

Módulo 5

DETECCIÓN IMPERFECTA Y MODELOS DE OCUPACIÓN

Curso de Posgrado: “Modelado y estimación de ocupación para poblaciones y comunidades de especies bajo enfoque Bayesiano”

CCT CONICET Mendoza
24 - 28 Abril 2023



Instituto Nacional de
Tecnología Agropecuaria
Argentina



GTBA
Grupo Transdisciplinario de
Biodiversidad y Agroecosistemas

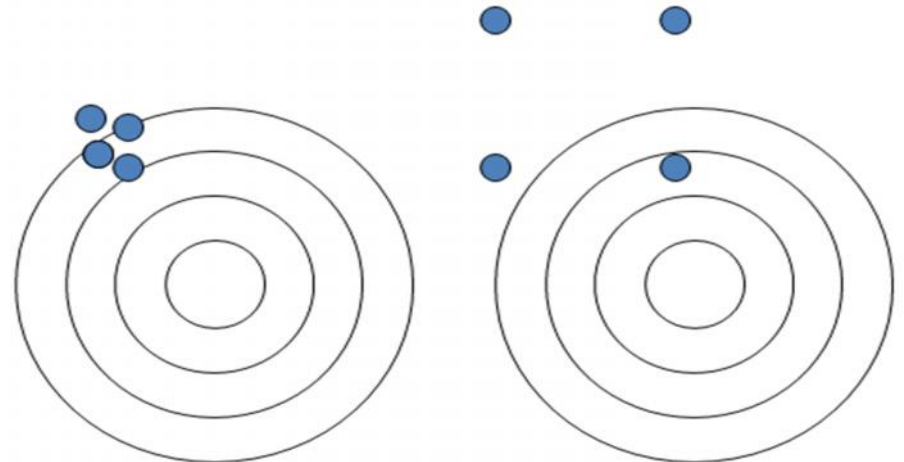


DETECCIÓN IMPERFECTA

- En prácticamente cualquier estudio es imposible asegurarse que todos los animales presentes sean detectados.
- Asumir detección total es hacer un “censo” de una población (o comunidad)
 - Todos los individuos son registrados sin error.
- Cuando hacemos un muestreo
 - No muestreamos el área completa
 - Perdemos registros de algunos animales presentes

DETECCIÓN IMPERFECTA

- Desafortunadamente muchos diseños de muestreo en conservación
 - Ignoran la detección incompleta de animales
 - Asumen que la probabilidad de detección es homogénea en espacio y tiempo.
- Ignorarla nos puede llevar a sesgos importantes y conclusiones falsas



DETECCIÓN IMPERFECTA

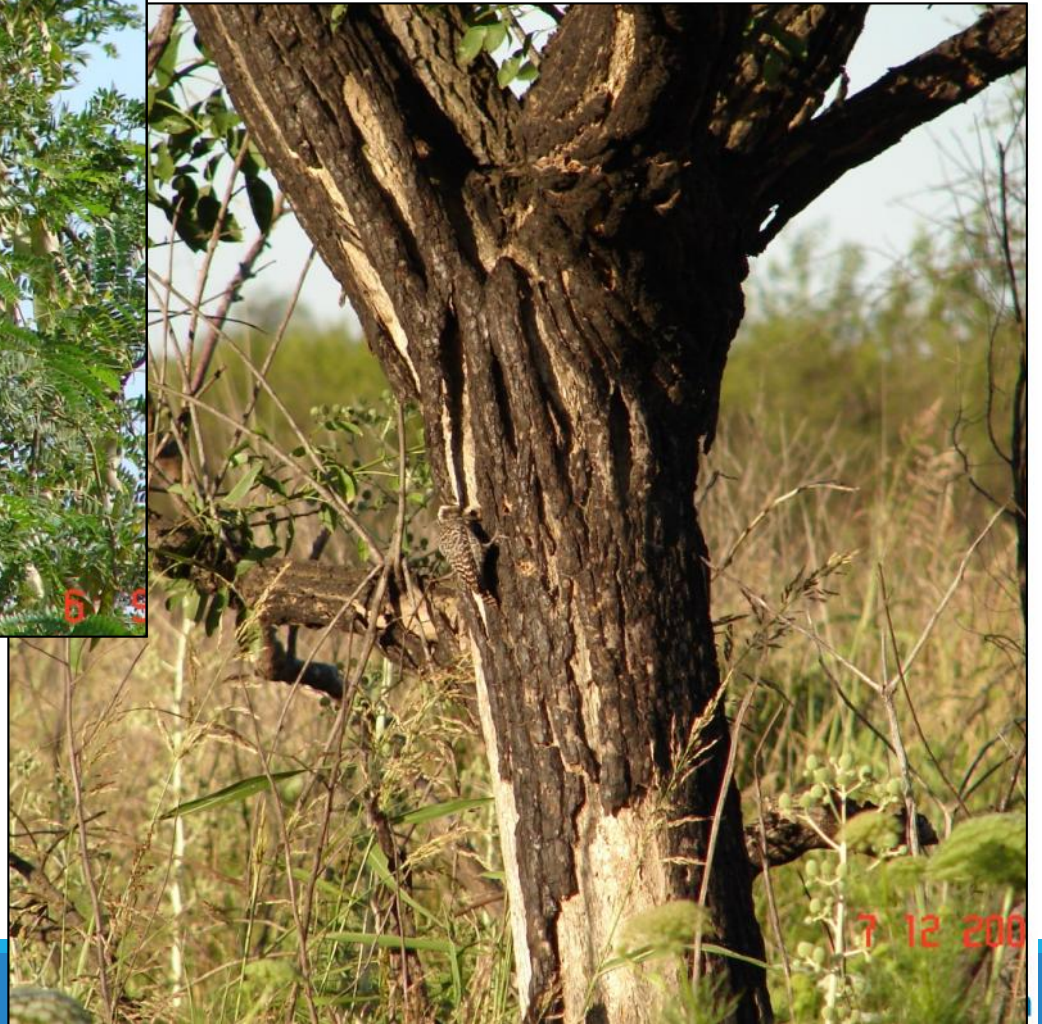
If a tree falls in the forest, but there is no one there to see it, does it make any noise?

Zen koan

If a bird sings in the forest, but the investigator fails to detect it, is the forest occupied?

