the sample.

From Galapagos doves to passerines: Spillover of *H. multipigmentatus*

Maricruz Jaramillo, Sage Rohrer and Patricia G. Parker

University of Missouri – St. Louis, Department of Biology



Zenaidura macroura, *Setophaga ruticilla*, *Mimus parvulus*, *Geospiza fortis*

Photo credit: Jaisson Zamudio

Yes/No

x% (95% CI)

ESTIMATE DISEASE PREVALENCE

Example:

Researchers want to know the prevalence of *Plasmodium* in sampled birds of Galapagos?

From Galapagos doves to passerines: Spillover of *H. multipigmentatus*

Maricruz Jaramillo, Sage Rohrer and Patricia G. Parker

University of Missouri – St. Louis, Department of Biology



Zenaidura macroura, *Setophaga ruticilla*, *Mimus parvulus*, *Geospiza fortis*

Photo credit: Jaisson Zamudio

Collect a sample from each individual once and test it once to determine the presence of the parasite in the sample.

Yes/No

X%



Assuming that test is perfect and probability of detection is 100% in infected individuals (sensitivity of test is 100%)

ESTIMATE DISEASE PREVALENCE

Reality:

- Most tests are imperfect
- Result in occupancy/prevalence estimates that are biased low
- Underestimation of pathogen transmission rates
- Flawed predictions regarding infection dynamics and epidemic spread.



		DISEASE	
		Present	Absent
TEST	Positive	True positive a	False positive b
	Negative	False negative c	True negative d

ESTIMATE DISEASE PREVALENCE

Reality:

- Most tests are imperfect
- Result in occupancy/prevalence estimates that are biased low
- Underestimation of pathogen transmission rates
- Flawed predictions regarding infection dynamics and epidemic spread.

Need to correct

- Comparison of used test with a reference test
 - Several diagnostic methods need to be available.

OCCUPANCY MODELS

To estimate the proportion of all ‘sites’ that are occupied by the taxon, **with the explicit recognition that the taxon may be missed even if present during a survey.**

ECOLOGY
ECOLOGICAL SOCIETY OF AMERICA

[Explore this journal >](#)

Regular Article

ESTIMATING SITE OCCUPANCY RATES WHEN DETECTION PROBABILITIES ARE LESS THAN ONE

Darryl I. MacKenzie, James D. Nichols, Gideon B. Lachman, Sam Droege, J. Andrew Royle, Catherine A. Langtimm

First published: 1 August 2002 [Full publication history](#)

DOI: 10.1890/0012-9658(2002)083[2248:ESORWD]2.0.CO;2 [View/save citation](#)



[View issue TOC](#)
Volume 83, Issue 8
August 2002
Pages 2248–2255

OCCUPANCY MODELS

Estimated parameters and notation

- Probability that taxon is present at a site = Occupancy (ψ Psi)
= *prevalence*

OCCUPANCY MODELS

Estimated parameters and notation

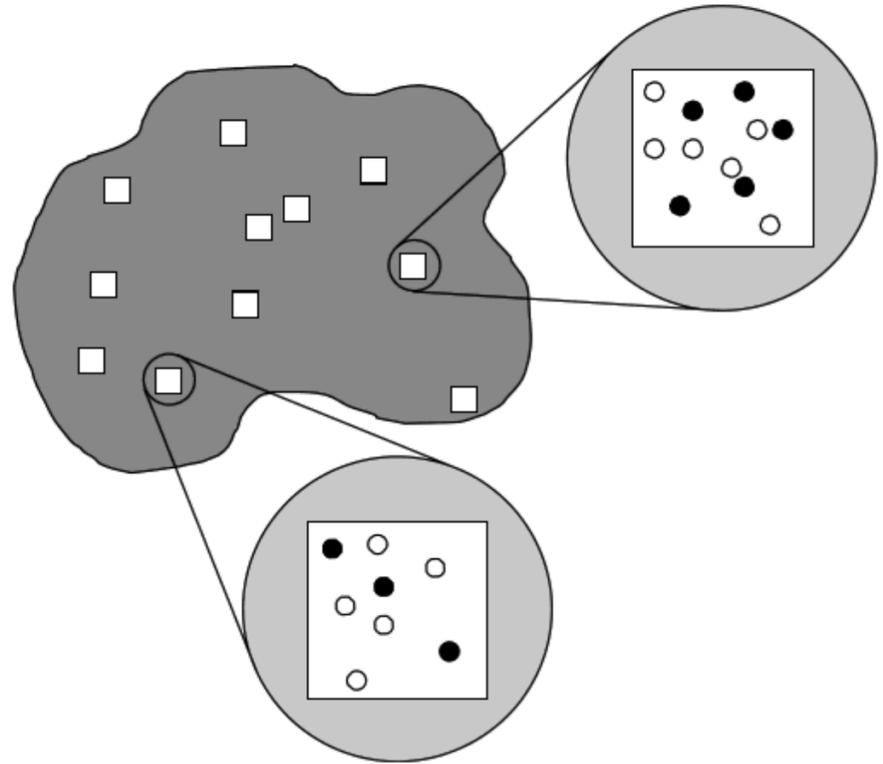
- Probability that it is detected when present (p)
= detection probability during survey
= *sensitivity of test*

- Probability that taxon is present = Occupancy (Ψ)
= prevalence

HOW DO I DESIGN A STUDY TO USE OCCUPANCY MODELS

Key is replicates:

