Modeling of species distribution change and range dynamics

Marc Kéry
Swiss Ornithological Institute

Sempach, 14 September 2012

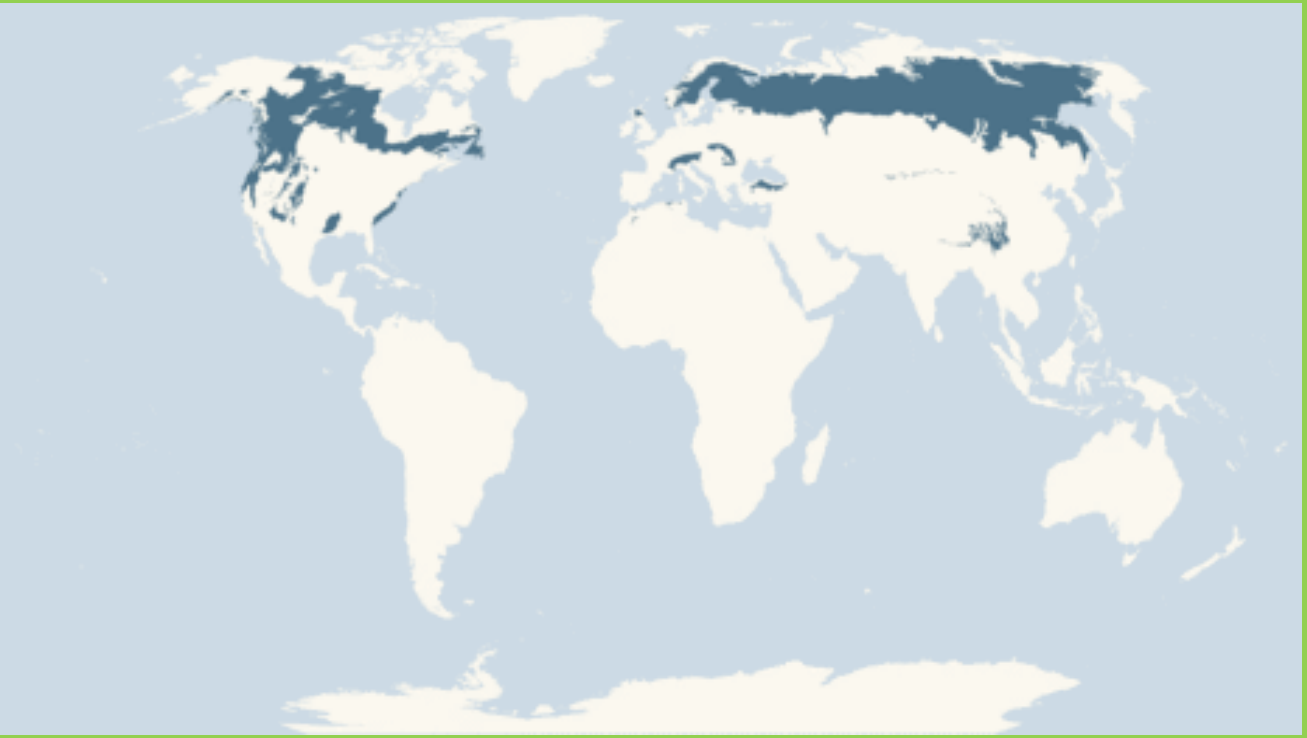
Outline

- Intro: What is a species distribution ?
- Dynamic occupancy models (MacKenzie et al., *Ecology*, 2003)
- Application: Crossbills in Switzerland 1999–2007
  (Kéry, Guillera-Arroita & Lahoz-Monfort, *Journal of Biogeography*, in review)

What is a species distribution ?

What is a species distribution ?

- Google “common crossbill distribution”
- We immediately understand what this means ...



http://www.bbc.co.uk/nature/life/Common_Crossbill

What is a species distribution ?

- Or do we really ??? What **exactly** is meant by that picture ?

What is a species distribution ?

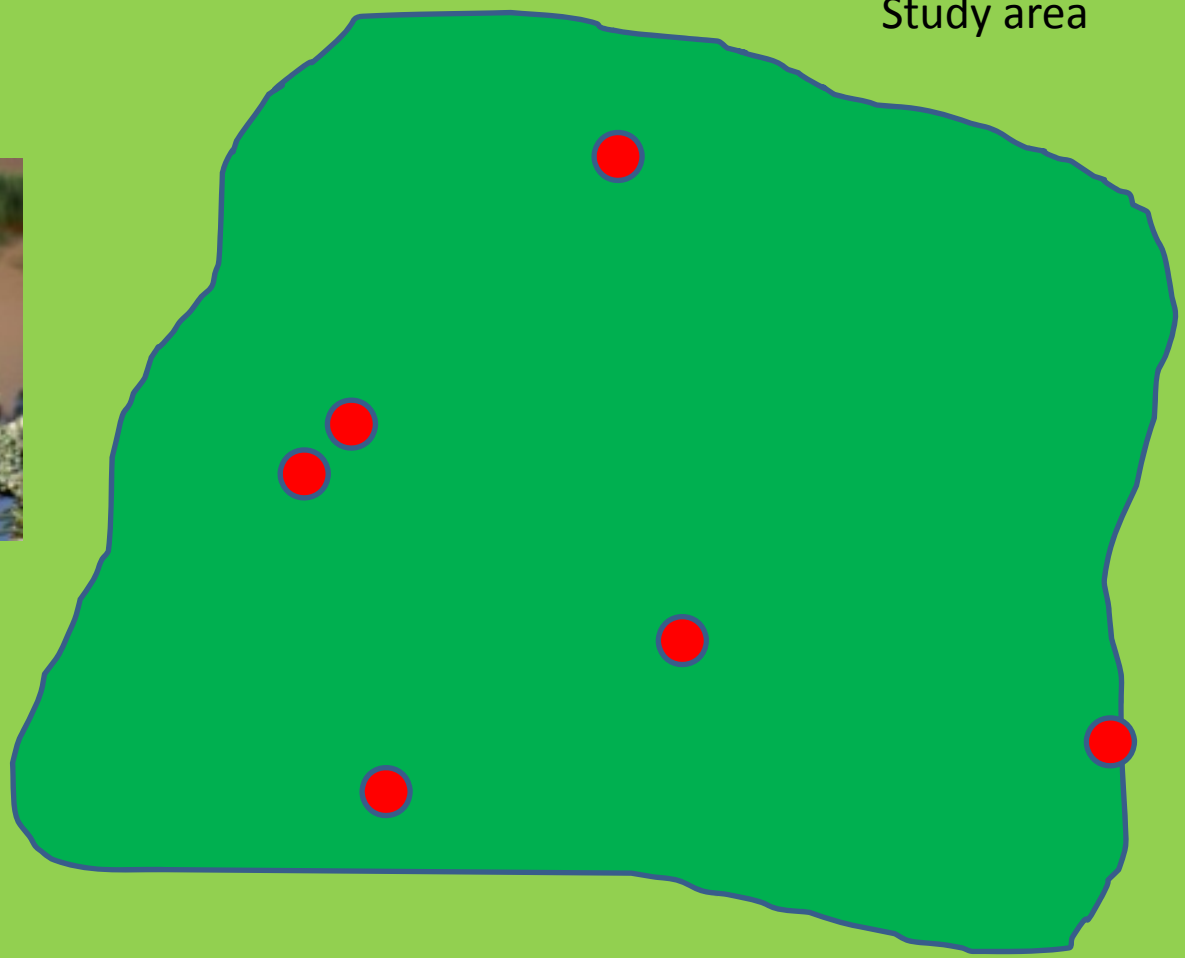
- Or do we really ??? What **exactly** is meant by that picture ?

Male crossbill

What is a species distribution ?

- What **exactly** is meant by such distribution maps ?

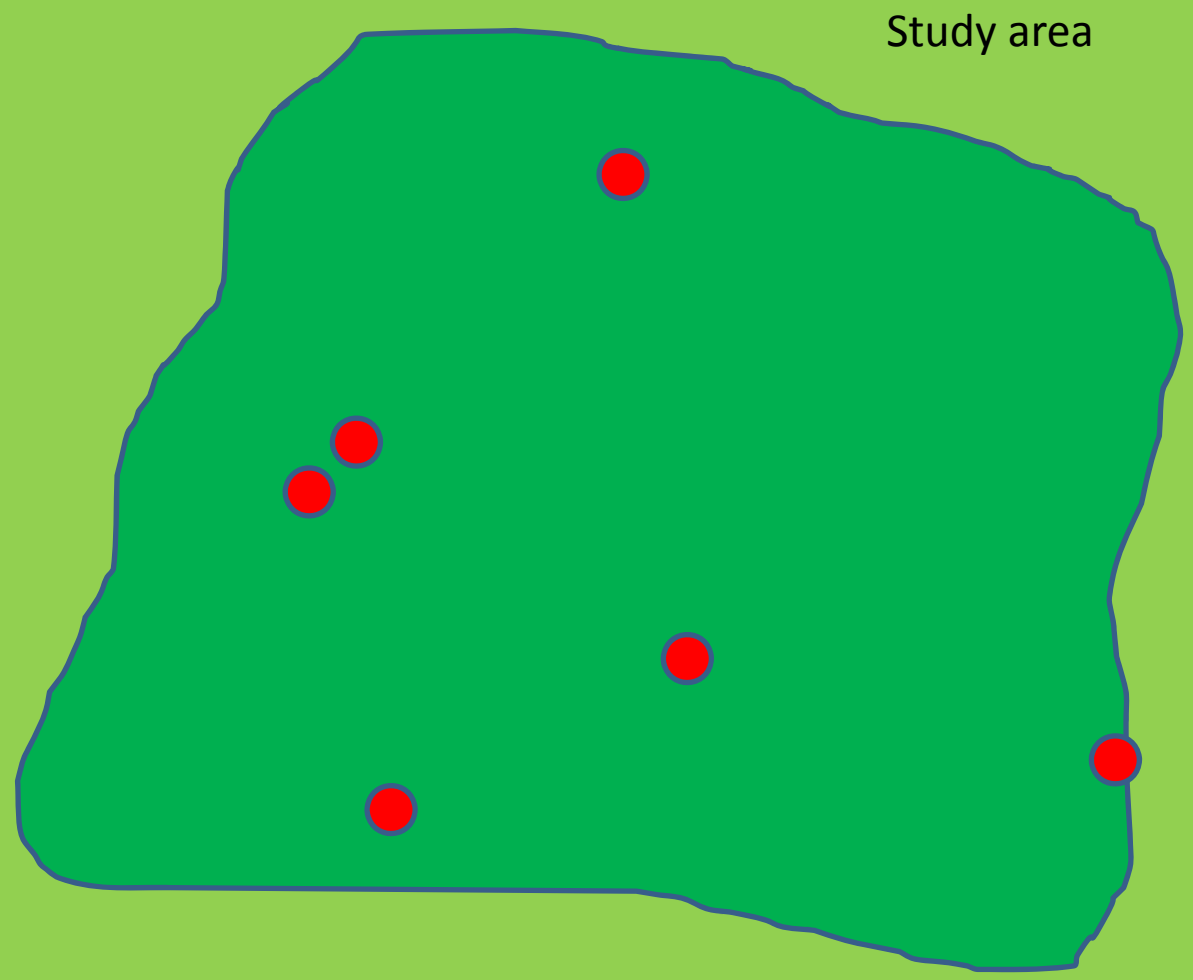
Male crossbill



What is a species distribution ?