Evan Cooch





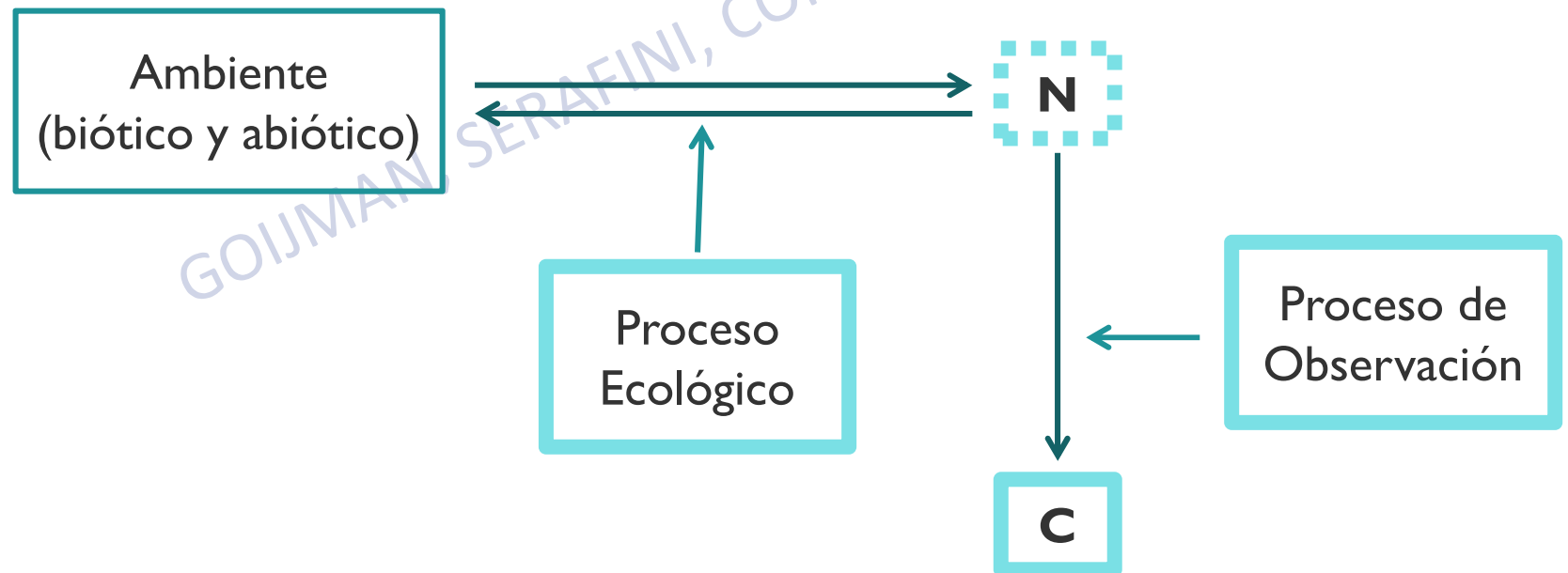


VISIÓN JERÁRQUICA DE LA ECOLOGÍA

Todas las observaciones son producto de dos procesos relacionados:

1^{er} submodelo Ecológico (nuestro interés)

2^{do} submodelo Observación (condicionante)



(Royle & Dorazio 2008, Kery & Schaub 2012)

PROCESO DE OBSERVACIÓN

2 componentes de error en el proceso de observación para conteos:

- “Individuos” no son registrados cuando están presentes
 - FALSOS NEGATIVOS
- “Individuos” identificados incorrectamente/ doble conteo
 - FALSOS POSITIVOS

GOIJMAN, CERAFIN, CONTRERAS 2023

PROCESO DE OBSERVACIÓN

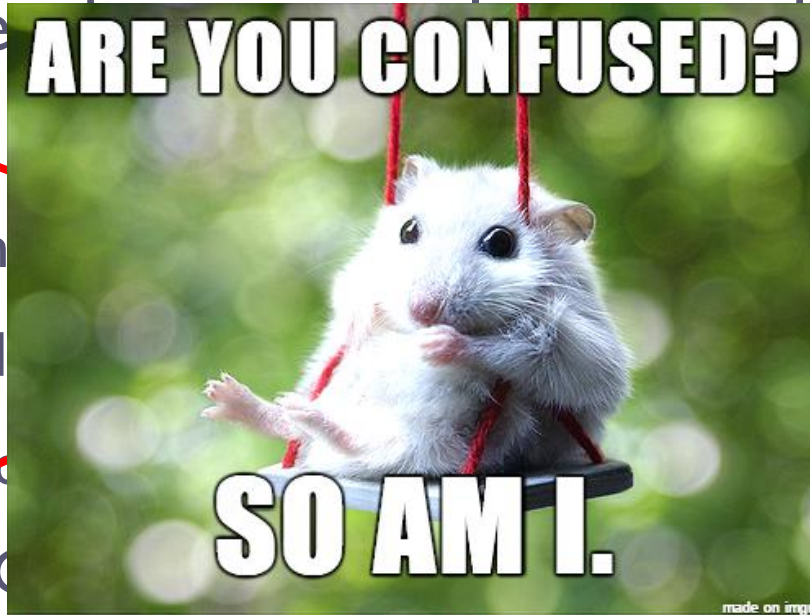
2 componentes de error en el proceso de observación para conteos:

- “Individuos” no son registrados cuando están presentes
 - FALSOS NEGATIVOS
- “Individuos” identificados incorrectamente/ doble conteo
 - FALSOS POSITIVOS
- “No existe”
- “No afecta a nuestro estudio”
- “Los errores se cancelan el uno a otro”

PROCESO DE OBSERVACIÓN

~~2 componentes de la observación para conteos:~~

- ~~– “Individuos” no presentes~~
 - ~~➤ FALSOS NEGATIVOS~~
- ~~– “Individuos” idóneos para el conteo~~
 - ~~➤ FALSOS POSITIVOS~~



- “No existe”
- “No afecta a nuestro estudio”
- “Los errores se cancelan el uno a otro”

PROCESO DE OBSERVACIÓN

- “Individuos” no son registrados cuando están presentes

- FALSOS NEGATIVOS

- └─ Modelado teniendo en cuenta la **probabilidad de detección** en los análisis

- “Individuos” identificados incorrectamente/ doble conteo

- FALSOS POSITIVOS

- └─ Controlable con buen diseño/ observadores
Aunque también puede ser modelado

ESTUDIOS QUE INCORPORAN PROBABILIDAD DE DETECCIÓN

Estudios Observacionales: Por lo general no se identifican individuos

- Registros de sp. (no individuos), rastros, sonidos, etc.
 - Presente y Detectado / Presente y No Detectado/ Ausente y No Detectado / Ausente y Detectado
- Registros de especies, rastros, etc. + Distancias
 - Probabilidad de detección en función de distancias

Estudios CMR (Captura-Marcado-Recaptura): Por lo general se identifican individuos

CLASIFICACIÓN y SELECCIÓN de MODELOS

- Objetivo
 - Distribución
 - Relación con hábitat/ paisaje / recursos
 - Abundancia
 - Supervivencia, otros parámetros demográficos
 - Tiempo
- Organismo
 - Puede ser capturado
 - Puede ser recapturado
 - Puede ser identificado a nivel individual
 - Marcas permanentes
- Disponibilidad de datos

CLASIFICACIÓN y SELECCIÓN de MODELOS

- POBLACION CERRADA

- Asume abundancia constante durante el estudio

- No hay ganancias ni pérdidas en la población
- Supuesto aproximado

- POBLACION ABIERTA

- Permite cambios en la abundancia durante el estudio

- Pérdidas (muertes o emigraciones) o ganancias (reproducción o inmigración)
- Foco de la estimación en tasas como las de supervivencia, reproducción, y abundancias

$$N(t + 1) = N(t) + B(t) + I(t) - D(t) - E(t)$$

	MÉTODOS POBLACIONES		PARÁMETROS
OBSERVACIÓN	OCUPACIÓN	ESTACION SIMPLE	Ocupación (Ψ), Detección (p)
		ESTACIONES MULTIPLES (POBLAC.ABIERTA)	Ψ , p , Colonización (), Extinción ()
		ROYLE-NICHOLS	Ψ , p , Abundancia (N)
	MODELOS N- MIXTOS	ESTACION SIMPLE	N , p
		ESTACIONES MULTIPLES (POBLAC.ABIERTA)	N , p , etc.
	MUESTREO DE DISTANCIAS	“DISTANCE SAMPLING”	Densidad, N
CMR	POBLACIONES CERRADAS		
	POBLACIONES ABIERTAS		

