Hierarchical models for distribution and abundance

Marc Kéry

AHM Workshop at
MSU, East Lansing,
22–26 July 2024

Fundamental questions in population/community ecology

- How many are there ? -> abundance
- Where are they ? -> species distribution/abundance
- Spatial, temporal and spatiotemporal patterns in Dist&Abu ? -> Trends, maps
- What environmental drivers determine Dist&Abu ? -> niche modeling
- What demographic drivers determine Dist&Abu ? -> survival/recruitment, colonization/exinction
- Statistical associations in Dist&Abu of multiple species ? -> Species interactions, …

Fundamental questions in population/community ecology

- How many are there ? -> abundance
- Where are they ? -> species distribution/abundance
- Spatial, temporal and spatiotemporal patterns in Dist&Abu ? -> Trends, maps
- What environmental drivers determine Dist&Abu ? -> niche modeling
- What demographic drivers determine Dist&Abu ? -> survival/recruitment, colonization/exinction
- Statistical associations in Dist&Abu of multiple species ? -> Species interactions, …

- Programs JAGS/NIMBLE/Stan and the R packages `unmarked/ubms/spOccupancy/spAbundance` let you to tackle these questions via the fitting of hierarchical statistical models (HMs)

=> what are HMs and what is distribution and abundance ?

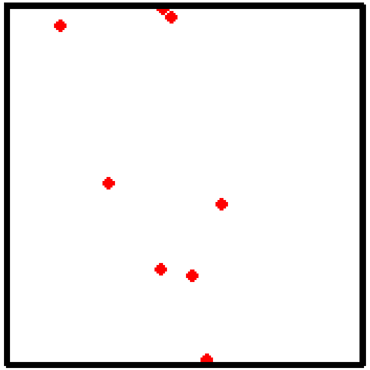
Overview

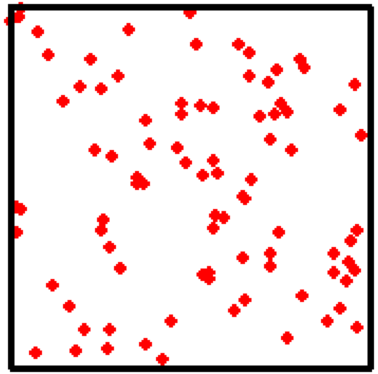
- What are hierarchical models ?
- What is distribution (z) ? …. abundance (N) ? … relationship between the two ?
  - Point patterns (PPs) and measurement errors in PPs
  - Discretization of space & summarization of PPs: yields Dist and Abu
  - Deterministic relationship between Dist and Abu
  - Measurement errors when assessing Dist&Abu
  - HMs for Dist&Abu

(Use of "*Experimental Statistics*": simulate stuff to "*see how things are*")

Fundamental model for animal/plant pops: point patterns

- PP: Outcome of point process, random process that produces "points" in 1D, 2D, 3D, …
- Both number of points and locations are random
- PP described in terms of intensity: limiting expected density of points
- e.g. Homogeneous PP (HPP)

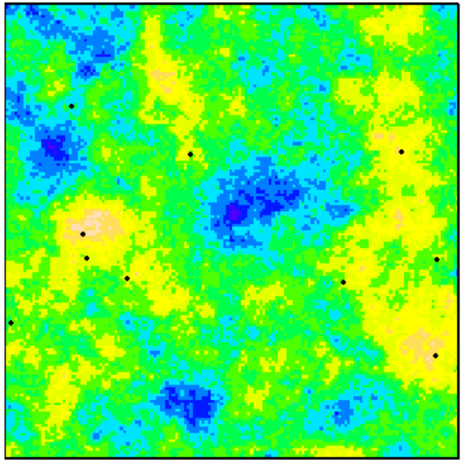


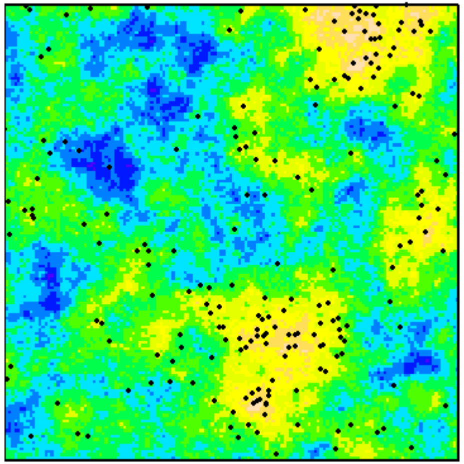


`AHMbook::sim.fn()`

Fundamental model for animal/plant pops: point patterns

- e.g. Inhomogeneous PP (IPP)

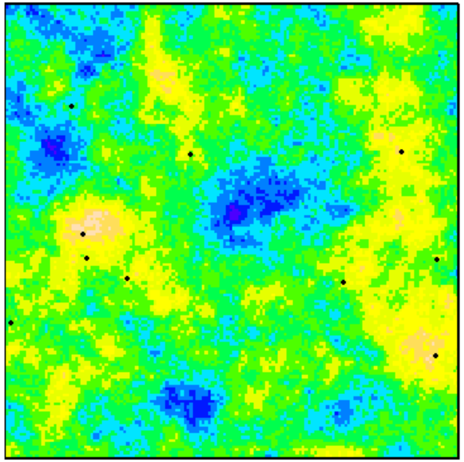


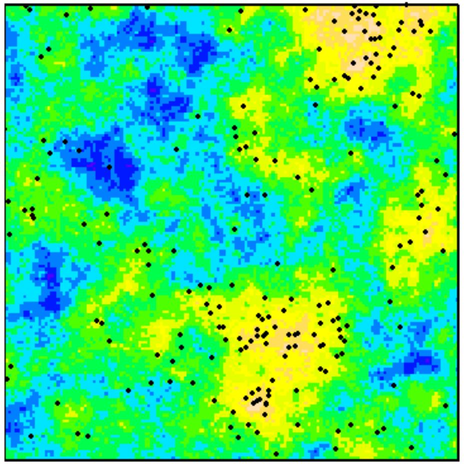


AHMbook::simPPe()

Fundamental model for animal/plant pops: point patterns

- Three types of measurement errors for PPs
- ?, ?, ?