- **Distribution:** “area where a species occurs”

$$\frac{Area\ with\ crossbill}{Total\ area}$$

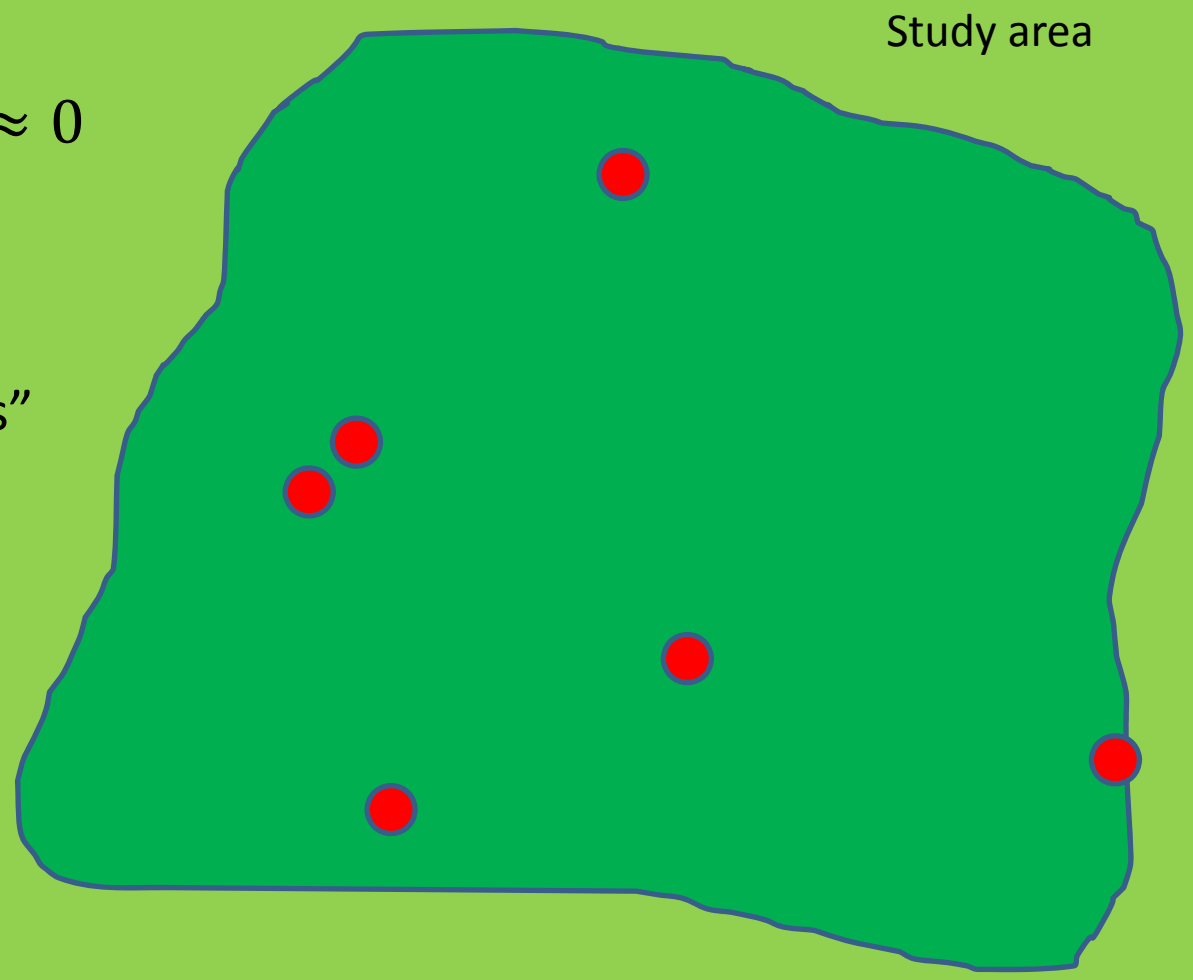


What is a species distribution ?

- **Distribution:** “area where a species occurs”

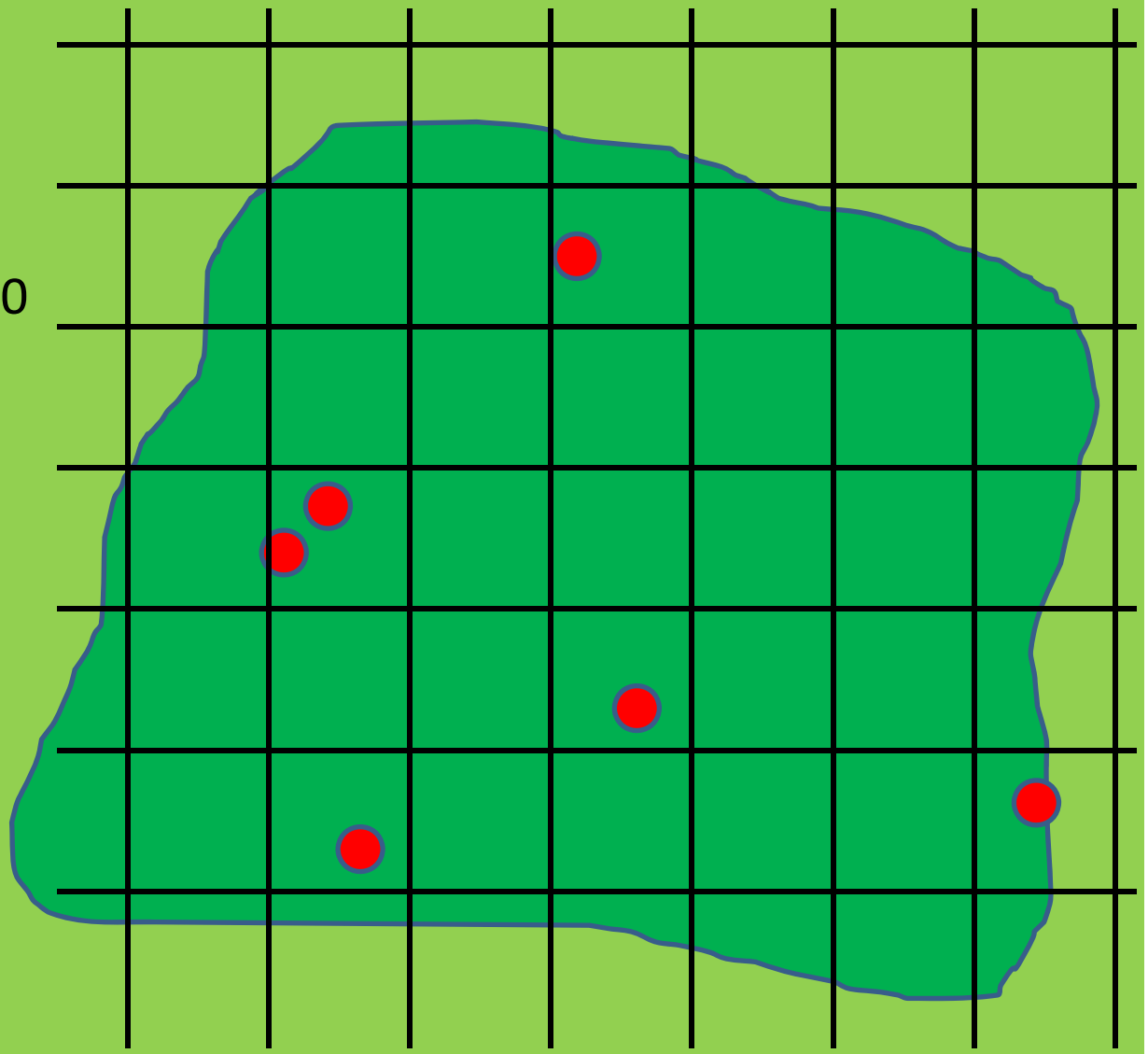
$$\frac{Area\ with\ crossbill}{Total\ area} \approx 0$$

- Birds small & discrete
- Approximately “points”
- Realization of spatial point process