	MÉTODOS POBLACIONES
	MÉTODOS COMUNIDADES
	OCUPACIÓN

GOIJMAN, SERAFIN

Marc Kéry
J. Andrew Royle



APPLIED HIERARCHICAL MODELING IN ECOLOGY

Analysis of distribution, abundance and
species richness in R and BUGS

Volume 1
Prelude and Static Models



Marc Kéry
Michael Schaub



BAYESIAN POPULATION ANALYSIS USING **WinBUGS**

A hierarchical perspective



Marc Kéry
J. Andrew Royle



APPLIED HIERARCHICAL MODELING IN ECOLOGY

Analysis of distribution, abundance and
species richness in R and BUGS

Volume 1
Prelude and Static Models

Marc Kéry
Michael Schaub



Marc Kéry
J. Andrew Royle

APPLIED HIERARCHICAL MODELING IN ECOLOGY

Analysis of distribution, abundance and
species richness in R and BUGS

Volume 2
Dynamic and Advanced Models

ANALYSIS OF DISTRIBUTION
USING
BUGS
perspective



BASES OCUPACIÓN

- Definida como la proporción de área ocupadas por una especie o unidades de muestreo donde la especie está presente
- Requiere menor esfuerzo que estimación de abundancia (datos de presencia/ausencia)
- Útil en estudios a gran escala
- Útil en especies raras donde la abundancia no puede ser estimada
- No asume que las especies son contadas en cada muestreo (como la abundancia)

OCUPACIÓN

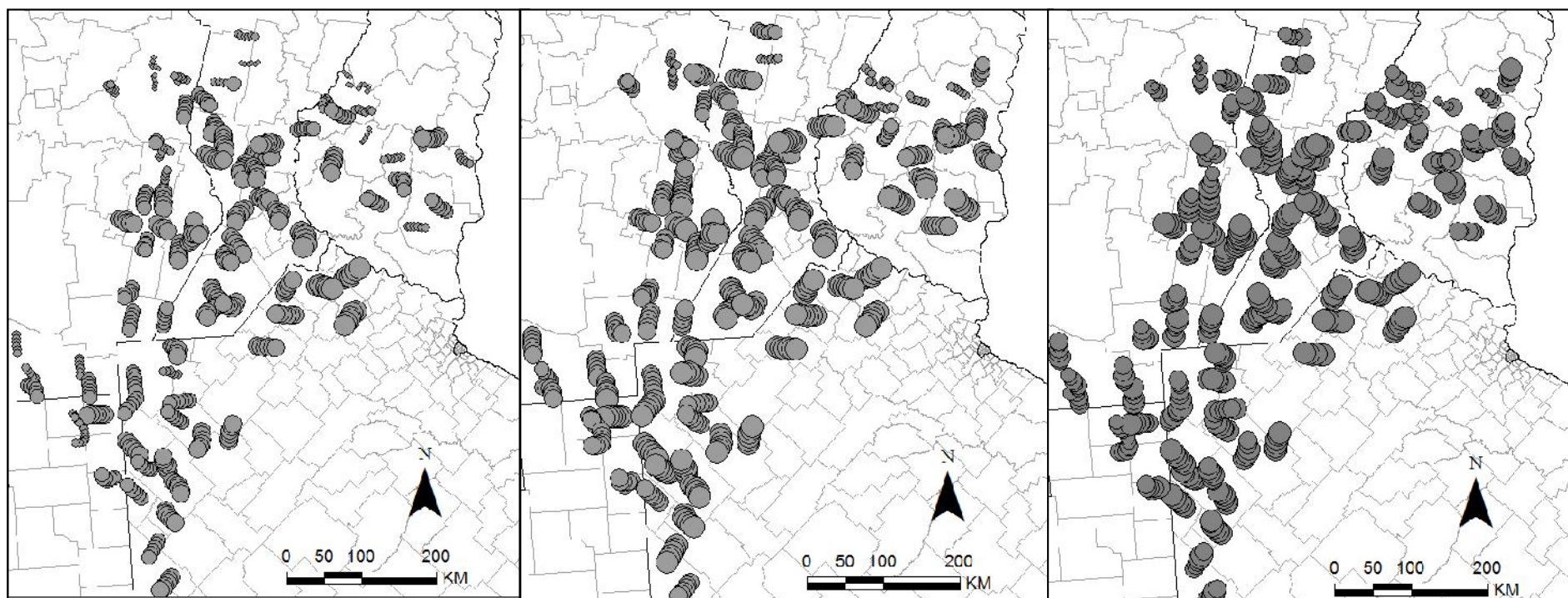
Preguntas ecológicas basadas en datos de ocupación:

- Dinámica de metapoblaciones
- Estudios de distribución y rangos geográficos
- Relaciones con el hábitat y selección de recursos
- Monitoreos de gran escala
- Interacción entre especies
- Epidemiología
- Etc

EJ. GRAN ESCALA



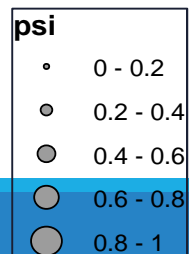
Columba picazuro



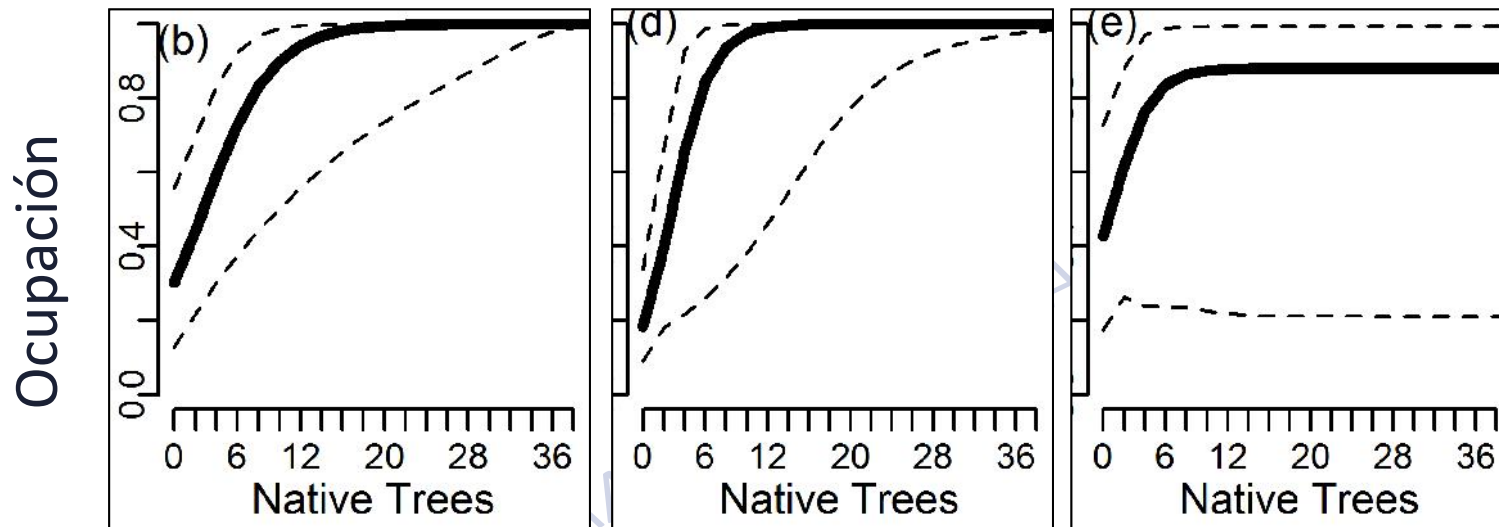
2006

2009

2012



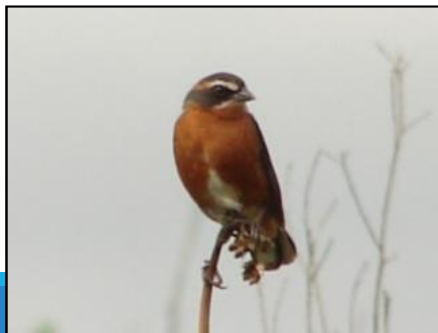
EJ. HABITAT



Granívoros de
follaje x espiguelo

Insectívoros de
follaje x espiguelo

Insectívoros
aéreos y de
vuelo elástico



OCUPACIÓN

Definiciones

ψ es la probabilidad que un sitio seleccionado al azar, o una unidad de muestreo en un área de interés, esté ocupado por una especie.