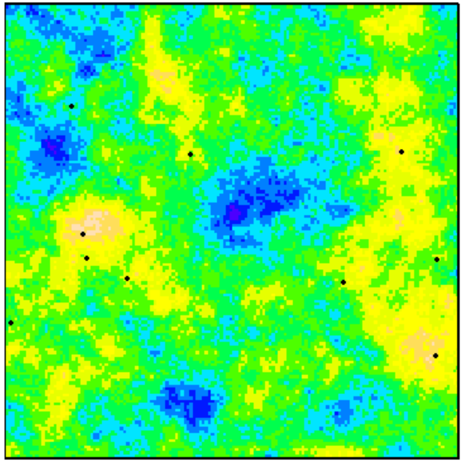


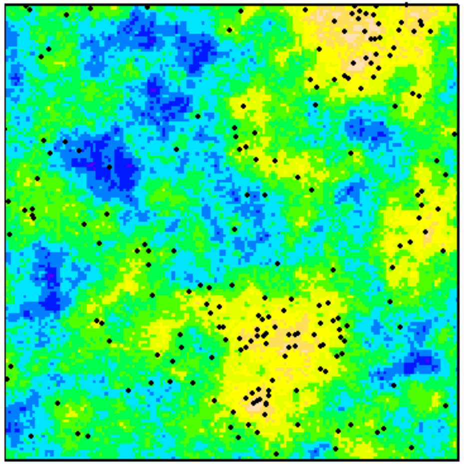


AHMbook::simPPe()

Fundamental model for animal/plant pops: point patterns

- Three types of measurement errors for PPs
- Location error, false negatives, false positives (also mark errors)





`AHMbook::simPPe()`

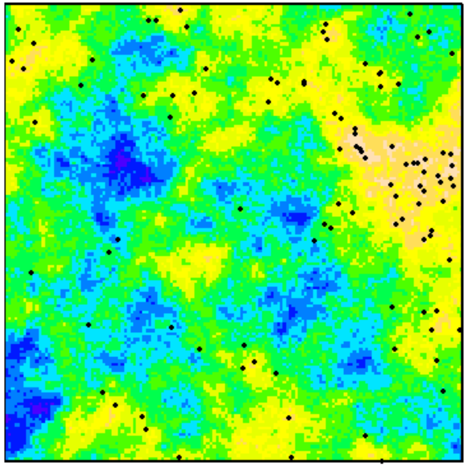
Why don't we do all pop./comm. ecology with PPs ?

- Location may be hard to measure precisely for every individual
- PPs defined in continuous space: conceptually harder, especially for non-statisticians,
  e.g., contain integrals, and biologists hate integrals
- Plus, humans like to put things into boxes
- Data may already come in some aggregated form (e.g., administrative units)

=> most times we work with data for discretized space

What happens when we discretize space ?

Point pattern with
core and buffer area

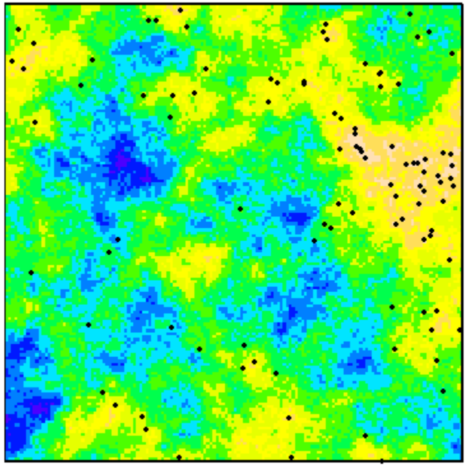


Mean intensity (lambda) = 0.00444

What happens when we discretize space ?

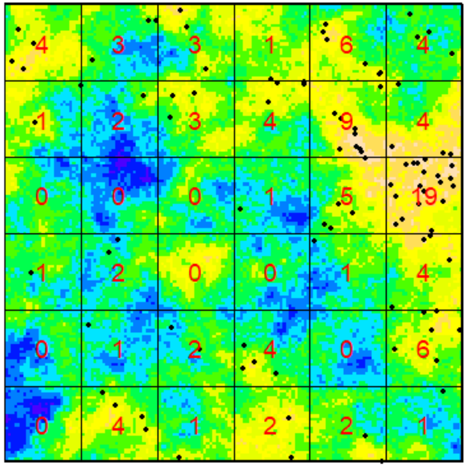
- Abundance and Distribution/Occurrence/Pres-Abs: discrete-space summaries of PPs

Point pattern with
core and buffer area



Mean intensity (lambda) = 0.00444

Abundance, N

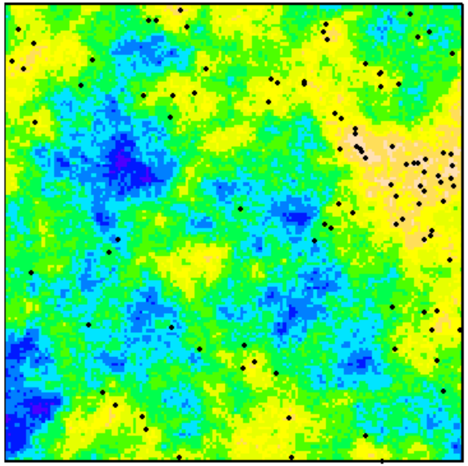


Mean(N) = 2.78, var(N) = 12.23

What happens when we discretize space ?