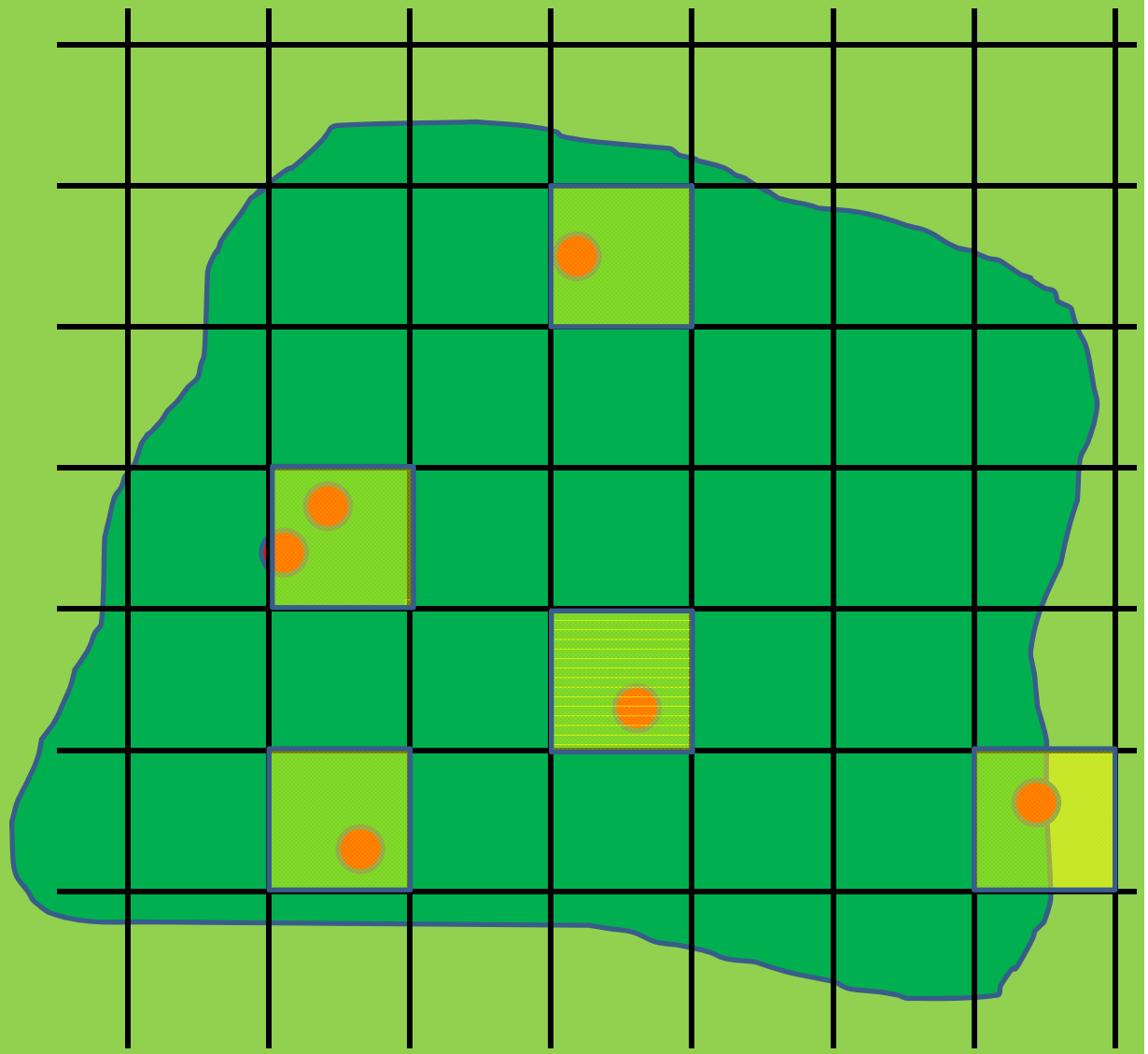
What is a species distribution ?

- In practice, have to impose “grid” to summarize spatial point process
- Otherwise bird area = 0 (and distribution = 0) !



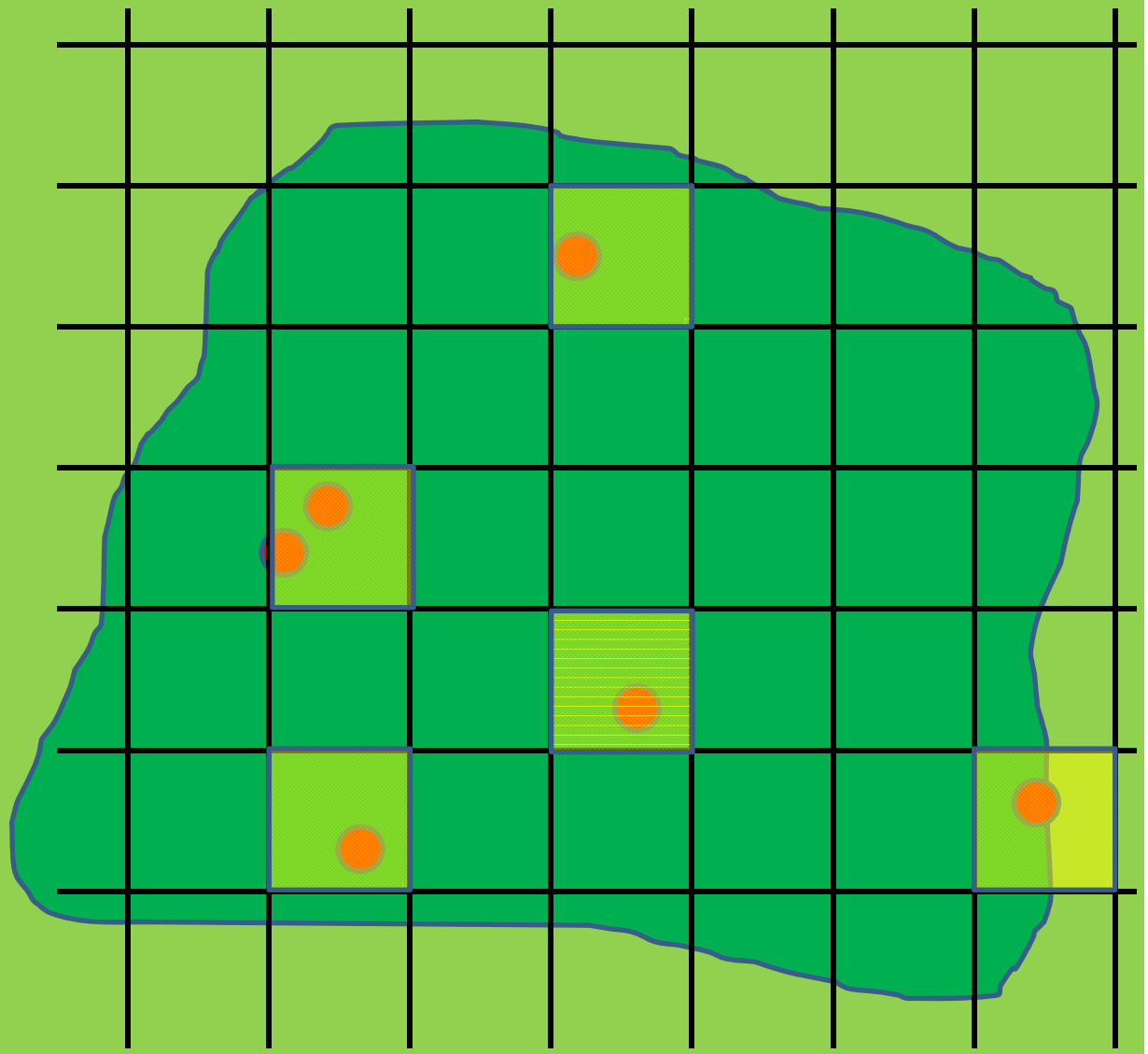
What is a species distribution ?

- Occurrence (z):
  areal summary
  of pixel abundance N
- Distribution:
  pattern of z,
  sum of occurrences
  (5 pixels)



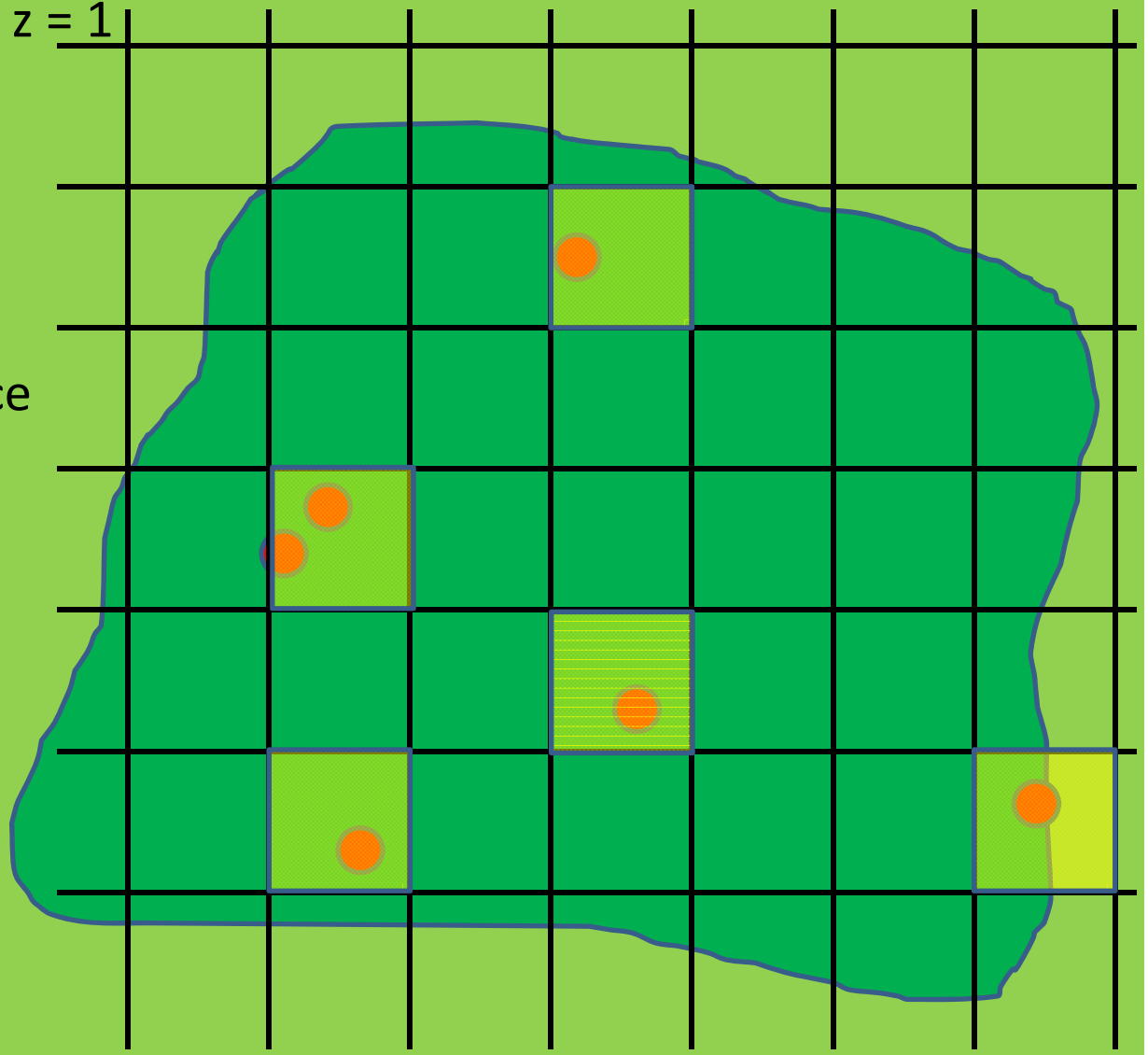
Mathematical relationship between distribution and abundance

- Occurrence/Presence: $z = 1$
- Presence $z = 1$ if $N > 0$
- Absence $z = 0$ if $N = 0$