$$\hat{\psi} = \frac{\hat{x}}{s}$$

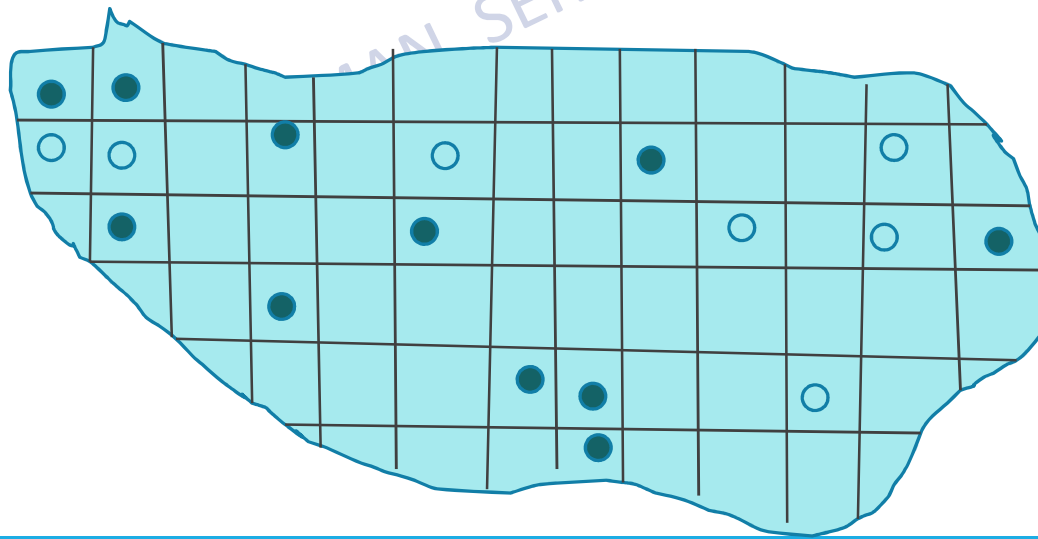
“naive” – subestiman ocupación

\hat{x} es el número de sitios ocupados, pero es estimado, porque número real de sitios “s” ocupados, **no lo vemos realmente**

OCUPACIÓN

$$\hat{\psi} = \frac{\hat{x}}{s}$$

\hat{x} es el número de sitios ocupados, pero es estimado, porque número real de sitios “s” ocupados, **no lo vemos realmente**



- Ocupado y detectado
- Ocupado y no detectado

OCUPACIÓN

- El tamaño y naturaleza de la unidad de muestreo puede ser definida natural o arbitrariamente.
- Generalmente nos interesan un número de unidades potenciales, asique la cantidad es una porción de un set más grande de unidades ocupadas.
- Interés puede estar en **la proporción de sitios ocupados** o en **la probabilidad de que una unidad esté ocupada**

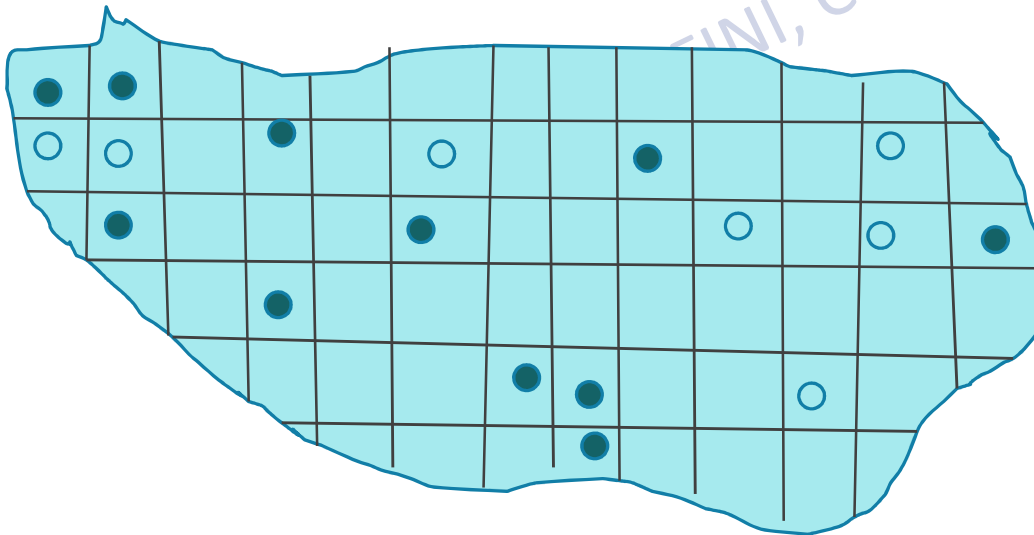
GOIJMAN, SERAFINI, CONTRERAS 2023

OCUPACIÓN

Proporción de área ocupada (hecho concreto, realización)

vs.

Probabilidad de ocupación (expectativa *a priori* de que un sitio esté ocupado)



- Ocupado y detectado
- Ocupado y no detectado

OCUPACIÓN

Proporción de área ocupada y probabilidad de ocupación

Cómo la probabilidad es desconocida, la proporción observada puede usarse para estimar esta probabilidad, y frecuentemente estos términos se usan indistintamente

GOIJMAN, SERAFINI, CONTRERAS 2023

OCUPACIÓN

- El protocolo básico de muestreo consiste simplemente en **revisitar sitios** y pasar tiempo buscando individuos de una especie deseada o evidencias de la presencia de la especie.
- Este tipo de muestreo generalmente es referido como presencia-ausencia (pero es detección-no detección)

- Visual
- Capturas en trampas o redes
- Huellas o rastros
- Vocalizaciones
- Trampas cámara



Enfoques de muestreo

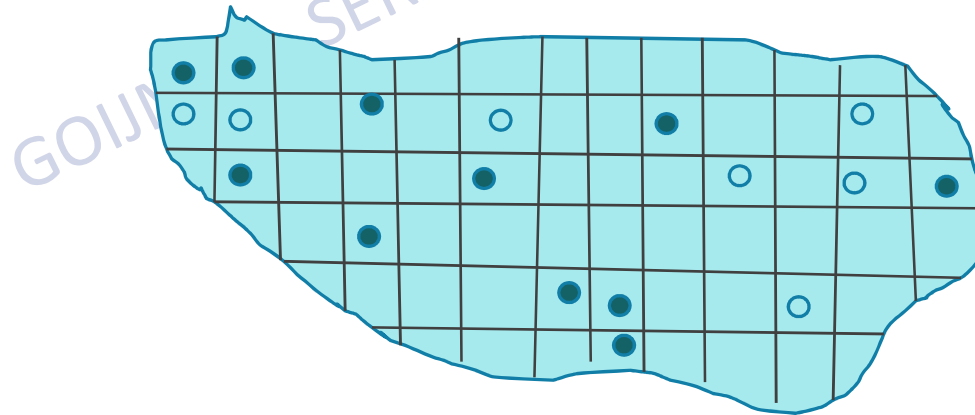
OCUPACIÓN

- El resultado del muestreo es una lista de unidades “ocupadas” (especie no detectada) o “desocupadas” (especie detectada)
- Historicamente se ha considerado el total de sitios ocupados como proporción de ocupación sobre el total de sitios y esto se tomaba como estimación de ocupación

¿Cuál es el problema de esto?