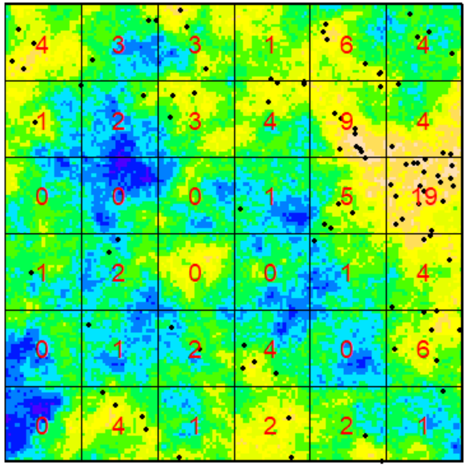
- Abundance and Distribution/Occurrence/Pres-Abs: discrete-space summaries of PPs

Point pattern with core and buffer area



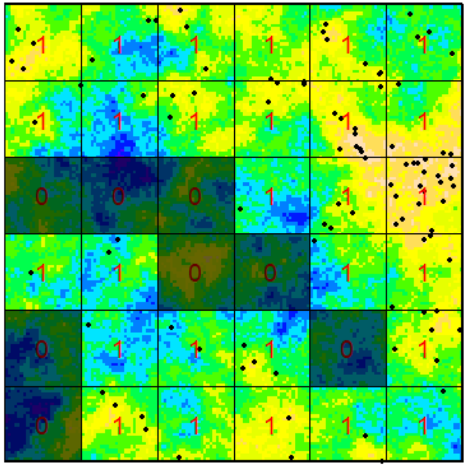
Mean intensity (lambda) = 0.00444

Abundance, N



Mean(N) = 2.78, var(N) = 12.23

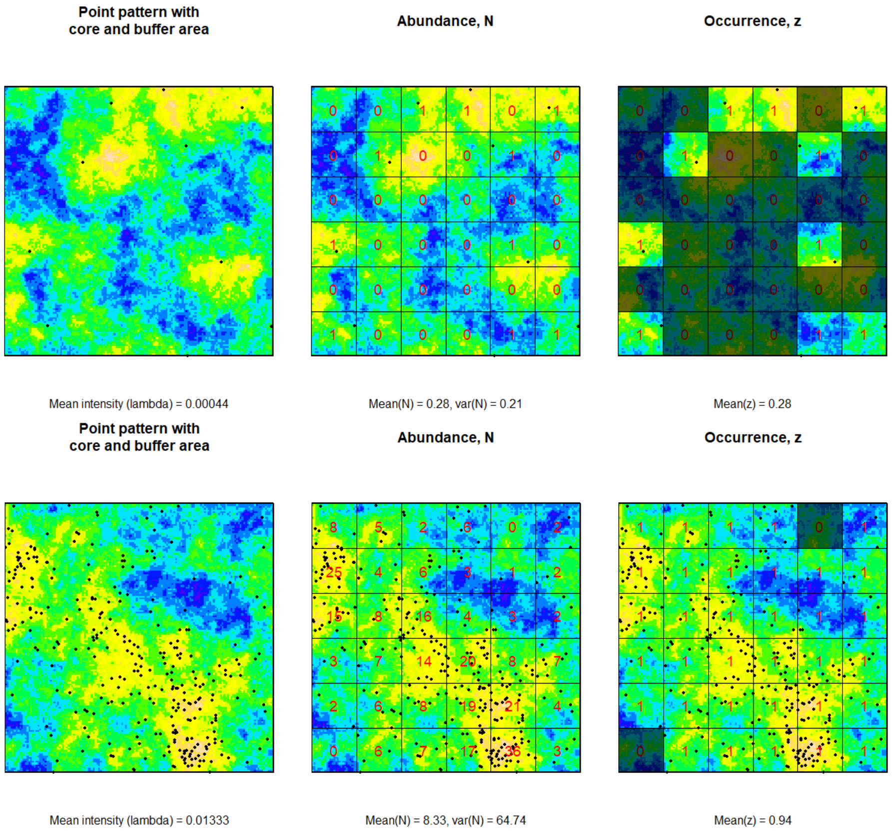
Occurrence, z



Mean(z) = 0.78

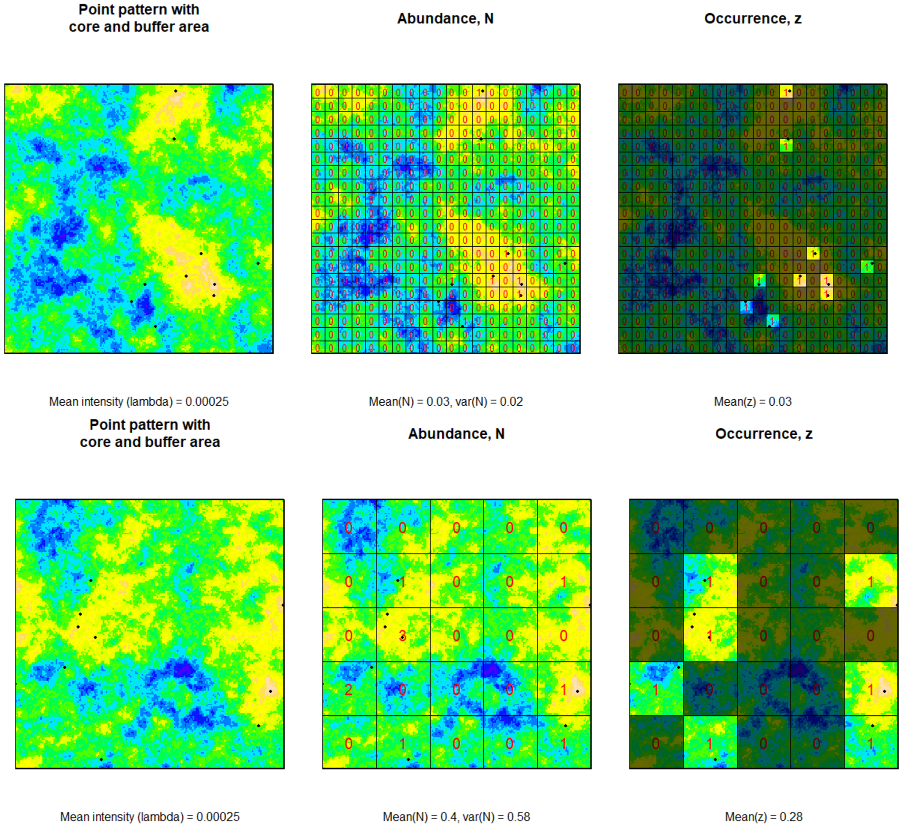
Dist&Abu depend on PP intensity and spatial scale