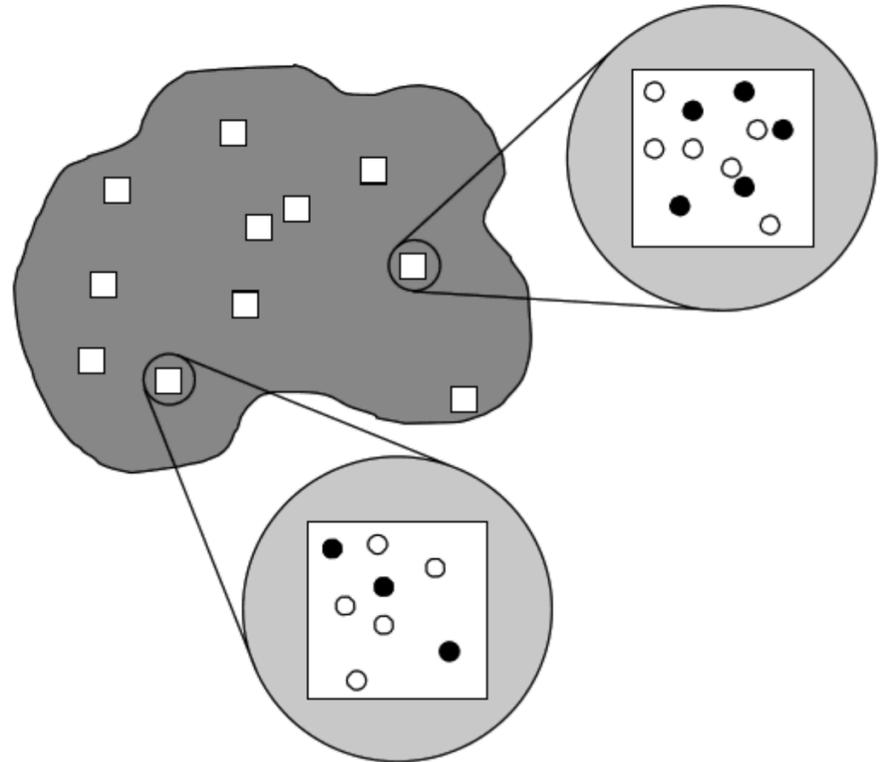
- Spatial: **multiple sites**
site : basic landscape unit at which the presence of the taxon is determined. Could be naturally occurring (pond, patch, etc...) or not



HOW DO I DESIGN A STUDY TO USE OCCUPANCY MODELS

Key is replicates:

- Spatial: **multiple sites**
site : basic landscape unit at which the presence of the taxon is determined. Could be naturally occurring (pond, patch, etc...) or not
- Temporal: **multiple visits to each site**
survey: a single opportunity for the detection of target species



OCCUPANCY MODELS: UNDER THE HOOD

- Record of whether a species was detected or not detected during each survey of each site (**detection histories**)
- Probability of a specific detection history e.g. 111, 011 or 000
- Maximum likelihood techniques are used to estimate parameters.

OCCUPANCY MODELS: UNDER THE HOOD

Example (1) survey 30 sites 4 times

	Survey 1	Survey 2	Survey 3	Survey 4
Site 1	1	0	0	1

We can write the probability of this detection history as:

$$\Pr(H_i = 1001) = \psi \times p_1 (1 - p_2)(1 - p_3) p_4$$

OCCUPANCY MODELS: UNDER THE HOOD

	Survey 1	Survey 2	Survey 3	Survey 4
Site 2	0	0	0	0
Site ..30	0	0	0	0

Sites 2 and 30 represent a case where the target species was never detected (detection history = 0000)

OCCUPANCY MODELS: UNDER THE HOOD

	Survey 1	Survey 2	Survey 3	Survey 4
Site 2	0	0	0	0
Site ..30	0	0	0	0

Sites 2 and 30 represent a case where the target species was never detected (detection history = 0000)

These sites could either be unoccupied, which mathematically is

$$(1-\psi),$$

or they could be occupied, but we never detected the target species, which mathematically is:

$$\psi(1-p_1)(1-p_2)(1-p_3)(1-p_4) \text{ or } \psi \prod_{j=1}^4 (1-p_j)$$

Thus, we can write the probability of detection history (0000) as:

$$\Pr(H_i = 0000) = \psi \prod_{j=1}^4 (1-p_j) + (1-\psi)$$

OCCUPANCY MODELS: UNDER THE HOOD

Mathematical statements of all detection histories are combined into model likelihood, such as:

$$L(\psi, p | H_1, \dots, H_30) = \prod_{i=1}^{30} \Pr(H_i)$$

Product of math equations forms the model likelihood for the observed data

Using MLE determine Ψ and p

OCCUPANCY MODELS

Assumptions:

- (1): Site is “closed”: species presence does not change during a sampling season (but can change between seasons- “multiple season”)

OCCUPANCY MODELS

Assumptions:

- (1): Site is “closed”: species presence does not change during a sampling season (but can change between seasons- “multiple season”)
- (2): Sites are independent: detection of species at one site is independent from detection at another site.

OCCUPANCY MODELS

Assumptions:

- (1): Site is “closed”: species presence does not change during a sampling season (but can change between seasons- “multiple season”)
- (2): Sites are independent: detection of species at one site is independent from detection at another site.
- (3): No false detection

ENOUGH TALK: WORKING EXAMPLE

Start up your Rstudio

- Working example:
Researchers want to know
:
 - What is the **prevalence** of *antibodies to Toxoplasma* in sampled galidinae of Madagascar given imperfect test?
 - Does sex, species influence the infection and detection of *Toxoplasma* in galidinae of Madagascar?



ENOUGH TALK: WORKING EXAMPLE

Working example:

Researchers want to know :

- What is the prevalence of Toxoplasma in sampled Galidinae of Madagascar given imperfect test?
- Does age and sex influence the infection and detection of antibodies to Toxoplasma in galidinae of Madagascar?

Material and methods

- 41 individual animals captured
- Blood samples
- ELISA for detection of antibodies to parasite
- Three replicate test for each blood sample

ENOUGH TALK: WORKING EXAMPLE



Journal of Statistical Software

August 2011, Volume 43, Issue 10.