OCUPACIÓN

- Podríamos no detectar una especie cuando esta se encuentra presente.
- Sitios “ocupados” podrían estar siendo clasificados como “desocupados”



OCUPACIÓN

Parámetros

- probabilidad de ocupación
- p probabilidad de detección

Estimadores

- $\hat{\Psi} = \frac{s_d}{s}$ proporción de sitios con detecciones/sitios
(detección perfecta)
- $\hat{\Psi} = \frac{s_d}{s\hat{p}}$ proporción de sitios con detecciones/sitios
(detección imperfecta)

OCUPACIÓN

Diseño

- Plantear bien el objetivo
- Pensar que podría estar afectando al ocupación
- La definición del sitio depende del organismo y de la homegeneidad del sitio
- Réplicas por sitio para estimación de detectabilidad
 - En el tiempo
 - Múltiples observadores
 - Replicas espaciales

OCUPACIÓN

Número de sitios vs numero de réplicas

TABLE 6.1 Optimum Number of Surveys to Conduct at Each Site (K) for a Standard Design Where All Sites Are Surveyed an Equal Number of Times with No Consideration of Survey Costs, for Selected Values of Occupancy (ψ) and Detection Probabilities (p)

p	ψ								
	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.1	14	15	16	17	18	20	23	26	34
0.2	7	7	8	8	9	10	11	13	16
0.3	5	5	5	5	6	6	7	8	10
0.4	3	4	4	4	4	5	5	6	7
0.5	3	3	3	3	3	3	4	4	5
0.6	2	2	2	2	3	3	3	3	4
0.7	2	2	2	2	2	2	2	3	3
0.8	2	2	2	2	2	2	2	2	2
0.9	2	2	2	2	2	2	2	2	2

Source: MacKenzie and Royle (2005).

OCUPACIÓN

Modelos avanzados

- Estaciones múltiples, dinámicos (Diseños Robustos)
- Abundancia
- Multi-estado
- Multi-escala
- Multi-especies

COLIMAN, SERAFINI, CONTRERAS 2023

OCUPACIÓN

Historia de detección/sitio	Probabilidad
101	$p_1(1-p_2)p_3$
111	$p_1p_2p_3$
100	$p_1(1-p_2)(1-p_3)$
110	$p_1p_2(1-p_3)$
...	
000	$(1-p_1)(1-p_2)(1-p_3) + (1-)$

La historia de detección se construye por sitios, y no por individuo

OCUPACIÓN

- Sabemos cuantos sitios existen
- Queremos saber cuantos están ocupados
- Modelo del binomial inflado por ceros (Zero Inflated Binomial)

$$P[\{x_i\} | \Psi, p] = \prod_{i=1}^s \left[\Psi_i \binom{k_i}{x_i} p^{x_i} (1-p)^{k_i-x_i} + (1-\Psi_i) I(x_i = 0) \right]$$

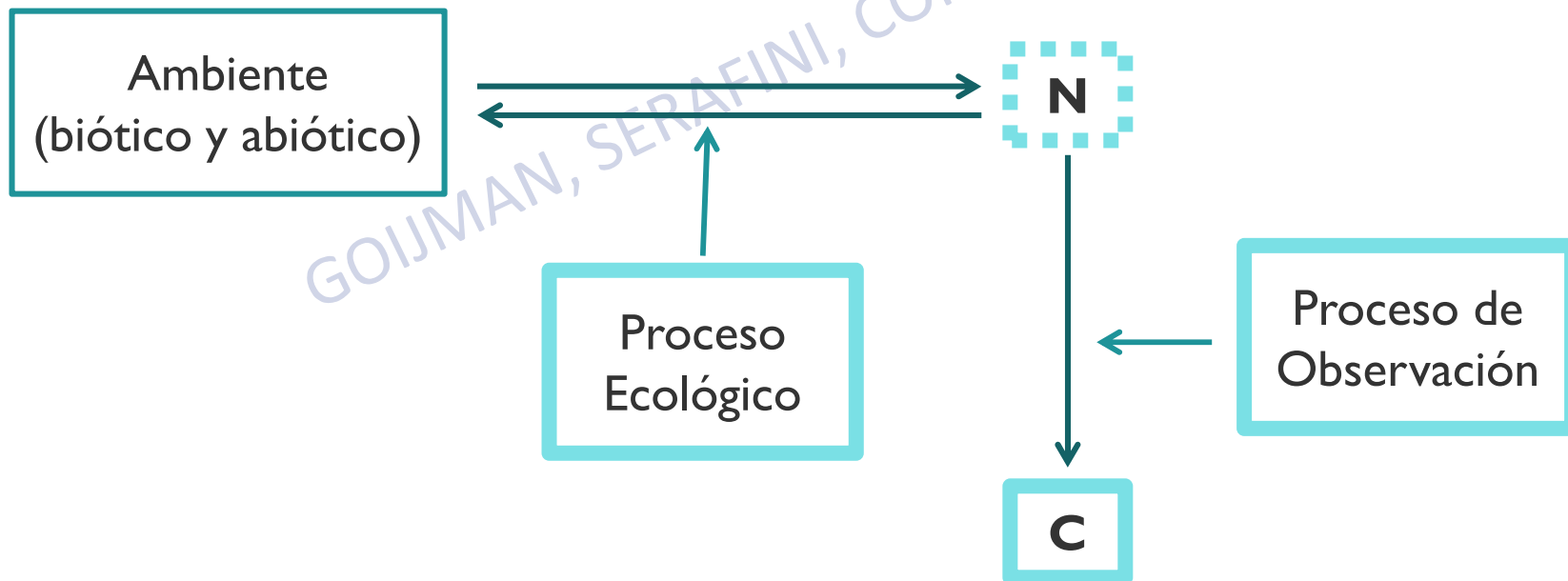
LIKELIHOOD

VISIÓN JERARQUICA DE LA ECOLOGÍA

Todas las observaciones son producto de dos procesos relacionados:

1^{er} submodelo Ecológico (nuestro interés)

2^{do} submodelo Observación (condicionante)



(Royle & Dorazio 2008, Kery & Schaub 2012)

MODELOS JERÁRQUICOS (“*State-Space*”)