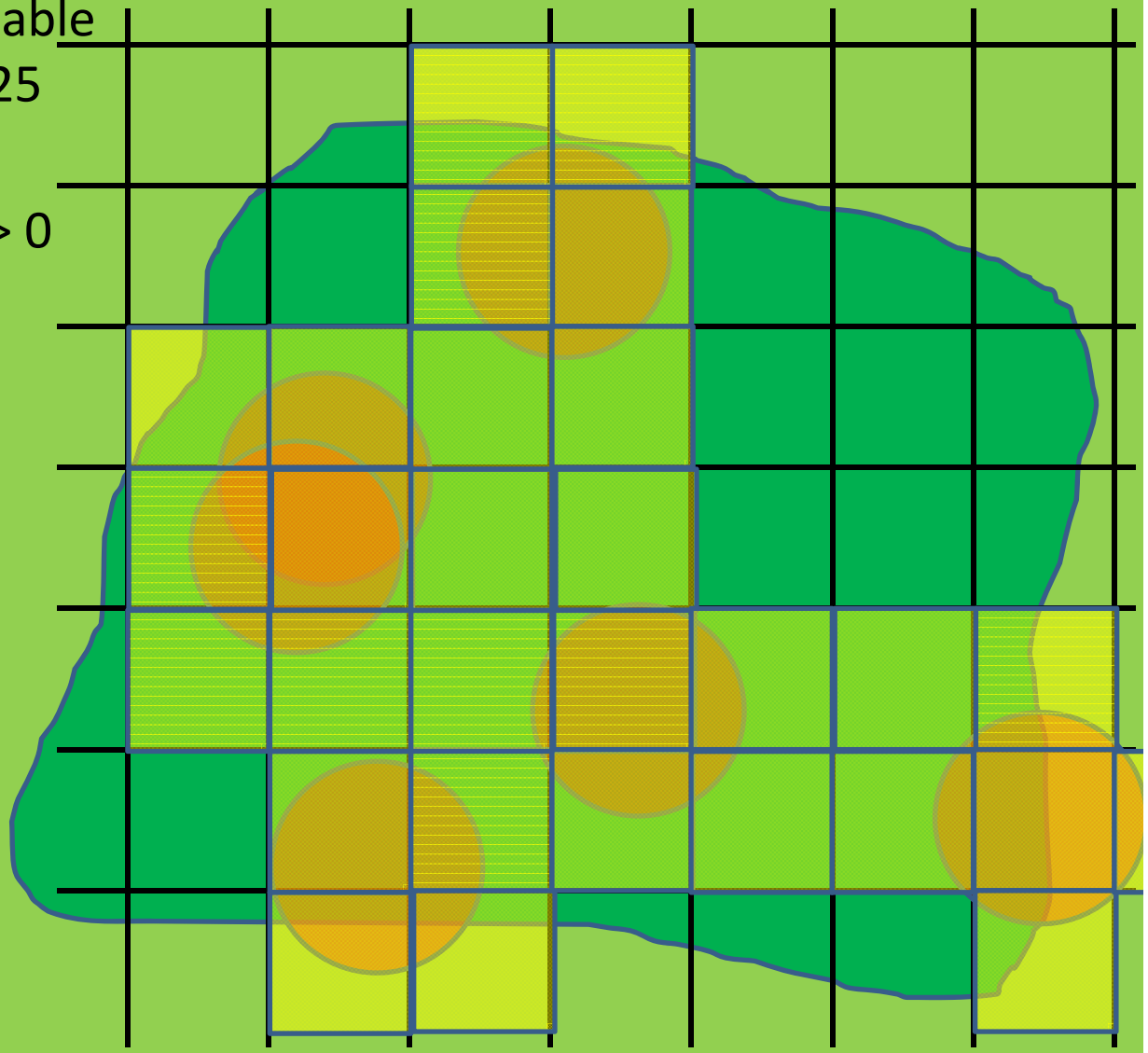
Mathematical relationship between distribution and abundance

- Occurrence/Presence: z = 1
- Presence z = 1 if N > 0
- Absence z = 0 if N = 0
- Distribution = information-reduced summary of abundance for some areal unit
- **AND:**
  Both distribution & abundance areal-summary of spatial point process



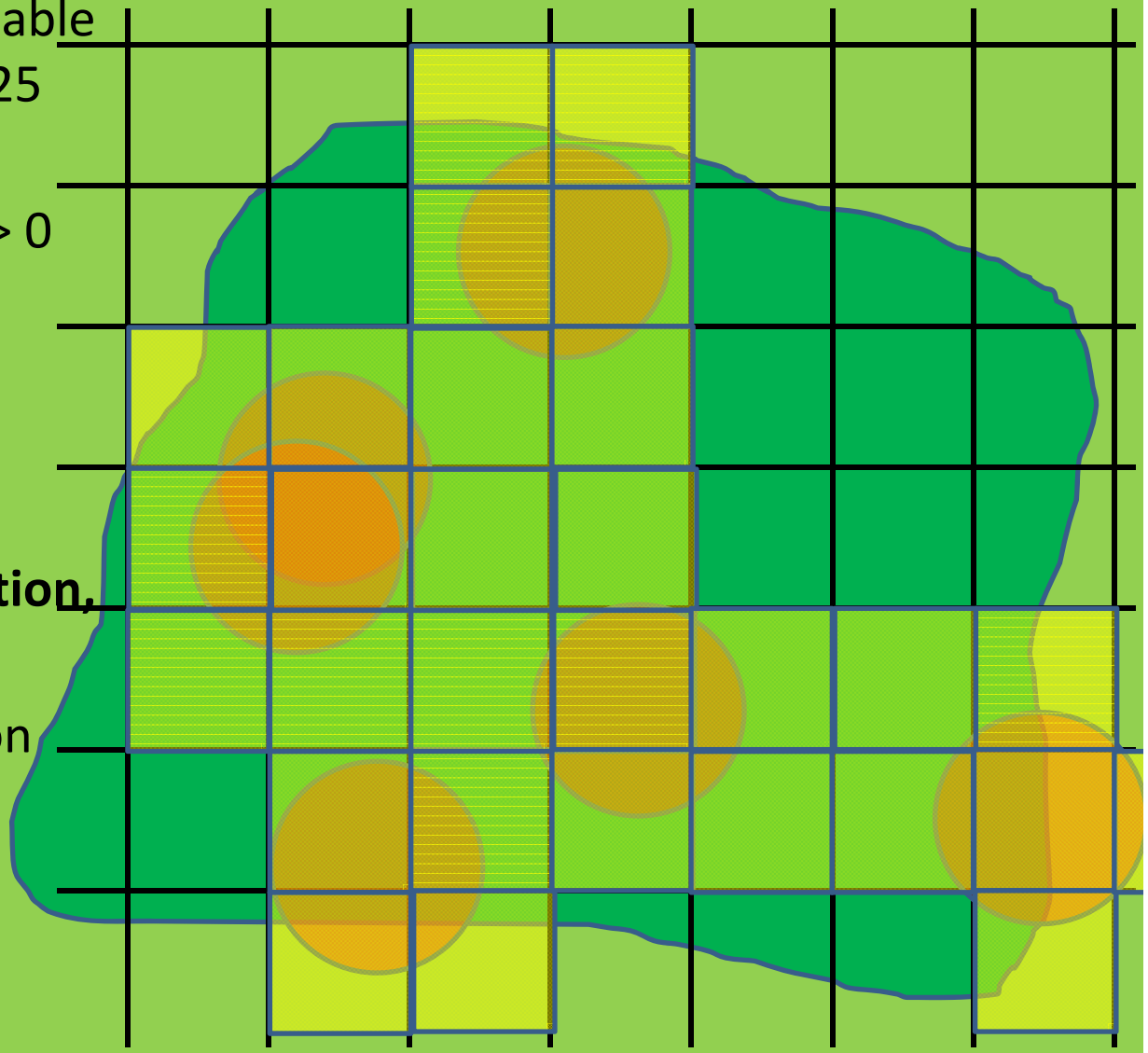
And what if there is movement ?

- **New:** N is random variable E(N) fractional, e.g. 0.25
- Same: Presence z = 1 if E(N) > 0
- Same: Distribution = information-reduced summary of expected abundance



And what if there is movement ?

- **New:** N is random variable E(N) fractional, e.g. 0.25
- Same: Presence z = 1 if E(N) > 0
- Same: Distribution = information-reduced summary of expected abundance
- **New: greater distribution, "use" vs. "presence"**
- Same: Both distribution & abundance areal-summary of spatial point process



Section summary: What is a species distribution ?

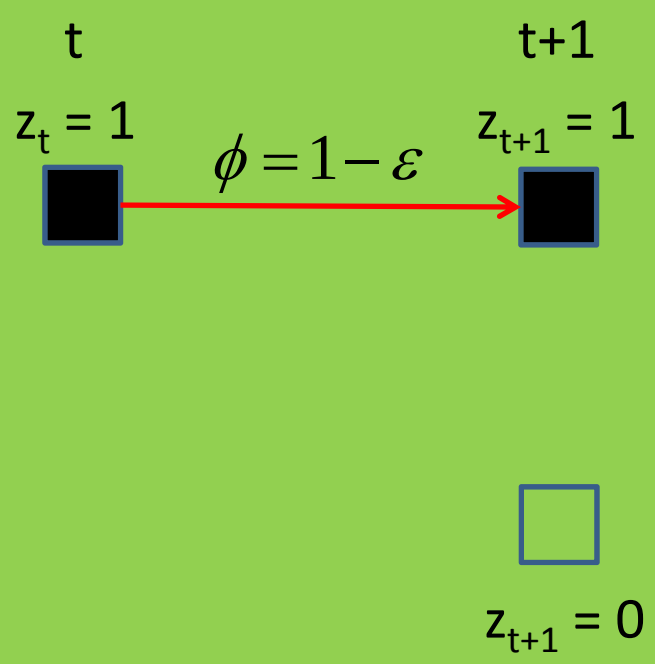
- Fundamentally, abundance and species distribution both areal summary of (a realization of) spatial point process
- Point process: individual animals and plants small & discrete units (not continuous like sea surface temperature)
- Any continuous species distribution map **implies** raster, within which point pattern is summarized
- Concept of "distribution" meaningless without:
  (i) definition of occurrence
  (ii) definition of raster (pixel size)
  (iii) spatial scale of organism movement (if any)
  [(iv) temporal scale]
- Distribution is information-poor summary of abundance

A model for distribution dynamics

- Start: define grid/pixel
  define state of occurrence
- Markov model: occurrence state at t+1 depends only on state at t
- Parameters for dynamic rates: occupied -> unoccupied
  unoccupied -> occupied