<http://www.jstatsoft.org/>

**unmarked: An R Package for Fitting Hierarchical
Models of Wildlife Occurrence and Abundance**

Ian J. Fiske

North Carolina State University

Richard B. Chandler

USGS Patuxent Wildlife Research Center

ENOUGH TALK: WORKING EXAMPLE

```
#install unmarked if not installed
```

```
install.packages("unmarked")
```

```
#load library
```

```
library(unmarked)
```

```
#set here the directory where your data files are located and  
import data
```

```
setwd("~/Dropbox/E2M2/occupancy")
```

```
data<-read.csv("toxoc.csv")
```

ENOUGH TALK: WORKING EXAMPLE

#install unmarked if not installed

```
install.packages("unmarked")
```

#load library

```
library(unmarked)
```

#set here the directory where your data files are located and
import data

```
setwd("~/Dropbox/E2M2/occupancy")
```

```
data<-read.csv("toxos.csv")
```

```
View(data)
```

	AnimalID	y.1	y.2	y.3	weight	sex	species	cq1	cq2	cq3
1	1	1	1	1	0.800	M	galidia	1	2	2
2	2	1	1	0	0.770	M	galidia	2	2	0
3	3	0	0	0	0.760	F	galidictis	0	0	0
4	4	0	1	0	0.980	F	galidictis	0	2	0
5	5	1	0	1	1.100	F	galidia	1	0	1
6	6	0	0	1	0.890	F	galidictis	0	0	2
7	7	0	0	1	0.630	M	galidia	0	0	1
8	8	1	1	1	0.860	M	galidia	1	3	1
9	9	1	1	0	0.560	F	galidictis	4	2	0
10	10	1	1	1	1.030	F	galidictis	4	3	4

ENOUGH TALK: WORKING EXAMPLE

Each animal is a
site.

Each replicate is a
survey.

	AnimalID	y.1	y.2	y.3	weight	sex	species	cq1	cq2	cq3
1	1	1	1	1	0.800	M	galidia	1	2	2
2	2	1	1	0	0.770	M	galidia	2	2	0
3	3	0	0	0	0.760	F	galidictis	0	0	0
4	4	0	1	0	0.980	F	galidictis	0	2	0
5	5	1	0	1	1.100	F	galidia	1	0	1
6	6	0	0	1	0.890	F	galidictis	0	0	2
7	7	0	0	1	0.630	M	galidia	0	0	1
8	8	1	1	1	0.860	M	galidia	1	3	1
9	9	1	1	0	0.560	F	galidictis	4	2	0
10	10	1	1	1	1.030	F	galidictis	4	3	4

OCCUPANCY WITH “UNMARKED”

If we surveyed
each individual
once (y.1)

	n	Prevalence (95% CI)
<i>Galidia</i>	29	0.48 (0.30-0.67)
<i>Galidictis</i>	12	0.33 (0.11-0.65)
Males	14	0.36 (0.14-0.64)
Females	27	0.48 (0.29-0.68)
Total	41	0.44 (0.29-0.60)

SINGLE SEASON: DATA PREP

#####format dataframe to
unmarkedframe

Unmarked uses a specific type
of data frame (S4)

SINGLE SEASON: DATA PREP

```
#####format dataframe to unmarkedframe
```

```
#detection data rows are sites columns are detection replicates
```

```
y<-data[,2:4]
```

```
#here we indicate that column 2 to 4 corresponds to detection  
history
```

```
#site level (individual) covariates (species, sex and weight)
```

```
toxosite<-data[,5:7]
```

	AnimalID	y.1	y.2	y.3	weight	sex	species	cq1	cq2	cq3
1	1	1	1	1	0.800	M	galidia	1	2	2
2	2	1	1	0	0.770	M	galidia	2	2	0
3	3	0	0	0	0.760	F	galidictis	0	0	0
4	4	0	1	0	0.980	F	galidictis	0	2	0
5	5	1	0	1	1.100	F	galidia	1	0	1
6	6	0	0	1	0.890	F	galidictis	0	0	2
7	7	0	0	1	0.630	M	galidia	0	0	1
8	8	1	1	1	0.860	M	galidia	1	3	1
9	9	1	1	0	0.560	F	galidictis	4	2	0
10	10	1	1	1	1.030	F	galidictis	4	3	4

SINGLE SEASON: DATA PREP

#put everything together in unmarked data frame

#note that covariate can come from separate files

toxos <- unmarkedFrameOccu(y = y, siteCovs = toxos.site)

#summary of unmarked data frame

summary(toxos)

ENOUGH TALK: WORKING EXAMPLE

```
> #summary of unmarked data frame
```

```
> summary(toxo)
```

```
unmarkedFrame Object
```

```
41 sites
```

```
Maximum number of observations per site: 3
```

```
Mean number of observations per site: 3
```

```
Sites with at least one detection: 33
```

```
Tabulation of y observations:
```

```
  0    1 <NA>
63  60    0
```

```
Site-level covariates:
```

	weight	sex	species
Min.	:0.5000	F:27	galidia :29
1st Qu.:	:0.7400	M:14	galidictis:12
Median	:0.8500		
Mean	:0.8407		
3rd Qu.:	:0.9500		
Max.	:1.2000		

SINGLE SEASON: OCCUPANCY

Model/function structure

```
model<-occu(~detection_formula ~occupancy_formula,  
dataframe)
```

#Run model with no covariates

```
fm1<- occu(~1~1,toxo)
```

SINGLE SEASON: OCCUPANCY

Model/function structure

#Run model with no
covariates

```
fm1<- occu(~1~1,toxo)
```

Look at the output

```
fm1
```

```
> fm1

Call:
occu(formula = ~1 ~ 1, data = toxo)

Occupancy:
  Estimate    SE      z P(>|z|)
    2.04 0.751 2.71 0.00666

Detection:
  Estimate    SE      z P(>|z|)
    0.206 0.242 0.851 0.395

AIC: 172.1898
```