Conteos en el tiempo

Proceso de estado

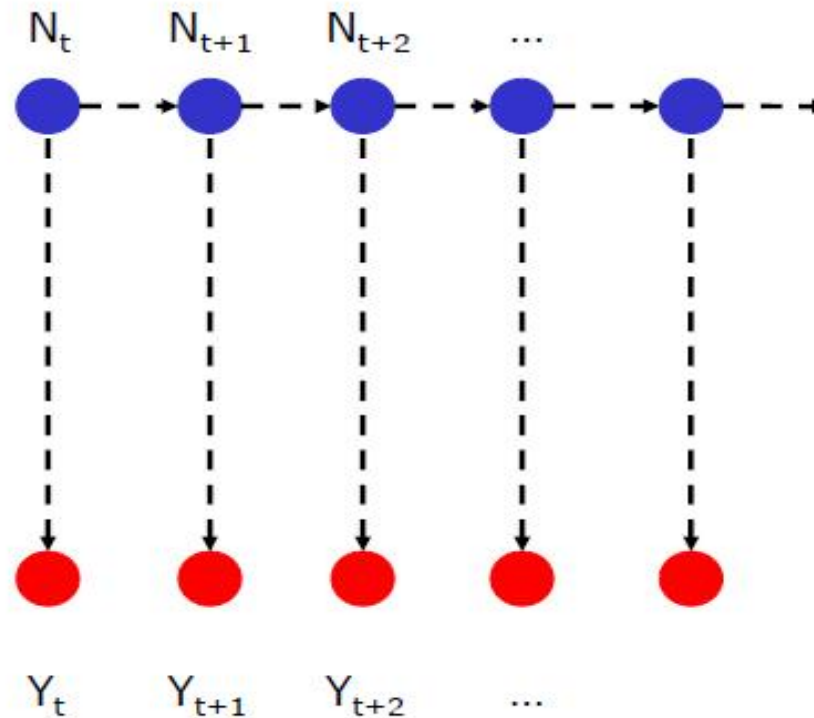
State process

True population size

Proceso de observación

Observation process

Observed count



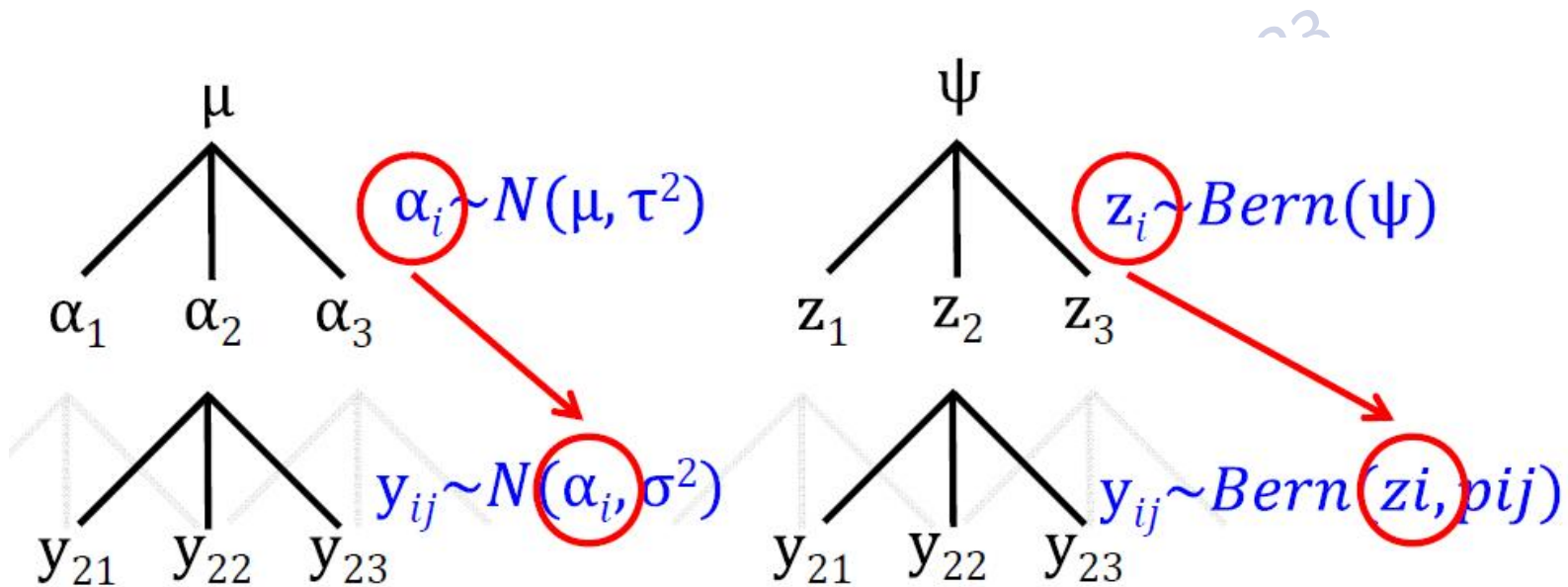
MODELOS JERÁRQUICOS (“*State-Space*”)

Compuesto de 2 sets de ecuaciones:

- Ecuación de estado: describe el estado real de un sistema
- Ecuación de observación: relaciona el el estado real a la observación
- Marco muy general y flexible
 - Datos de Captura-recaptura
 - Datos de Ocupación

MODELOS JERÁRQUICOS

- Secuencia dependiente de variables aleatorias (observadas y no observadas)



Ocupación como Modelo Jerárquico

Esencialmente los modelos de ocupación de sitio (vistos como “*site-occupancy models*”) son **modelos de regresión logística jerárquicos** que modelan en conjunto la probabilidad de detección y ocupación de animales o plantas.

Ocupación – el caso más simple

Una especie
Estación simple

Probabilidad que un sitio seleccionado al azar en un área esté ocupado por una especie

$z = 1$ Sitio está ocupado

$z = 0$ Sitio no ocupado

$z_i \sim \text{Bernoulli}(\psi)$

Proceso ecológico

$y_{ij} | z_i \sim \text{Bernoulli}(z_i * p)$

Proceso de observación

$y = \text{datos}$

$p = \text{probabilidad de detección}$

sitio i , réplica j

OCUPACIÓN

Una especie – Estación simple

Supuestos

- **El estado de ocupación de los sitios no cambia durante el período de muestreo (población cerrada - “closure assumption”)**

Durante la duración de los muestreos, el estado de ocurrencia en un sitio no cambia. Un sitio está ocupado o desocupado, no puede cambiar (no hay colonización o extinción). El período de tiempo considerado cerrado dependerá de la especie.

Si hay emigraciones temporarias durante el período, tampoco es “desastrozo”. Se debe considerar la ocupación como “*probabilidad de uso durante un período de tiempo*” durante el periodo de estudio.

También se puede modelar (multi-escala, diseño robusto, ocupación dinámica)

OCUPACIÓN

Una especie – Estación simple

Supuestos

- **No hay falsos positivos**

La especie, cuando es detectada, es detectada con certeza (también puede modelarse)

- **La ocurrencia y la detección de una especie son independientes** (“*independence assumption*”)

- La probabilidad de ocupación de un sitio es independiente de la de otro sitio (se puede explicar con covariables)
- La probabilidad de detección es independiente en sucesivas visitas a un mismo sitio (puede ser modelado)

- **La detección es homogénea** (puede ser modelado)

OCUPACIÓN

Incorporación de covariables

- La pb. de ocupación es igual para todos los sitios
 - La pb. de detectar una sp. en el muestreo, dado que está presente, es igual para todos los sitios
- En realidad ésto no siempre es probable, de hecho, muchas veces, es es mismo foco de estudio:
- “¿Cómo varía la ocupación de una especie en distintos sitios dado x condiciones?”*
- “¿Varía la detectabilidad de la especie?”*