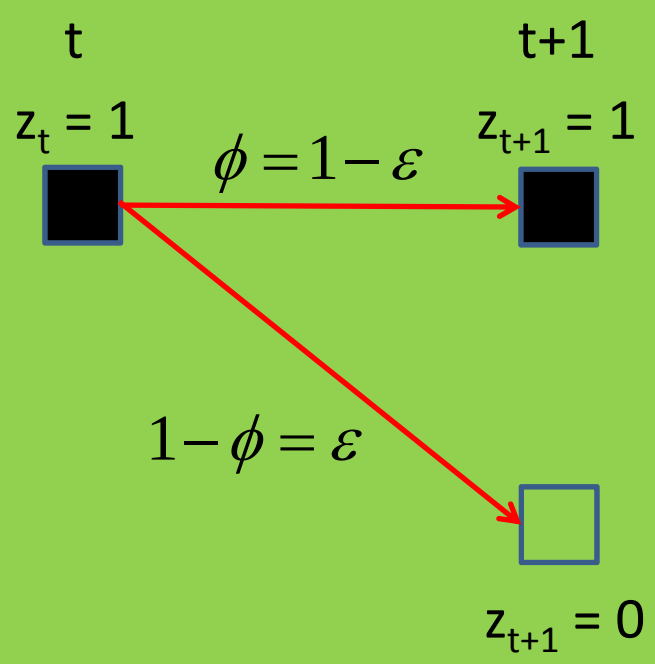
A model for distribution dynamics

- Start: define grid/pixel
  define state of occurrence
- Markov model: occurrence state at t+1 depends only on state at t
- Parameters for dynamic rates: occupied -> unoccupied
  unoccupied -> occupied



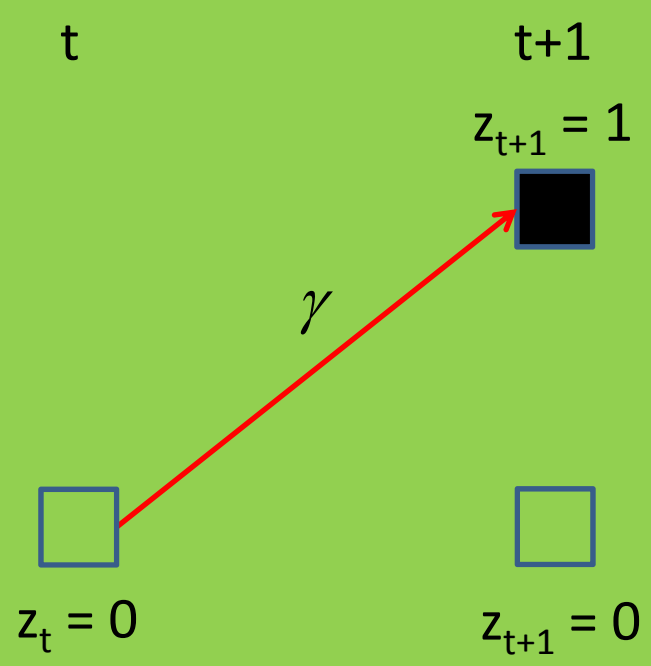
A model for distribution dynamics

- Start: define grid/pixel
  define state of occurrence
- Markov model: occurrence state at t+1 depends only on state at t
- Parameters for dynamic rates: occupied -> unoccupied
  unoccupied -> occupied



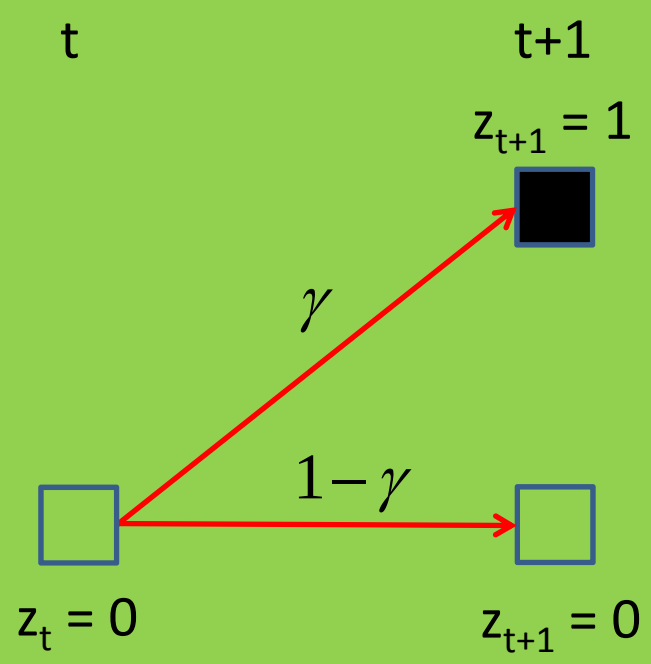
A model for distribution dynamics

- Start: define grid/pixel
  define state of occurrence
- Markov model: occurrence state at t+1 depends only on state at t
- Parameters for dynamic rates: occupied -> unoccupied
  unoccupied -> occupied

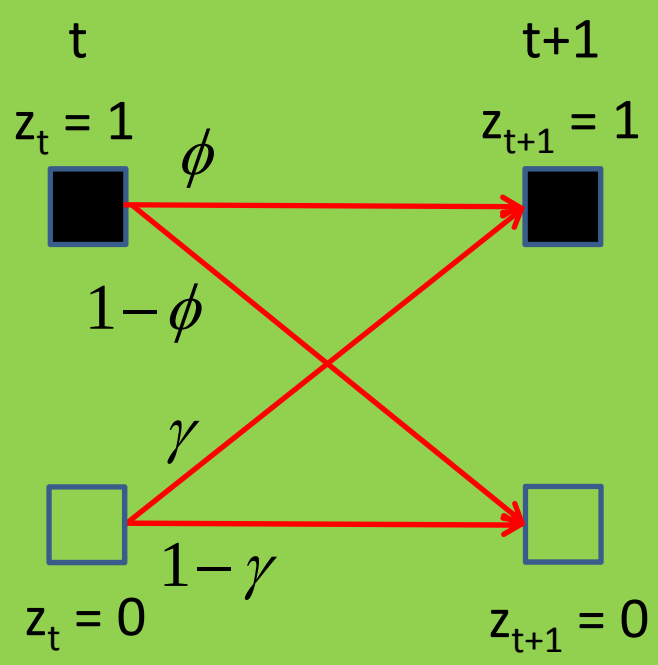


A model for distribution dynamics

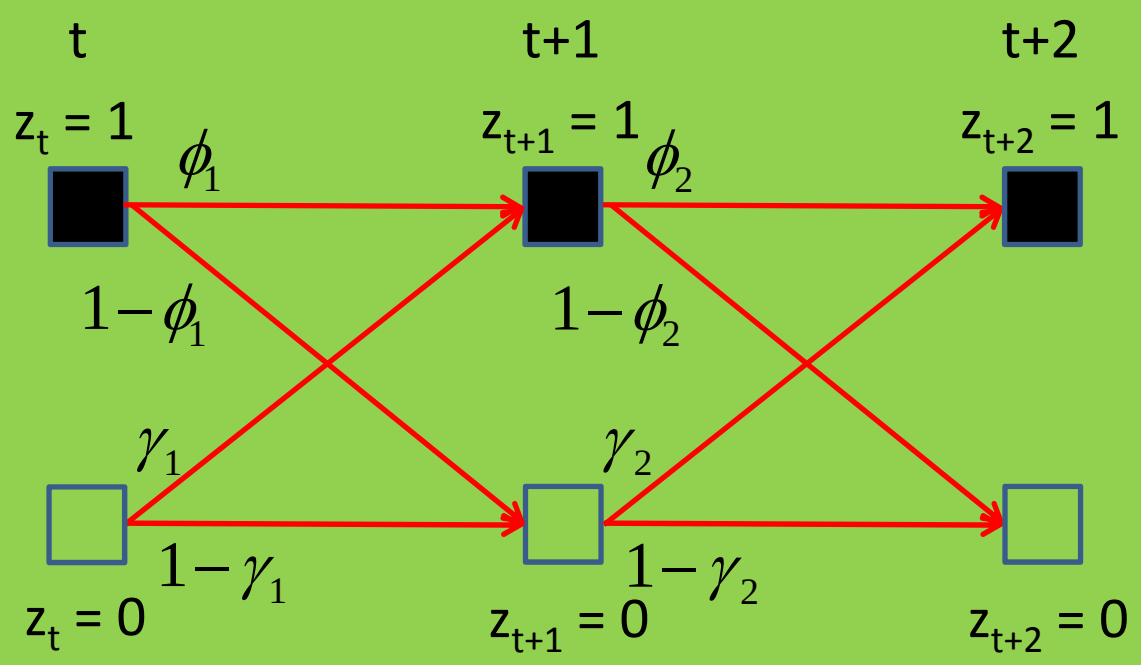
- Start: define grid/pixel
  define state of occurrence
- Markov model: occurrence state at t+1 depends only on state at t
- Parameters for dynamic rates: occupied -> unoccupied
  unoccupied -> occupied