SINGLE SEASON: OCCUPANCY

Estimate for occupancy and detection are on logit scale and need transformation

```
backTransform(fm1,'det')
```

```
backTransform(fm1,"state")
```

'Det' for detection or 'state' for occupancy

```
> #back transformations
> backTransform(fm1,'det')
Backtransformed linear combination(s) of Detection estimate(s)

Estimate      SE LinComb (Intercept)
      0.551 0.0599      0.206          1

Transformation: logistic
> backTransform(fm1,"state")
Backtransformed linear combination(s) of Occupancy estimate(s)

Estimate      SE LinComb (Intercept)
      0.885 0.0766      2.04          1

Transformation: logistic
```

SINGLE SEASON: OCCUPANCY

Estimate for occupancy and detection are on logit scale
and need transformation back transformations

```
backTransform(fm1,'det')
```

```
backTransform(fm1,"state")
```

'Det' for detection or 'state' for occupancy

```
> backTransform(fm1,'det')
```

Backtransformed linear combination(s) of Detection
estimate(s)

Estimate	SE	LinComb	(Intercept)
0.551	0.0599	0.206	1

Transformation: logistic

ADDING COVARIATES

Does the species affect the probability of detection of the parasite?

```
fm2<-occu(~species~1, toxo)
```

Does the species and/or sex affect the probability of infection of the parasite in galidinae?

```
fm3<-occu(~1~species, toxo)
```

```
fm4<-occu(~1~sex, toxo)
```

```
fm5<-occu(~1~sex+species, toxo)
```

```
fm6<-occu(~species~species, toxo)
```

CHOOSING THE BEST MODEL

We now have 6 models:

Which one(s) is/are the best?

#####which of these models are the best

#built-in model selection

#labeled list of models

```
fmlist<-fitList("null"=fm1,  
               "psi_sp"=fm2,  
               "psi_sx"=fm3,  
               "psi_sx+sp"=fm4,  
               "p_sp"=fm5,  
               "psi_p_sp"=fm6)
```

#model selection -- ranked list of models and AIC

```
modSel(fmlist)
```

CHOOSING THE BEST MODEL

modSel(fmList)

	nPars	AIC	delta	AICwt	cumltvWt
null	2	172.19	0.00	0.285	0.28
psi_sp	3	172.66	0.47	0.225	0.51
psi_sx	3	173.12	0.93	0.179	0.69
p_sp	3	173.92	1.73	0.120	0.81
psisx+sp	4	174.10	1.91	0.110	0.92
psi_p_sp	4	174.66	2.47	0.083	1.00

ASSESSING FIT OF MODELS

**How valid are our
models?**

ASSESSING FIT OF MODELS

How valid are our models?

Simulate a dataset based upon a fitted model, refits the model, and evaluates a user-specified fit-statistic for each simulation.

Remember these equations?

$$\Pr(H_i = 101) = \psi \times p_1 (1 - p_2)^{p_3}$$

- We can now calculate “expected” frequencies of a 101 if we survey 41 sites

ASSESSING FIT OF MODELS

How valid are our models?

Simulate a dataset based upon a fitted model, refits the model, and evaluates a user-specified fit-statistic for each simulation.

Remember these equations?

$$\Pr(H_i = 101) = \psi \times p_1 (1 - p_2) p_3$$

- We can now calculate “expected” frequencies of a “101” if we survey 43 sites

Comparing this sampling “expected” distribution to the observed statistic provides a means of evaluating goodness-of-fit or assessing uncertainty in a quantity of interest.

ASSESSING FIT OF MODELS

```
63 #use a chisq and parametric bootstrapping to check the adequacy of model fit
64 - chisq<-function(fm) {
65     observed<-getY(fm@data)
66     expected <-fitted(fm)
67     sum((observed-expected)^2/expected)
68 }
69 (pb<-parboot(fm1,statistic=chisq, nsim=200, parallel=FALSE))
```

ASSESSING FIT OF MODELS

```
63 #use a chisq and parametric bootstrapping to check the adequacy of model fit
64 chisq<-function(fm) {
65   observed<-getY(fm@data)
66   expected <-fitted(fm)
67   sum((observed-expected)^2/expected)
68 }
69 (pb<-parboot(fm1,statistic=chisq, nsim=200, parallel=FALSE))
```

```
Call: parboot(object = fm1, statistic = chisq, nsim = 200, parallel = FALSE)
```

```
Parametric Bootstrap Statistics:
```

	t0	mean(t0 - t_B)	StdDev(t0 - t_B)	Pr(t_B > t0)
1 63		0.0946	5.89	<u>0.478</u>

```
t_B quantiles:
```

	0%	2.5%	25%	50%	75%	97.5%	100%
t*1	50	52	59	63	68	73	78

Fail to reject the null hypothesis: there is no evidence of lack of fit

SINGLE SEASON: OCCUPANCY

**We have our best
models that fit the data.**

What now?

- what is the probability of detection?
- what is the probability of occupancy

SINGLE SEASON: OCCUPANCY

Estimates are on log
scale

Transform on original
scale using built-in
function “backTransform”

‘det’ for detection or
‘state’ for occupancy

SINGLE SEASON: OCCUPANCY

Estimates are in logit scale.