Características biológicas de las especies

Detectabilidad

Edad

Movimiento

Sexo

Densidad

Hora del día

Comportamiento de los grupos

Cambios estacionales

Sistema social

Características del hábitat

Estructura de la vegetación

Composición de la vegetación

Topología del paisaje

Ruidos

Factores meteorológicos

Precipitaciones

Vientos

Temperatura

Niebla

Nieve

Humedad relativa



RECORDEMOS - GLM Binomial

Binomial

$$\varphi_i \sim \text{Binomial}(N_i, p_i)$$

$$\text{logit}(p_i) = \alpha + \beta * x_i$$

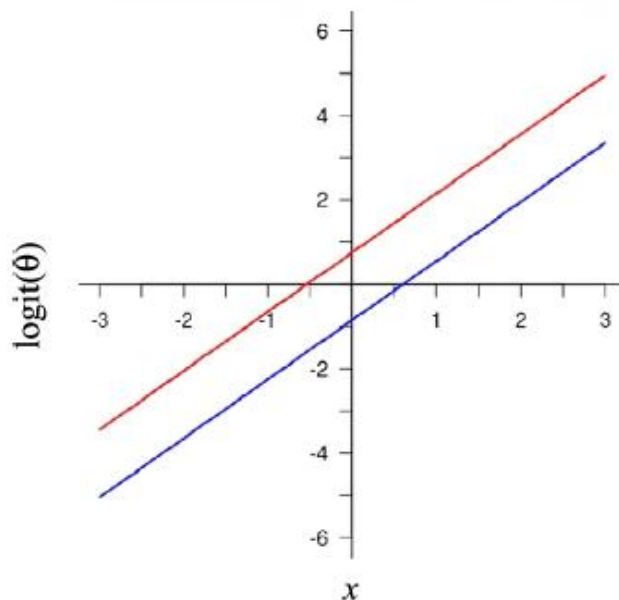
Función logit

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

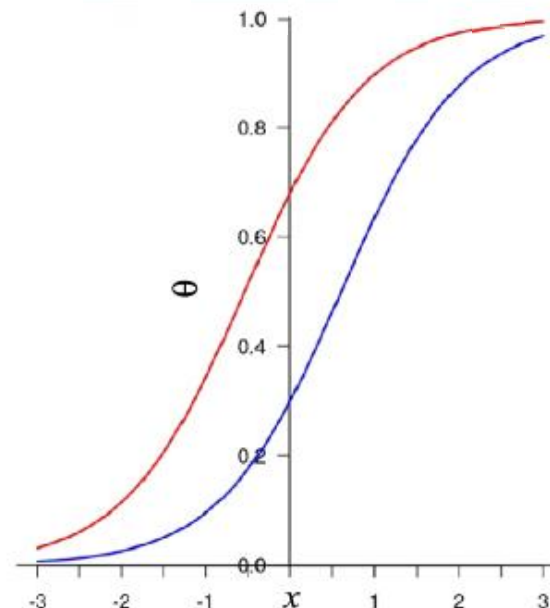
N_i es el número de intentos

p_i probabilidad de éxito (sólo dos eventos)

$$\text{logit}(\theta_i) = \alpha + \beta * x_i$$



$$\theta_i = \exp(\alpha + \beta * x_i)$$



OCUPACIÓN

Incorporación de covariables

- Covariables de detección
 - Pueden variar entre ocasiones para un mismo sitio
 - Características de la ocasión (ej. Variables climáticas, observador), o del sitio (ej. vegetación que afecta la detectabilidad)
- Covariables de ocupación
 - Constantes por sitio, dentro de una estación
 - Características del sitio (ej. habitat)

OCUPACIÓN

Incorporación de covariables

$z_i \sim \text{Bernoulli}(\psi)$

Proceso ecológico

$y_{ij} | z_i \sim \text{Bernoulli}(z_i * p)$

Proceso de observación

$$\text{logit}(\psi_i) = \alpha_{\text{psi}} + \beta_{x1} * x1_i$$

$$\text{logit}(p_{ij}) = \alpha_p + \beta_{x2} * x2_{ij}$$

OCUPACIÓN

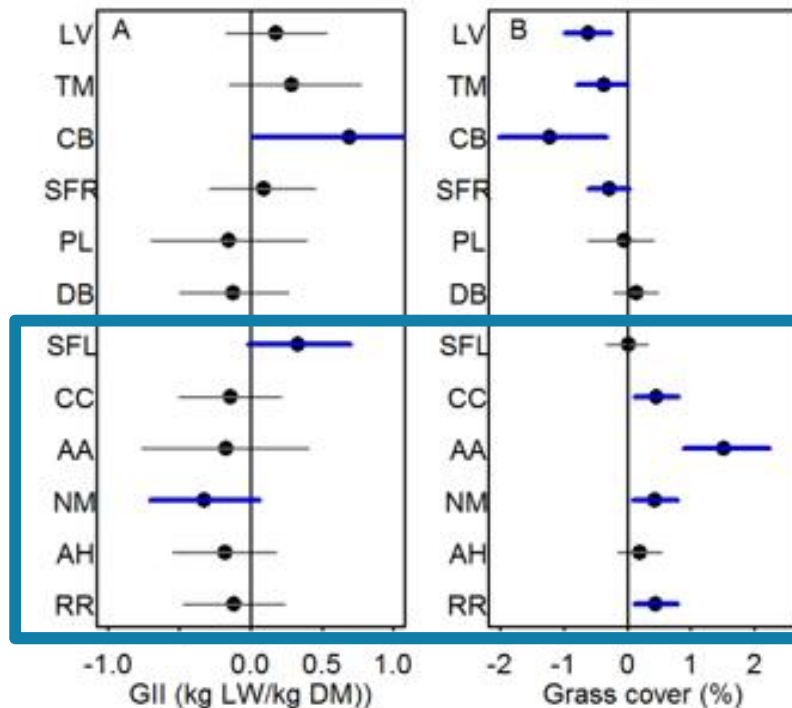
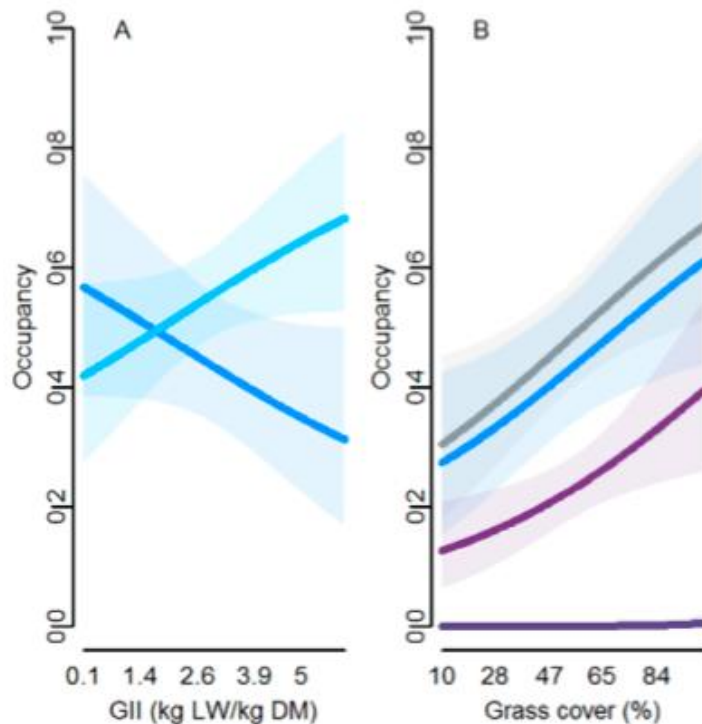
Incorporación de covariables