- Dependence on intensity



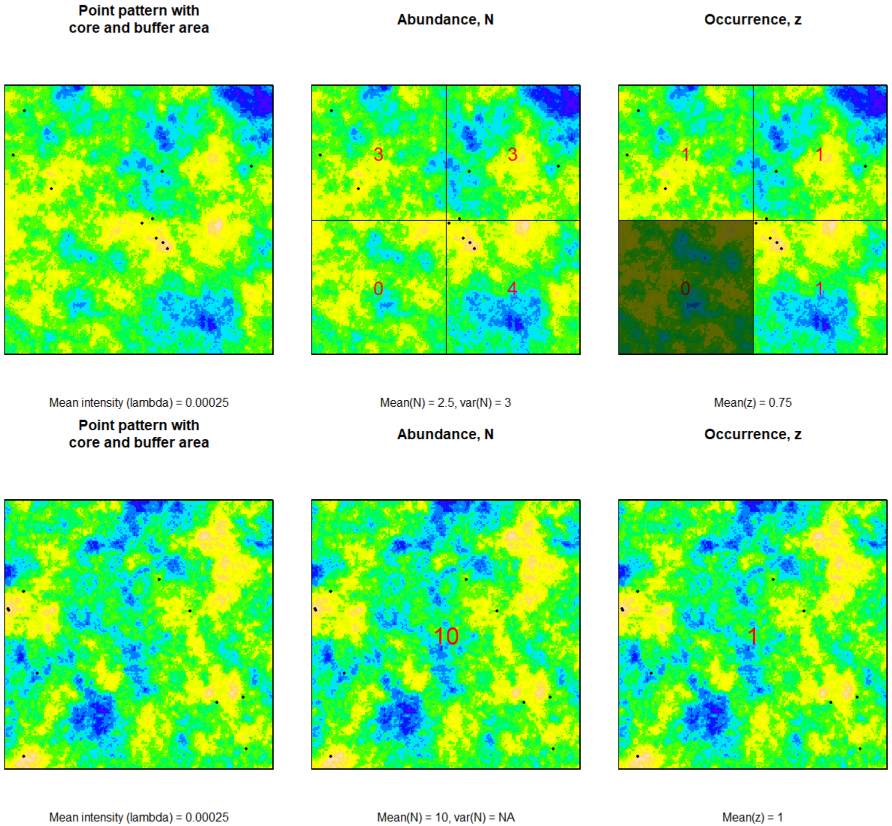
Dist&Abu depend on PP intensity and spatial scale

- Dependence on spatial scale



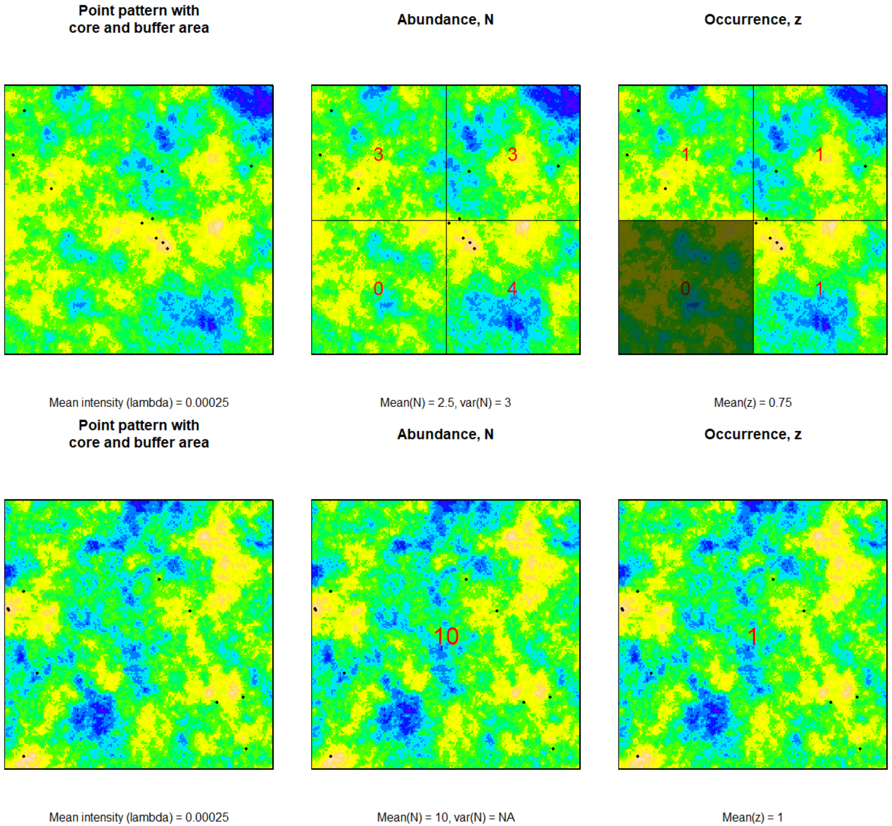
Dist&Abu depend on PP intensity and spatial scale

- Dependence on spatial scale



Dist&Abu depend on PP intensity and spatial scale

- Dependence on spatial scale



Therefore,

- If you need an occupancy prob. of ~0, just make your sites the size of a stamp
- If you need an occupancy prob. of 1 for every occurring species, just make your entire study area a single site