- Can use “back transform” when no covariates
- Otherwise, “predict”

```
#example obtain estimates of occupancy  
according to sex where fm3<-occu(~1~sex)
```

```
#first create a new dataframe
```

```
nd <- data.frame(sex=c("M", "F"))
```

```
predict(fm3, type='state',newdata=nd)
```

SINGLE SEASON OCCUPANCY

		Tested once	Occupancy
	n	Prevalence (95% CI)	
<i>Galidia</i>	29	0.48 (0.30- 0.67)	0.83 (0.65-1.00)
<i>Galidictis</i>	12	0.92 (0.60- 1.00)	1.00 (0.95-1.00)
Males	14	0.36 (0.14- 0.64)	0.79 (0.54-1.00)
Females	27	0.48 (0.29- 0.68)	0.94 (0.78-1.00)
Total	41	0.44 (0.29- 0.60)	0.89 (0.72-1.00)

SINGLE SEASON DETECTION

	Test once	Test three times
P*	0.55	0.90

$$p^* = 1 - (1 - p)^t$$

Probability of missing (false negative)

Test once 45%

Test three times 10%

BEYOND SINGLE SPECIES- SINGLE SEASON

BEYOND

Other functions available in unmarked

**Taking into account the abundance of focal animals
Royle-Nichols (occuRN)**

Estimating abundance from repeated counts (pcount)

Multiple season:

**probability of extinction and colonization of a site
(Colext)**

BEYOND SINGLE SPECIES: ROYLE NICHOLS

Detection heterogeneity can occur over space (not time)

e.g. Abundance (density) of animals varies among sites -> more likely to miss animals on sites with few animals than those with many animals

In Royle Nichols model:

$$p = 1 - (1 - r)^N$$

r: probability of detection of each animal

BEYOND SINGLE SPECIES: ROYLE NICHOLS MODEL

e.g. Royle Nichols and avian malaria

Methods in Ecology and Evolution

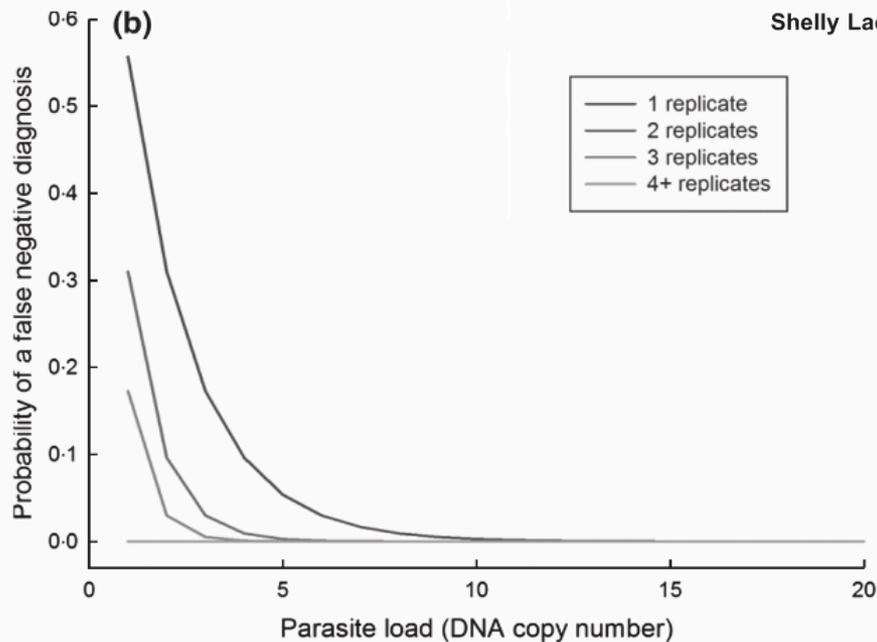


Methods in Ecology and Evolution 2012, 3, 339–348

doi: 10.1111/j.2041-210X.2011.00156.x

Site-occupancy modelling as a novel framework for assessing test sensitivity and estimating wildlife disease prevalence from imperfect diagnostic tests

Shelly Lachish^{1*}, Arjun M. Gopaldaswamy^{2,3,4}, Sarah C. L. Knowles^{1,5} and Ben C. Sheldon¹



BEYOND SINGLE SPECIES: ROYLE COUNT (N MIXTURE MODEL)

Fit a N mixture model

- **Ability to estimate abundance**

BEYOND SINGLE SPECIES: ROYLE COUNT (N MIXTURE MODEL)

Fit a N mixture model

- **Ability to estimate abundance**
- **Data frame preparation**
 - S4 dataframe, “unmarkedframePcount”
 - Need to specify count columns
 - `count<-data[,8:10]`

BEYOND SINGLE SPECIES: ROYLE COUNT (N MIXTURE MODEL)

Fit a N mixture model

- **Data frame preparation**
 - S4 dataframe, “unmarkedframePcount”
 - Need to specify count columns
 - `count<-data[,8:10]`
 - function is `pcount` (instead of `occu`)
 - Model structure is the same (double right-hand)
 - Distribution (poisson, Negative binomial, ZIP)