- Regresión logística

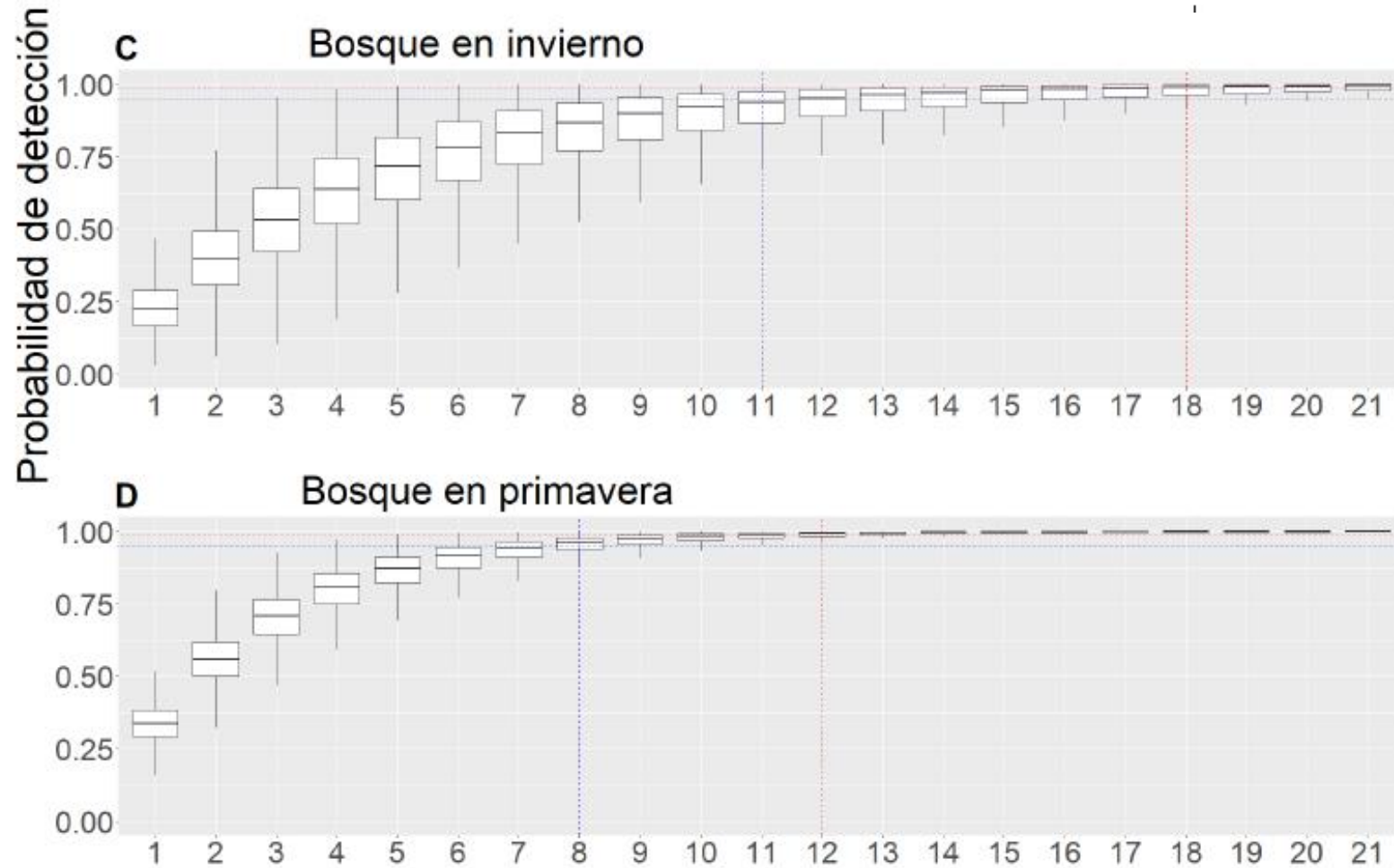
$$\text{logit}(\psi_i) = \alpha_{\text{psi}} + \beta_{x1} * x1_i$$

$$\text{logit}(p_{ij}) = \alpha_p + \beta_{x2} * x2_{ij}$$

(Dardanelli et al. 2022)



$$P^* = 1 - (1 - p)^n$$



(Jusim et al. 2023)

Figura III-4. Probabilidad de detección según esfuerzo de muestreo (medido en cámaras-noche) con CRI95%. Las líneas punteadas horizontales marcan un valor $\alpha=0,01$ (roja) y $\alpha=0,05$ (azul). Las líneas segmentadas verticales indican el número de cámaras-noche necesarias para que la probabilidad de no detectar castores si es que están presentes sea al valor α correspondiente al color. Figuras **A-D** corresponden a ambiente de bosque y figura **E** al ambiente de ecotono.

```
model {
```

Priors

```
psi ~ dunif(0, 1)
```

```
p ~ dunif(0, 1)
```

Likelihood

Modelo Ecológico para ocurrencia real

```
for (i in 1:R) {
```

```
z[i] ~ dbern(psi)
```

Modelo de Observación para observaciones replicadas detección/no detección

```
for (j in 1:T) {
```

```
y[i,j] ~ dbern(z[i] * p ) # y condicional a z=1 (observac condicional a la presencia)
```

```
  } #j repeticiones
```

```
} #i sitios
```

GOIJMAN, SERAFINI, CONTRERAS 2023

```
model {
```

Priors

```
alpha.occ ~ dunif(-10, 10)
```

```
beta.occ ~ dunif(-10, 10)
```

```
p ~ dunif(0, 1)
```

Likelihood

Modelo Ecológico para ocurrencia real

```
for (i in 1:R) {
```

```
logit(psi[i]) <- alpha.occ + beta.occ * COV[i]
```

```
z[i] ~ dbern(psi[i])
```

Modelo de Observación para observaciones replicadas detección/no detección

```
for (j in 1:T) {
```

```
y[i,j] ~ dbern(z[i] * p) # y condicional a z=1 (observac condicional a la presencia)
```

```
  } #j repeticiones
```

```
} #i sitios
```

REFERENCIAS

- <http://www.vogelwarte.ch/de/projekte/publikationen/bpa/>
- Dardanelli, S., Calamari, N. C., Canavelli, S. B., Barzan, F. R., Goijman, A. P., & Lezana, L. (2022). Vegetation structure and livestock grazing intensity affect ground-foraging birds in xerophytic forests of Central-East Argentina. *Forest Ecology and Management*, 521, 120439.
- Jusim, P., Goijman, A. P., Schiavini, A. *submitted*. Searching for remnants: Verifying for beaver eradication using camera traps.
- Kéry, M. 2009. Introduction to WinBUGS for Ecologists. Academic Press.
- Kéry, M., and M. Schaub. 2012. Bayesian population analysis using WinBUGS: a hierarchical perspective. Access Online via Elsevier.
- Link, W. A. and Richard J. Barker. 2010. Bayesian Inference with Ecological Applications. Academic Press.
- Mackenzie, D. I., and J. A. Royle. 2005. Designing occupancy studies: general advice and allocating survey effort. *Journal of Applied Ecology* 42: 1105–1114.
- MacKenzie, D. I., J. D. Nichols, A. R. Royle, K. H. Pollock, L. L. Bailey, and J. E. Hines. 2006. Occupancy estimation and modeling: inferring patterns and dynamics of species occurrence. Elsevier/Academic Press, Burlington, MA.
- MacKenzie, D. I., J. D. Nichols, G. B. Lachman, S. Droege, J. Andrew Royle, and C. A. Langtimm. 2002. Estimating site occupancy rates when detection probabilities are less than one. *Ecology* 83:2248–2255.
- MacKenzie, D. I., Nichols, J. D., Royle, J. A., Pollock, K. H., Bailey, L. L., & Hines, J. E. 2017. Occupancy estimation and modeling: inferring patterns and dynamics of species occurrence. Elsevier.
- Royle, J. A., and M. Kéry. 2016. Applied Hierarchical Modeling in Ecology: Analysis of distribution, abundance and species richness in R and BUGS: Volume 1: Prelude and Static Models Books. Academic Press.
- Royle, J. A., and R. M. Dorazio. 2008. Hierarchical modeling and inference in ecology: the analysis of data from populations, metapopulations and communities. Academic Press.
- Tyre, A. J., B. Tenhumberg, S. A. Field, D. Niejalke, K. Parris, and H. P. Possingham. 2003. Improving Precision and Reducing Bias in Biological Surveys: Estimating False-Negative Error Rates. *Ecological Applications* 13: 1790–1801.
- Williams, B., J. Nichols, and M. Conroy. 2002. Analysis and Management of Animal Populations: Modeling, Estimation, and Decision Making. Academic Press, San Diego, CA.