Side comment on "Distribution"

- Here, mostly use "distribution" synonymous with presence/absence, occurrence, or occupancy probability
- But can use "distribution" more generally and indeed use all of the following as a characterization of "species distribution":
    - Point pattern (i.e., realization of a point process) or intensity of point process (i.e., its expectation)
    - Realized or expected abundance
    - Realized or expected presence/absence (latter is occupancy probability or prob. of presence)
- Thus, a species distribution model (SDM), or a map based on it, may depict any of these

More on scale dependence of Dist & Abu

- Can't directly compare occupancy and abundance when spatial scale differs
- Downscaling/upscaling, "modifiable area unit problem", "change of support", "areal disaggregation regression"
- 4 important (and undercited) papers by Keil *et al.*, 2013-2014
- Great paper by Pacifici et al., *Ecology*, 2019
- Also recent paper by Murphy et al., *Eco. Apps*, 2023

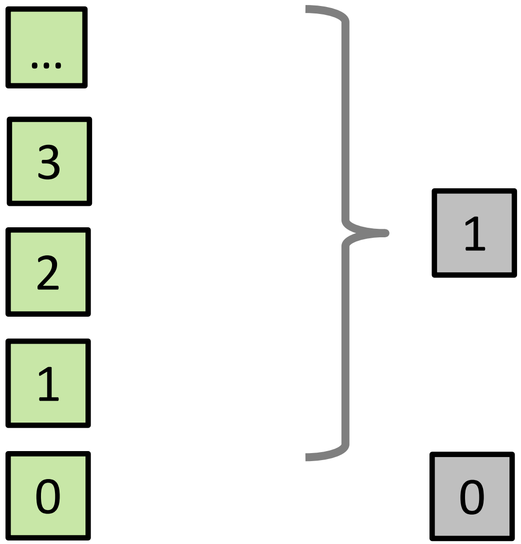
Relationship between distribution and abundance

- Not independent things: distribution is deterministic function of abundance
- Dist: Information-poor summary of abundance; *"the poor man's abundance"*

Relationship between distribution and abundance

- Not independent things: distribution is deterministic function of abundance
- Dist: Information-poor summary of abundance; *"the poor man's abundance"*

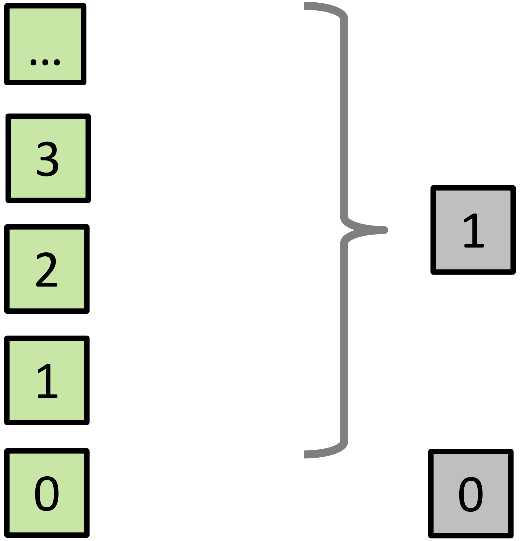
Abundance (*N*) Distribution (Presence/absence *z*)



Relationship between distribution and abundance

- Not independent things: distribution is deterministic function of abundance
- Dist: Information-poor summary of abundance; *"the poor man's abundance"*

Abundance (N) Distribution (Presence/absence z)



Thus:

p(z = 1) = 1 – p(N = 0)

Relationship between distribution and abundance

- Not independent things: distribution is deterministic function of abundance
- Dist: Information-poor summary of abundance; *"the poor man's abundance"*

Abundance (N)

…

3

2

1

0

Distribution (Presence/absence z)

1

0

Thus:

$p(z = 1) = 1 – p(N = 0)$

with Poisson abu. model:

$N \sim \text{Poisson}(\lambda)$

$P(N = 0) = e^{-\lambda}$

=> occ.prob. $\Psi = 1 - e^{-\lambda}$

Relationship between distribution and abundance

- Not independent things: distribution is deterministic function of abundance
- Dist: Information-poor summary of abundance; *"the poor man's abundance"*

<u>Abundance (N)</u> <u>Distribution (Presence/absence z)</u>

…

3

2

1

0

1

0

<u>Thus:</u>

$p(z = 1) = 1 – p(N = 0)$

with Poisson abu. model:

$N \sim \text{Poisson}(\lambda)$

$P(N = 0) = e^{-\lambda}$

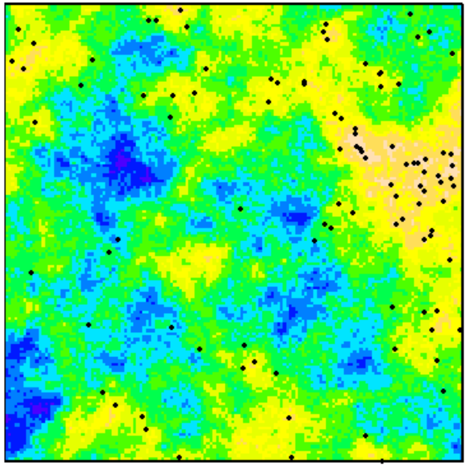
=> occ.prob. $\Psi = 1 - e^{-\lambda}$

- (Hence, typically, zero-inflated abundance model is a cheap modeling trick only)

Measurement errors in distribution and abundance

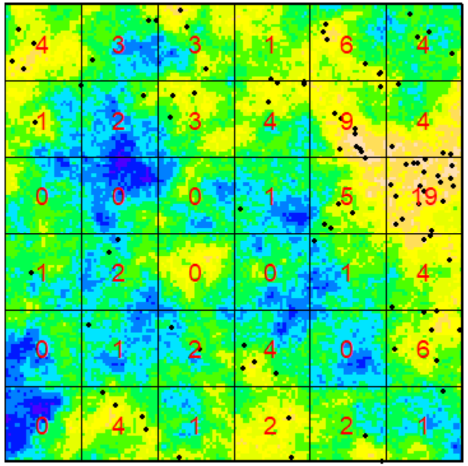
- We had 3 for the PP …. How many do we have for Dist&Abu ?