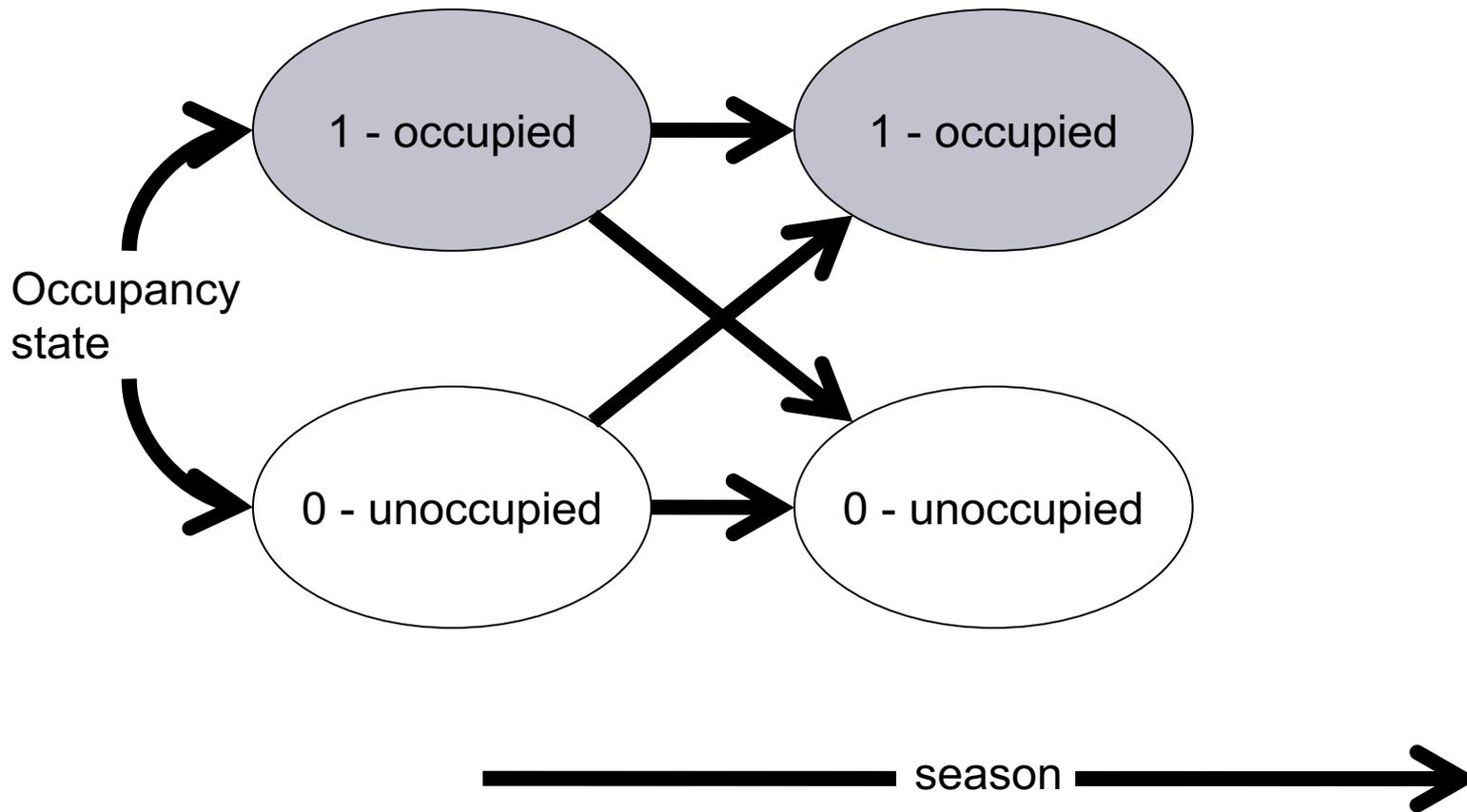
BEYOND SINGLE SPECIES: COLONIZATION-EXTINCTION

Dynamic occupancy models

Model the fate of an occupied/unoccupied site at $t+1$

OCCUPANCY DYNAMICS

$(1-\gamma)$ – not colonized



BEYOND SINGLE SPECIES: COLONIZATION-EXTINCTION

Example of detection history:

$$\Pr(10\ 00) = \psi p_1 (1 - p_2) (\varepsilon_1) + \psi p_1 (1 - p_2) (1 - \varepsilon) (1 - p_3) (1 - p_4)$$

- Occupied and detected on first, not detected on second and then unoccupied (extinct)
OR
- Occupied and detected on first, not detected on second and remained occupied but undetected on third and fourth.

BEYOND SINGLE SPECIES: COLONIZATION-EXTINCTION

Patch occupancy dynamics

Model changes in occupancy over time

Parameters of interest:

$$\psi_{t+1} = \psi_t \cdot (1 - \varepsilon) + (1 - \psi_t) \cdot \gamma.$$

ψ – probability of occupancy

ε_t – $Pr(\text{absence at time } t+1 \mid \text{presence at } t)$
– patch **extinction** probability

γ_t – $Pr(\text{presence at } t+1 \mid \text{absence at } t)$
– patch colonization probability

p_i – $Pr(\text{detection on occasion } i)$

BEYOND SINGLE SPECIES: COLONIZATION-EXTINCTION

Survey multiple sites multiple times during multiple seasons

Example:

Survey 267 sites during 9 years and 3 times each year.

Want to know the:

Probability of occupancy (ψ)

Probability of colonization (γ)

Probability of extinction (ϵ)

Probability of detection (p)

And factors affecting these estimates

COLONIZATION AND EXTINCTION

#####made up data survey 267 sites

#each site is surveyed 3x every year during 9 years (2010-2018)

#noted presence and absence each survey

#want to know if probability of occupancy varies according to forest cover

#want to know the probability of extinction and colonization of the taxon each year

```
data<-read.csv("varecia.csv")
```

```
View(data)
```

```
yvar <- as.matrix(data[,5:31])
years <- as.character(2010:2018)
years <- matrix(years, nrow(data), 9, byrow=TRUE) #create a
matrix to indicate the year(period) each site was survey
yearly
umf <- unmarkedMultFrame(y=yvar,
                          siteCovs=data[,2:3],
                          yearlySiteCovs=list(years=years),
                          numPrimary=9)
summary(umf)
```

Null model and More:

null dynamic occupancy model

```
fm0 <- colext(psiformula=~1,  
             gammaformula= ~1,  
             epsilonformula= ~1,  
             pformula= ~1, umf)
```

fm0

#null model where all estimates are constant

**#fit a couple more dynamic occupancy model with colonization
extinction and detection that are dependant on year**

```
fm1 <- colext(psiformula = ~1, # occupancy is constant  
  gammaformula = ~ years-1, # Colonization  
  epsilonformula = ~ years-1, # Extinction  
  pformula = ~ years-1, # Detection  
  data = umf)
```

```
fm2 <- colext(psiformula = ~forest, # Occupancy is forest dependent  
  gammaformula = ~ years-1, # Colonization  
  epsilonformula = ~ years-1, # Extinction  
  pformula = ~ years-1, # Detection  
  data = umf)
```