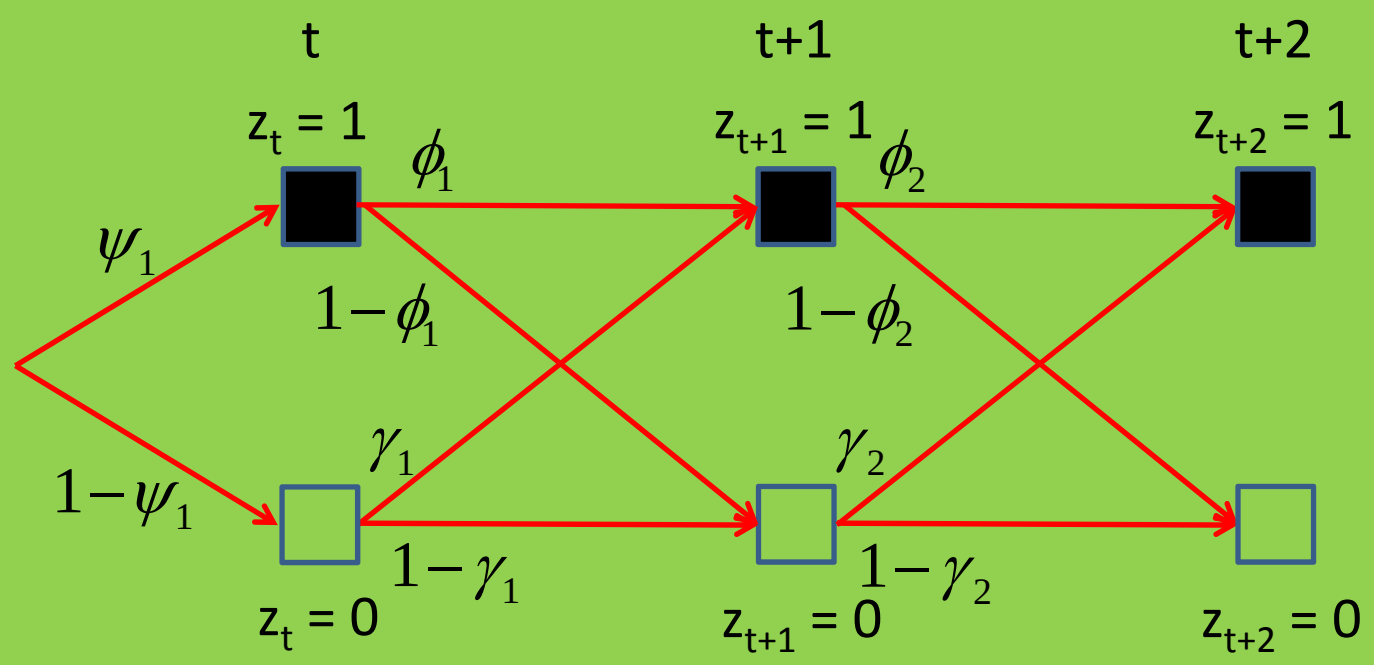
Markov model for distribution dynamics



Markov model for distribution dynamics



Markov model for distribution dynamics



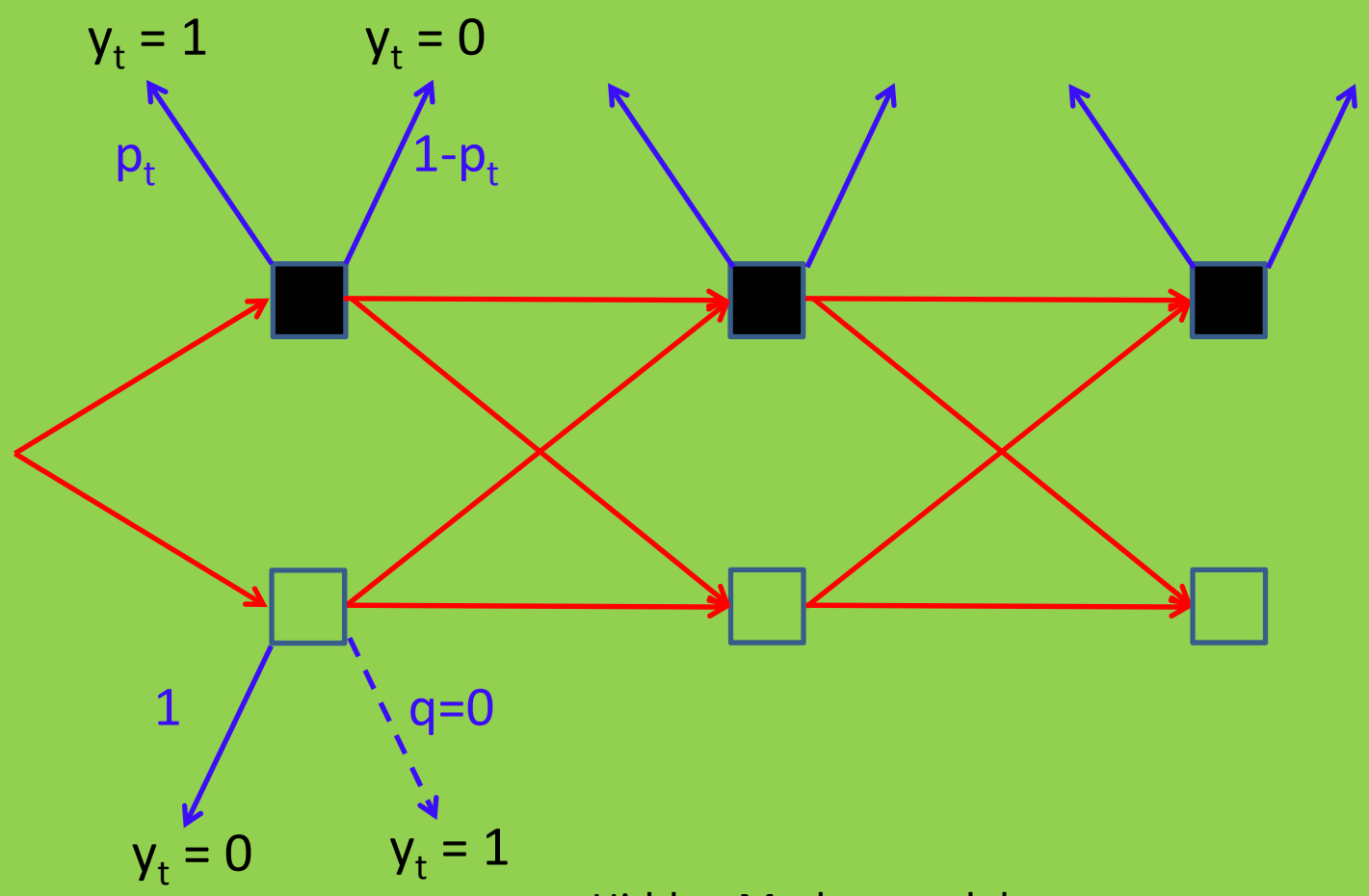
- Markovian dynamics model:
  (i) initial state z_t governed by initial occupancy probability ψ
  (ii) dynamics governed by survival (ϕ) and colonization (γ)
  (iii) states z and occupancy probability after t obtained recursively

Markov model for distribution dynamics

- **In reality, occurrence state seldom perfectly observable !**

Hidden Markov model for distribution dynamics

- **In reality, occurrence state seldom perfectly observable !**



Hidden Markov model:
(i) p: detection probability (false-negative errors)
(ii) q: false-positive error = 0 (by assumption)

Section summary: Dynamic occupancy model (MacKenzie et al. 2003)

- Hidden Markov model, 4 basic sets of parameters
- 4 logistic regressions:
  - z_1 ~ Bernoulli(ψ_1) # Inititial state
  - z_{t+1} ~ Bernoulli(ϕ) # Survival rate, if z_t = 1 (occupied)
  - z_{t+1} ~ Bernoulli(γ) # Colonisation rate, if z_t = 0 (unoccupied)
  - y_t ~ Bernoulli(p_t) # Detection error (for occupied site)
- Can add GLMs for covariates: logit(...) = a + b * X

- Parameters estimable if 'robust-design' data, e.g.

00	1	01000	NA	(more years ...)
NA	0010	NA	NA	
101	NA	000	00	

(more sites ...)

Section summary: Dynamic occupancy model (MacKenzie et al. 2003)

- Hidden Markov model, 4 basic sets of parameters
- 4 logistic regressions:
  - z_1 ~ Bernoulli(ψ_1) # Inititial state
  - z_{t+1} ~ Bernoulli(ϕ) # Survival rate, if z_t = 1 (occupied)
  - z_{t+1} ~ Bernoulli(γ) # Colonisation rate, if z_t = 0 (unoccupied)
  - y_t ~ Bernoulli(p_t) # Detection error (for occupied site)
- Can add GLMs for covariates: logit(...) = a + b * X

- Parameters estimable if 'robust-design' data, e.g.

00	1	01000	NA	(more years ...)
NA	0010	NA	NA	
101	NA	000	010	**Must assume 'closure'**

(more sites ...)

Application: European crossbills in Switzerland 1999-2007

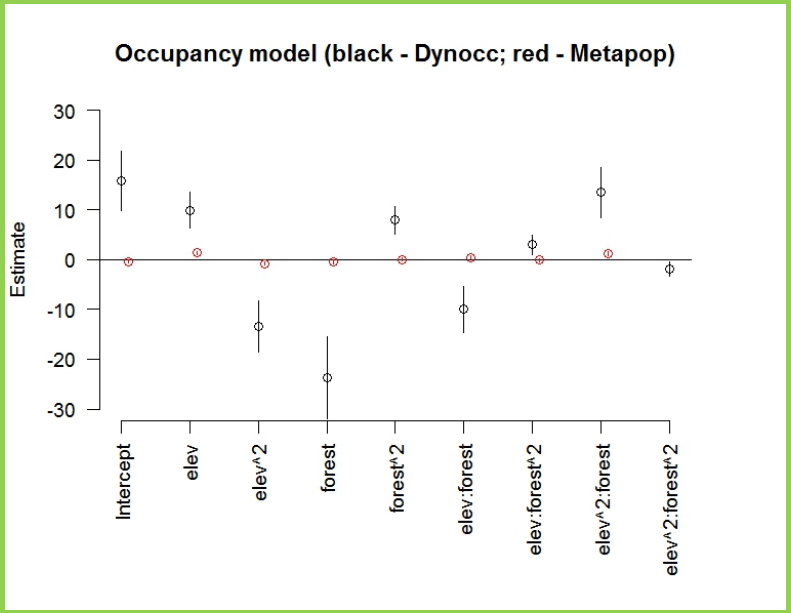
- Kéry, Guillera-Arroita & Lahoz-Monfort, *Journal of Biogeography*
- Birds eat spruce seeds; spruce cone production variable from year to year; bird distribution dynamic
- Data: Swiss Common Bird Monitoring MHB (Schmid et al. 2004); 267 1 km^2 quadrats with 2-3 surveys per breeding season
- Model fit using Maximum likelihood; R package **unmarked** (Fiske & Chandler, *JSS*, 2011)
- also possible MARK, PRESENCE, WinBUGS/OpenBUGS/JAGS, ...
- Find correlates of everything: year, forest cover, elevation, date of survey
- Model selection using Likelihood ratio tests
- Produce maps
- Compare with analysis that assumes perfect detection (p = 1) (metapopulation model)