Point pattern with core and buffer area



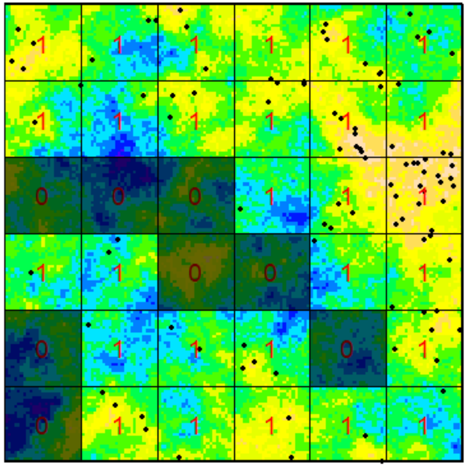
Mean intensity (lambda) = 0.00444

Abundance, N



Mean(N) = 2.78, var(N) = 12.23

Occurrence, z



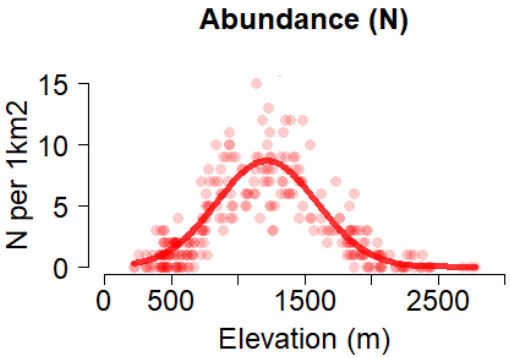
Mean(z) = 0.78

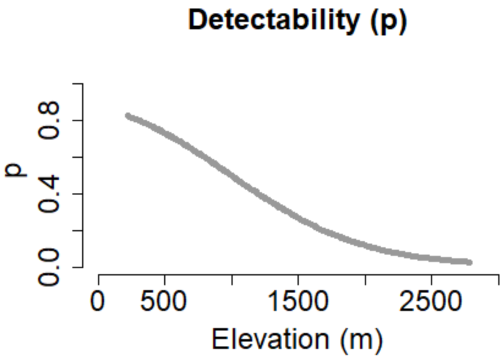
Why may it be a good idea to account for measurement errors ?

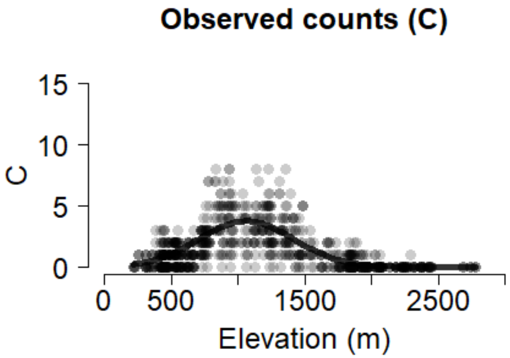
- Depends on the goal of model: e.g., do we need absolute quantities (true *N*) or are we happy with mere patterns ("relative *N*") ?
- Two examples where it would matter (simulated data)

Spatially variable detection probability

- Elevation gradient of abundance …. and of measurement error



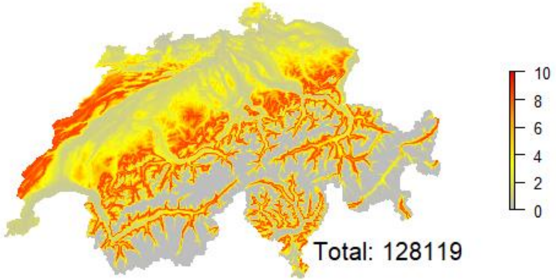




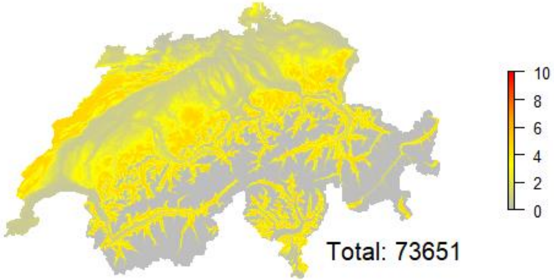
Spatially variable detection probability

- Spatial predictions and national population size

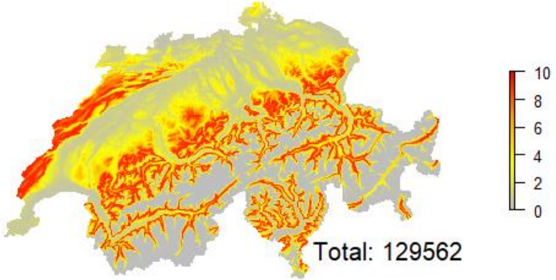
Bullfinch density (simulated truth)



Estimated bullfinch density (ignoring detection error)