Model selection AS WE DID IT BEFORE:

#####MODEL SELECTION

```
models<-fitList(fm0,fm1,fm2)
```

```
modSel(models)
```

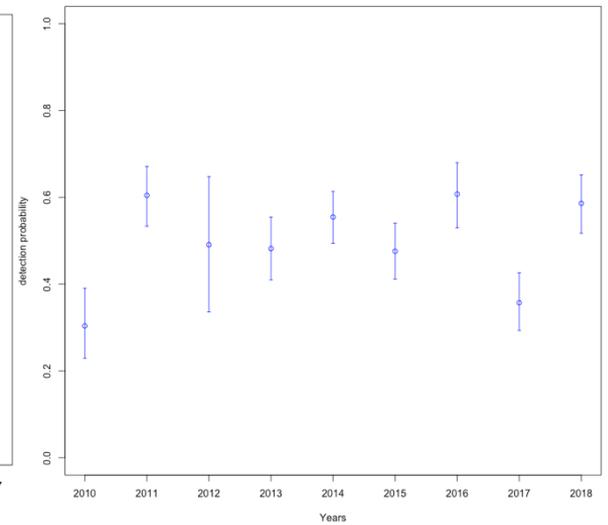
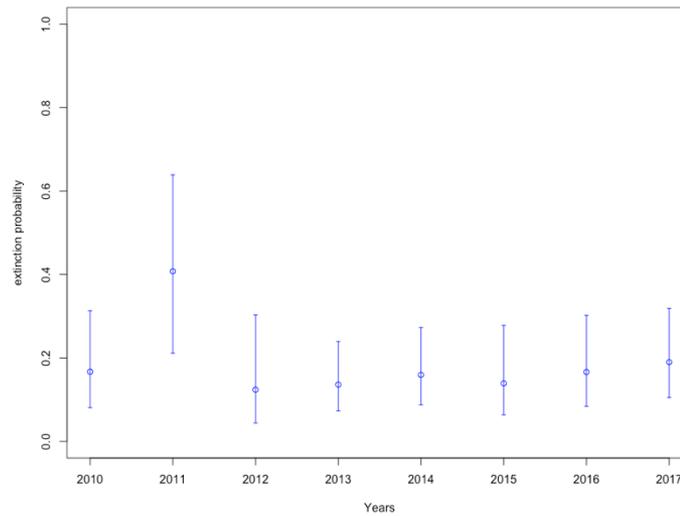
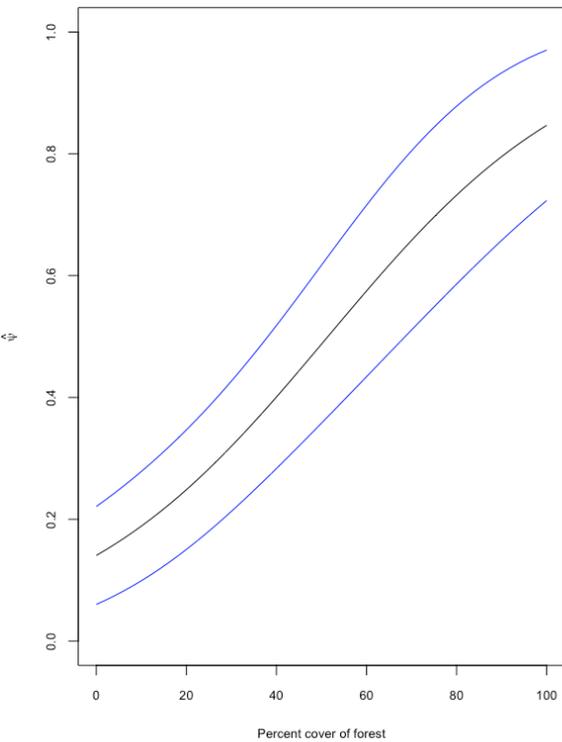
```
> modSel(models)
```

	nPars	AIC	delta	AICwt	cumltvWt
fm2	27	5217.20	0.00	1.0e+00	1.00
fm1	26	5250.92	33.72	4.8e-08	1.00
fm0	4	5318.93	101.74	8.1e-23	1.00

###let's say that fm2 (where occupancy varies with forest cover and colonization and extinction vary by year) is our best model, we could represent it graphically.

```
nd <- data.frame(forest=seq(0, 100, length=50))
E.psi <- predict(fm2, type="psi", newdata=nd, appendData=TRUE)
with(E.psi, {
  plot(forest, Predicted, ylim=c(0,1), type="l",
       xlab="Percent cover of forest",
       ylab=expression(hat(psi)), cex.lab=0.8, cex.axis=0.8)
  lines(forest, Predicted+1.96*SE, col=4)
  lines(forest, Predicted-1.96*SE, col=4)
})
```

PLOT DYNAMIC OCCUPANCY



SUMMARY TYPICAL WORK FLOW IN UNMARKED

PART I (import data and create unmarkedFrame)

```
pointCountData <- read.csv("myPointCountData.csv")
siteCovariates <- read.csv("mySiteCovariate.csv")
pointCountUMF <- unmarkedFramePCount(y=pointCountData,
  siteCovs = siteCovariates)
```