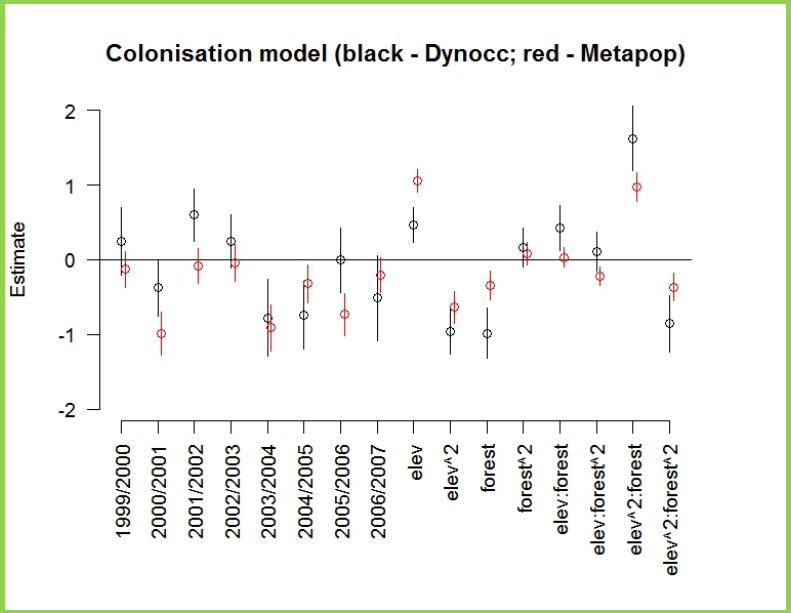
Crossbills: Model selection and fit

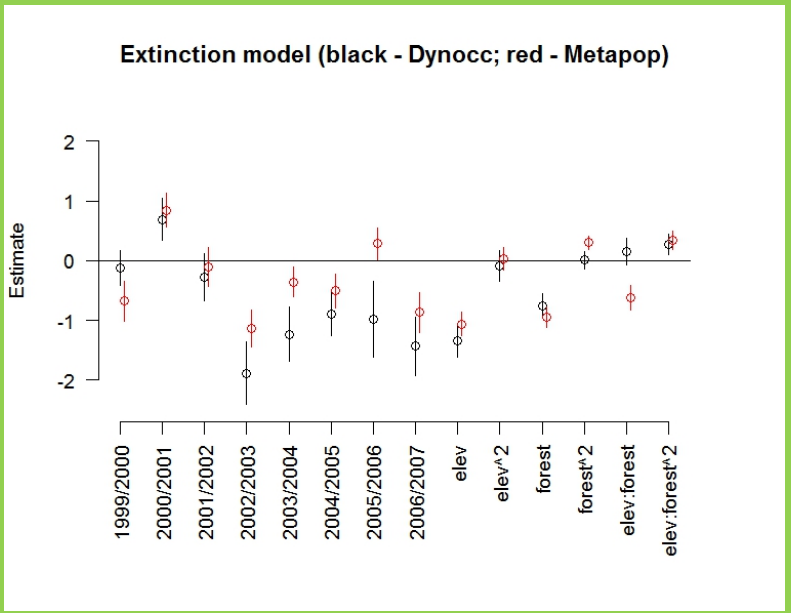
- Best dynamic occupancy model (with estimation of detection):
  - (1) Time-dependent params
  - (2) Detection: Year + (elev + elev^2)*(forest + forest^2)
  - (3) 1st-year Occupancy: (elev + elev^2)*(forest + forest^2)
  - (4) Extinction: Year + (elev + elev^2)*(forest + forest^2)
  - (5) Colonisation: Year + (elev + elev^2)*(forest + forest^2) – e2.f – e2.f2

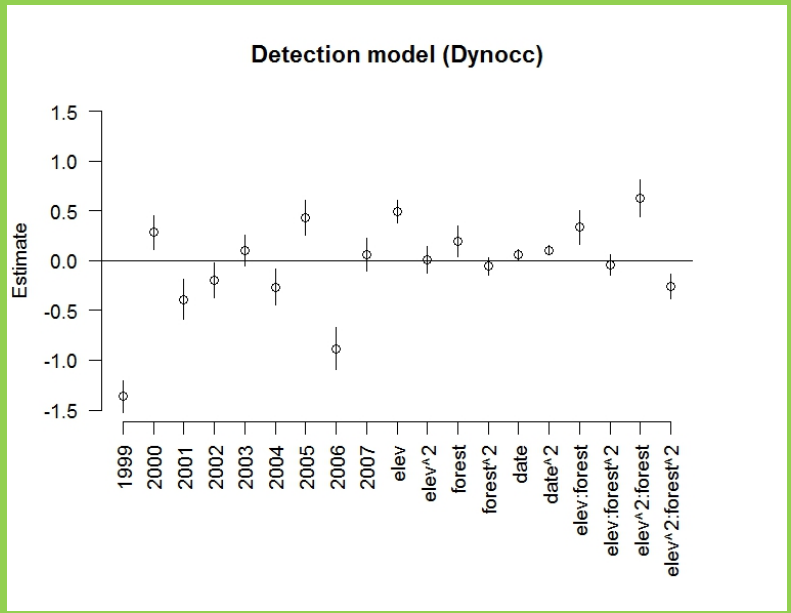
- Model fit: Parametric bootstrap, X^2 test statistic: p = 0.72
- Best naive metapopulation model (without estimation of detection): almost same params chosen

Crossbills: Parameter estimates dynocc + metapop models

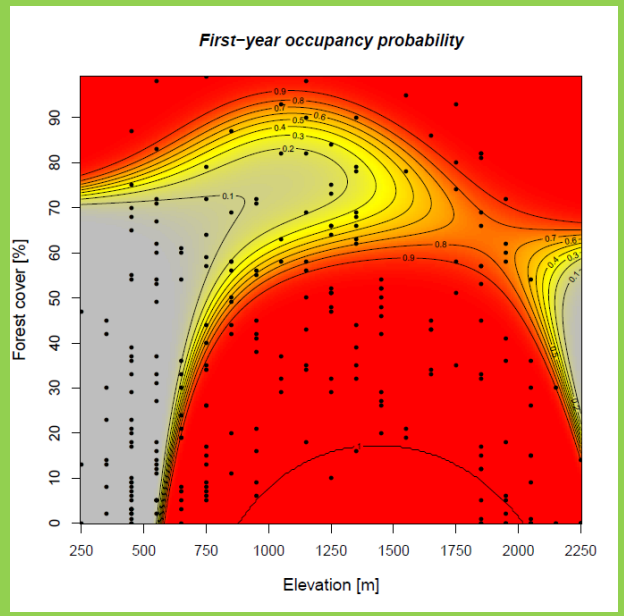


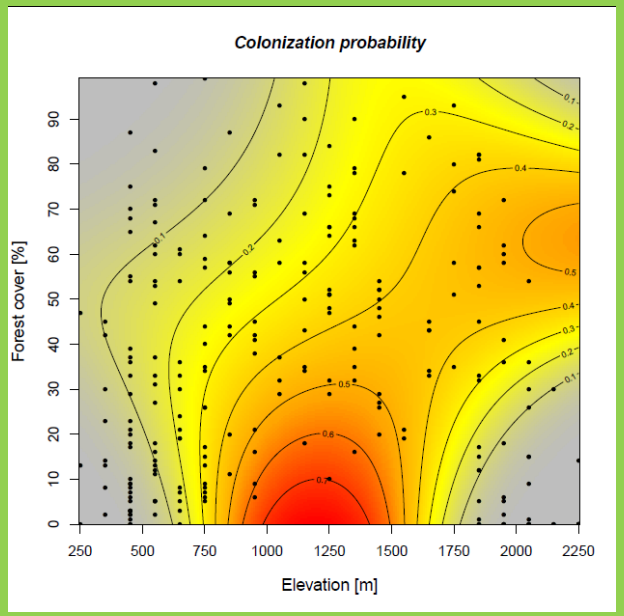


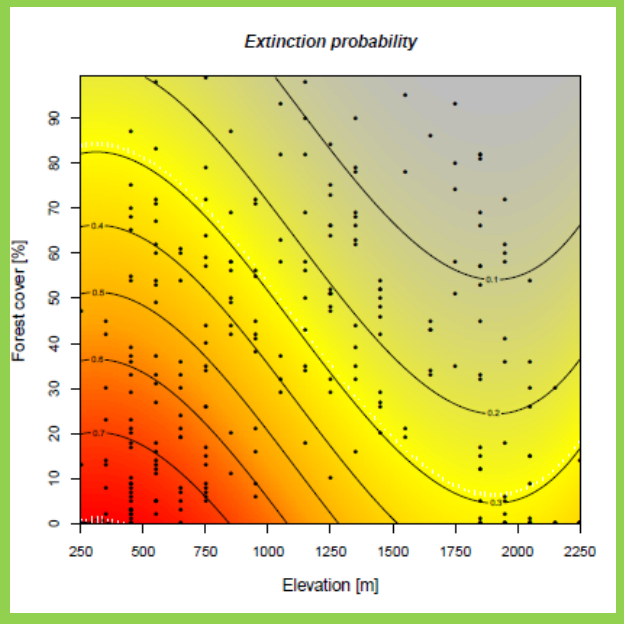


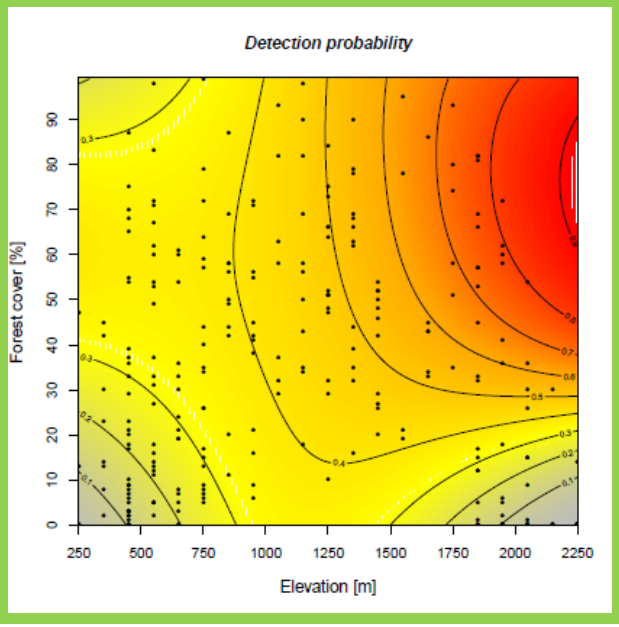


Crossbills: What do these parameter estimates mean ?
Predictive maps in environmental space.