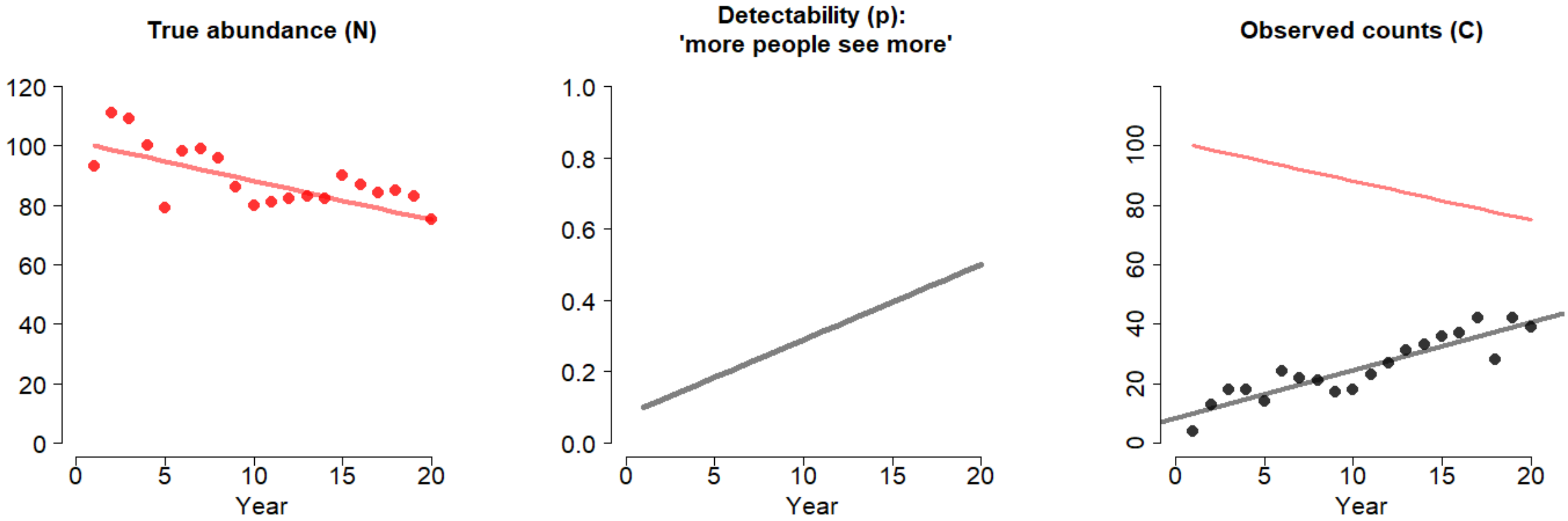


Estimated bullfinch density (correcting for detection error)



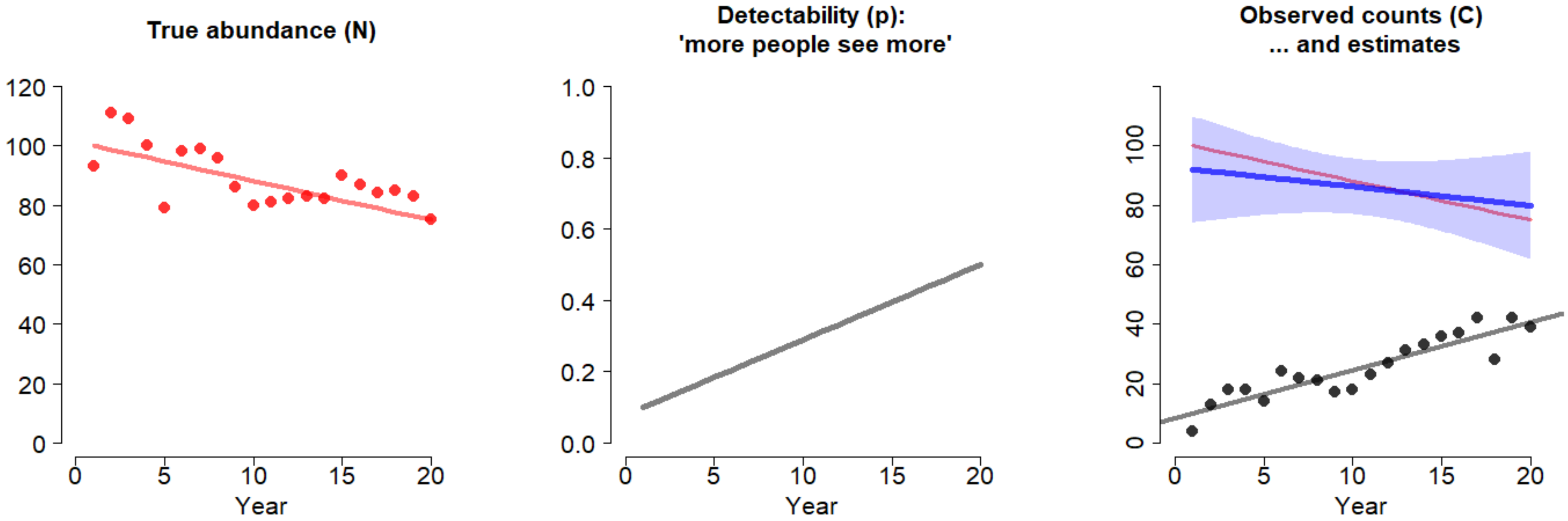
Temporally variable detection probability

- Time trends in abundance …. and of measurement error



Temporally variable detection probability

- Time trends in abundance …. and of measurement error



Hierarchical models, or HMs

- Describe sequence (or hierarchy) of random processes, or random variables (RVs)

$$x \sim f(\omega)$$
$$y \sim g(x, \theta)$$

- Factorization of joint distribution [x, y] into marginal ([x]) and conditional distribution, [y|x]: [x,y] = [x] [y|x]
- Estimands: parameters ω and θ; latent variables x
- **x**: latent variables = unobserved RVs = random effects
- Many advantages: e.g., modeling correlations, also enforcing clarity of thought
- Typically one submodel for each process underlying observed (or raw) data
- Plenty of statistical models can be formulated as HMs

Most common example of HM: mixed models

- MMs always have random effects that are continuous and drawn from normal distribution, appear additively in linear predictor of a LM or GLM

$$y_{ij} \sim Normal(\alpha_i, \sigma^2_{res})$$
$$\alpha_i \sim Normal(\mu_\alpha, \sigma^2_\alpha)$$