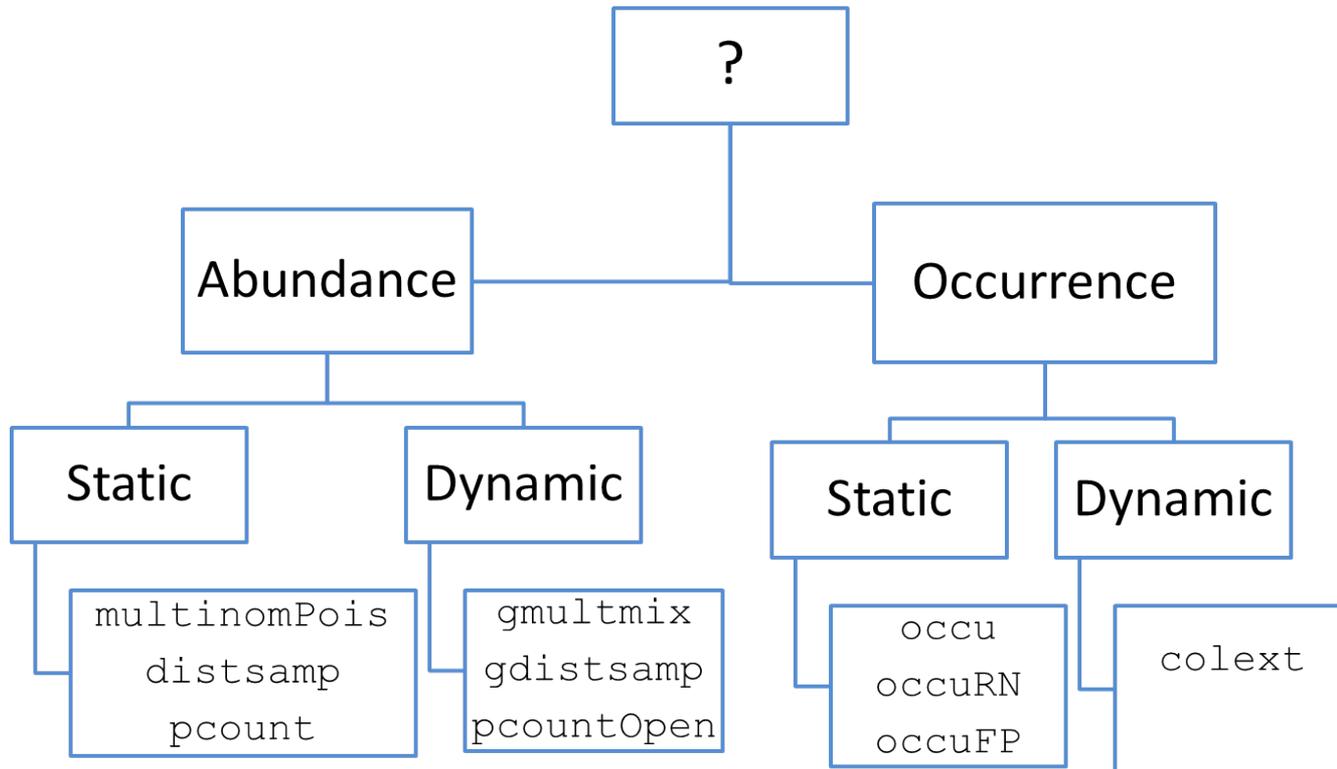
PART II (fit 2 models)

```
fm1 <- pcount(~1 ~1, data=pointCountUMF)
fm2 <- pcount(~vegHt ~vegHt, data=pointCountUMF)
```

PART III (model selection, model fit, prediction, mapping)

```
fms <- fitList(fm1, fm2)
modSel(fms) # model selection
parboot(fm2) # goodness of fit
predictions <- predict(fm2, type="state", newdata=mapdata)
levelplot(Predicted ~ x.coord+y.coord, data=predictions)
```

WHICH MODEL SHOULD I USE (UNMARKED)?



HOW DO I USE OCCUPANCY?

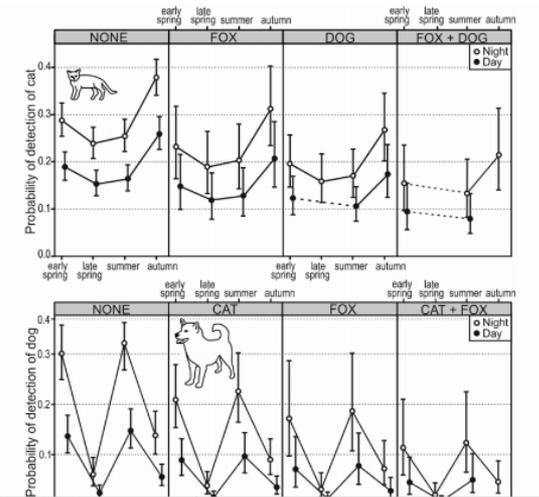
Interactions and disease transmission between carnivores in Betampona, Madagascar



Short communication

The global impacts of domestic dogs on threatened vertebrates

Tim S. Doherty^{a,*}, Chris R. Dickman^a, Alistair S. Glen^a, Thomas M. Newsome^{a,b,c,d}, Dale G. Nimmo^a, Euan G. Ritchie^a, Abi T. Vanak^{a,b,d}, Aaron J. Wirsing^e



RESOURCES

D.I. MacKenzie, J.D. Nichols, J.A. Royle, K.H. Pollock, L.L. Bailey, and J.E. Hines. 2006. Occupancy estimation and modeling. Academic Press. Burlington, MA.

D. I. MacKenzie, J. D. Nichols, J. E. Hines, M.G. Knutson, and A.B. Franklin. 2003. Estimating site occupancy, colonization, and local extinction when a species is detected imperfectly. Ecology 84:2200-2207.

Barbraud, C. , J. D. Nichols, J. E. Hines, and H. Hafner. 2003. Estimating rates of local extinction and colonization in colonial species and an extension to the metapopulation and community levels. – Oikos 101: 113–126.