Most common example of HM: mixed models

- MMs always have random effects that are continuous and drawn from normal distribution, appear additively in linear predictor of a LM or GLM

$$y_{ij} \sim Normal(\alpha_i, \sigma^2_{res})$$

$$\alpha_i \sim Normal(\mu_\alpha, \sigma^2_\alpha)$$

Most common example of HM: mixed models

- MMs always have random effects that are continuous and drawn from normal distribution, appear additively in linear predictor of a LM or GLM

$$y_{ij} \sim Normal(\alpha_i, \sigma^2_{res})$$
$$\alpha_i \sim Normal(\mu_\alpha, \sigma^2_\alpha)$$

or

$$y_{ij} \sim Normal(\alpha_i + \beta_i x_i, \sigma^2_{res})$$
$$\alpha_i \sim Normal(\mu_\alpha, \sigma^2_\alpha)$$
$$\beta_i \sim Normal(\mu_\beta, \sigma^2_\beta)$$

Most common example of HM: mixed models

- MMs always have random effects that are continuous and drawn from normal distribution, appear additively in linear predictor of a LM or GLM

$$y_{ij} \sim Normal(\alpha_i, \sigma^2_{res})$$
$$\alpha_i \sim Normal(\mu_\alpha, \sigma^2_\alpha)$$

or

$$y_{ij} \sim Normal(\alpha_i + \beta_i x_i, \sigma^2_{res})$$
$$\alpha_i \sim Normal(\mu_\alpha, \sigma^2_\alpha)$$
$$\beta_i \sim Normal(\mu_\beta, \sigma^2_\beta)$$

More general hierarchical models

- NOTE: random effects can be discrete …. even binary (i.e. 0/1)
- ... and can follow distribution other than normal … e.g., Poisson or Bernoulli
- Two examples of such more general HMs, where one describes state and other observation

$$y_{ij} \sim Binomial(N_i, p_{ij})$$
$$N_i \sim Poisson(\lambda_i)$$

$$y_{ij} \sim Bernoulli(z_i \, p_{ij})$$
$$z_i \sim Bernoulli(\psi_i)$$

More general hierarchical models

- NOTE: random effects can be discrete …. even binary (i.e. 0/1)
- ... and can follow distribution other than normal … e.g., Poisson or Bernoulli
- Two examples of such more general HMs, where one describes state and other observation

$$y_{ij} \sim Binomial(N_i, p_{ij})$$
$$N_i \sim Poisson(\lambda_i)$$

$$y_{ij} \sim Bernoulli(z_i p_{ij})$$
$$z_i \sim Bernoulli(\psi_i)$$

HMs can be thought of as linked GLMs !

$$y_{ij} \sim Binomial(N_i, p_{ij})$$

$$N_i \sim Poisson(\lambda_i)$$

$$y_{ij} \sim Bernoulli(z_i p_{ij})$$

$$z_i \sim Bernoulli(\psi_i)$$

HMs can be thought of as linked GLMs !

$$y_{ij} \sim Binomial(N_i, p_{ij})$$

$$N_i \sim Poisson(\lambda_i)$$

$$y_{ij} \sim Bernoulli(z_i p_{ij})$$

$$z_i \sim Bernoulli(\psi_i)$$

- And we can add effects of covariates ....

$$\text{logit}(p_{ij}) = \alpha_0 + \alpha_1 x_{1,i} + \alpha_2 x_{2,ij}$$

$$\log(\lambda_i) = \beta_0 + \beta_1 x_i$$

$$\text{logit}(p_{ij}) = \alpha_0 + \alpha_1 x_{1,i} + \alpha_2 x_{2,ij}$$

$$\text{logit}(\psi_i) = \beta_0 + \beta_1 x_i$$

HMs can be thought of as linked GLMs !

$$y_{ij} \sim Binomial(N_i, p_{ij})$$

$$N_i \sim Poisson(\lambda_i)$$

$$y_{ij} \sim Bernoulli(z_i p_{ij})$$

$$z_i \sim Bernoulli(\psi_i)$$

- … and spatial effects !

$$\text{logit}(p_{ij}) = \alpha_0 + \alpha_1 x_{1,i} + \alpha_2 x_{2,ij}$$

$$\log(\lambda_i) = \beta_0 + \beta_1 x_i + w_i$$

$$w_i \sim MVN(0, \Sigma)$$

$$\text{logit}(p_{ij}) = \alpha_0 + \alpha_1 x_{1,i} + \alpha_2 x_{2,ij}$$

$$\text{logit}(\psi_i) = \beta_0 + \beta_1 x_i + w_i$$

$$w_i \sim MVN(0, \Sigma)$$

Summary: What are Distribution and abundance ?

- Dist is z/ψ and abundance is N/λ
- Dist&Abu are summaries of a point pattern -> get a feel for PPs with `simPPe()`
- [PP: the fundamental demographic model of animals and plants]
- Dist&Abu depend on underlying intensity of PP … and on scale of spatial discretization!
- … they are scale dependent
- Dist is deterministic function of abundance
- Dist&Abu are the same when intensity and/or scale small
- Otherwise, Dist is information-poor summary of abundance
- Beyond some density, Dist becomes uninformative about abundance
  (when all quads occupied)

Summary: Measurement errors

- 3 for point patterns, only 2 for Dist&Abu
- Abundance measurements: counts
- Distribution (or presence/absence) Measurements: detection/nondetection data
- Can screw things when unaccounted for
- Accommodation requires repeated measurements

Summary: Hierarchical models for Dist&Abu

- HM: statistical models that area linked sequence (or hierarchy) of submodels, typically one for each process that underlie observed data
- Typical processes in Dist/Abu : state process (e.g., abundance or pres./abs.) & observation process (e.g., false negatives, false positives)
- Thus, with such HMs we can model Dist/Abu as random effects, while correcting for ubiquitous measurement errors in all field data on plant and animal distribution and abundance


Thank you.
François Mougeot