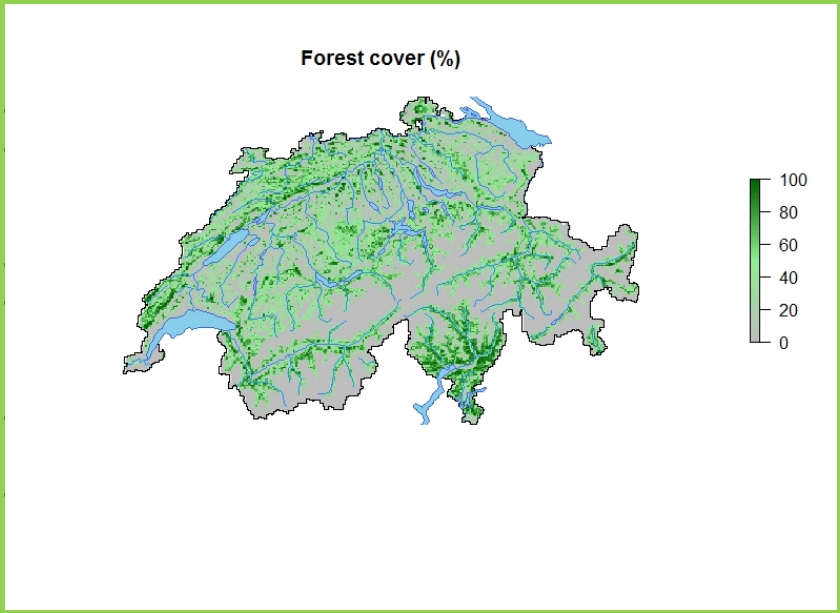






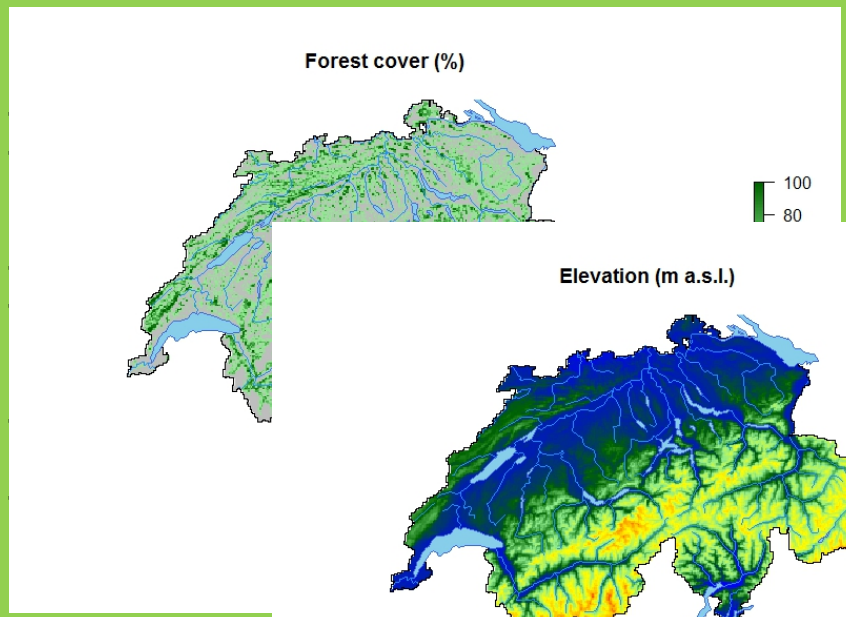
Analysis of results: Predictive maps in geographic space

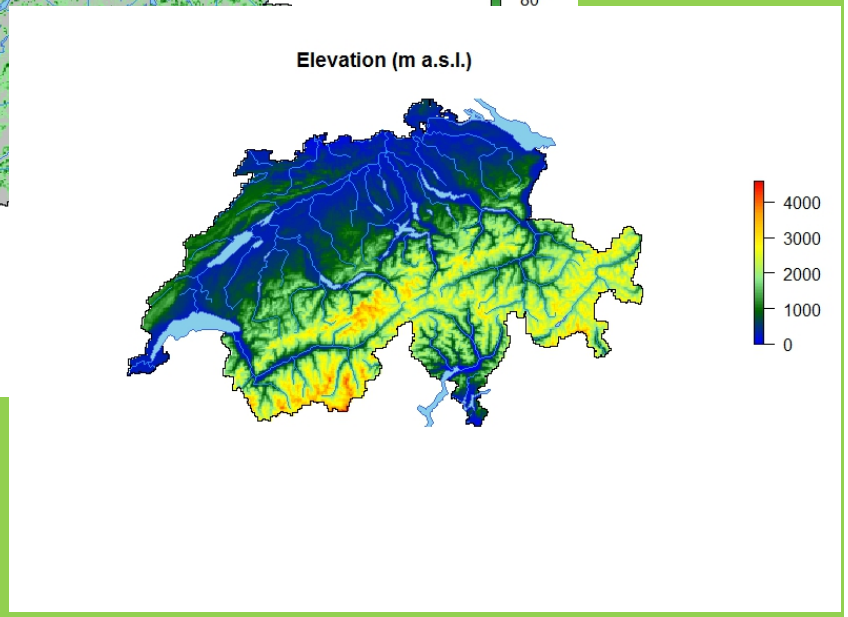
```
Parameter estimates (Colonization):
                        Estimate    SE
year1999                 -0.1305 0.237
year2000                 -0.9872 0.291
...
elev                      1.0538 0.149
I(elev^2)                -0.6380 0.208
forest                   -0.3489 0.193
I(forest^2)               0.0774 0.153
elev:forest               0.0306 0.127
elev:I(forest^2)         -0.2165 0.123
I(elev^2):forest          0.9730 0.196
I(elev^2):I(forest^2)    -0.3698 0.187
....
```



Analysis of results: Predictive maps in geographic space

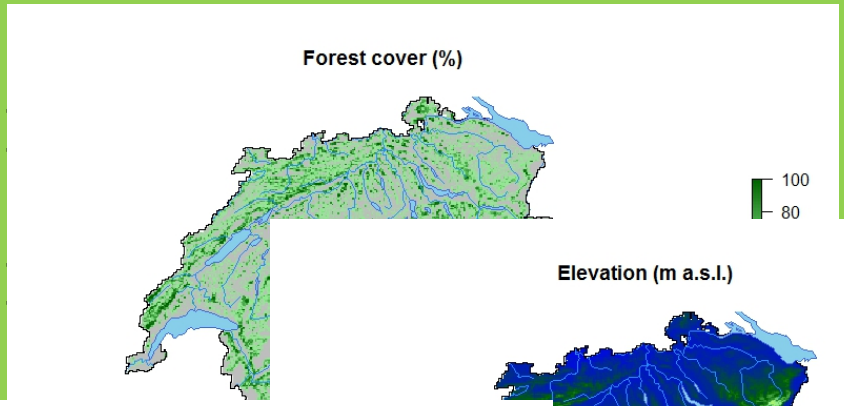
```
Parameter estimates (Colonization):
                        Estimate    SE
year1999                 -0.1305 0.237
year2000                 -0.9872 0.291
...
elev                      1.0538 0.149
I(elev^2)                -0.6380 0.208
forest                   -0.3489 0.193
I(forest^2)               0.0774 0.153
elev:forest               0.0306 0.127
elev:I(forest^2)         -0.2165 0.123
I(elev^2):forest          0.9730 0.196
I(elev^2):I(forest^2)    -0.3698 0.187
....
```

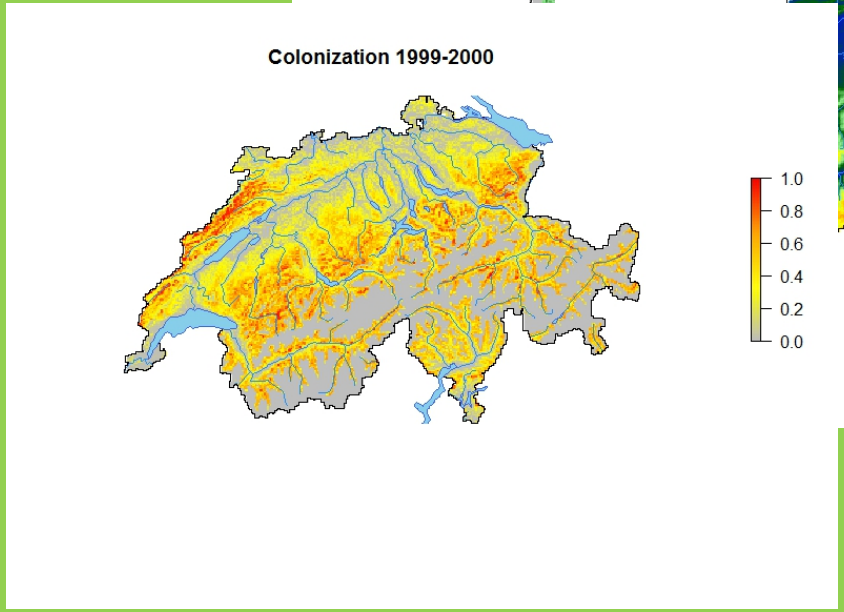


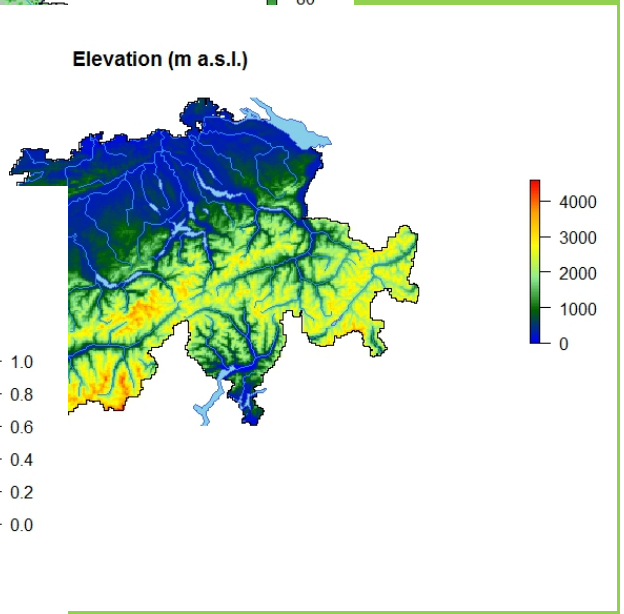


Analysis of results: Predictive maps in geographic space

```
Parameter estimates (Colonization):
                        Estimate    SE
year1999                 -0.1305 0.237
year2000                 -0.9872 0.291
...
elev                      1.0538 0.149
I(elev^2)                -0.6380 0.208
forest                   -0.3489 0.193
I(forest^2)               0.0774 0.153
elev:forest               0.0306 0.127
elev:I(forest^2)
I(elev^2):forest
I(elev^2):I(forest^2)
....
```

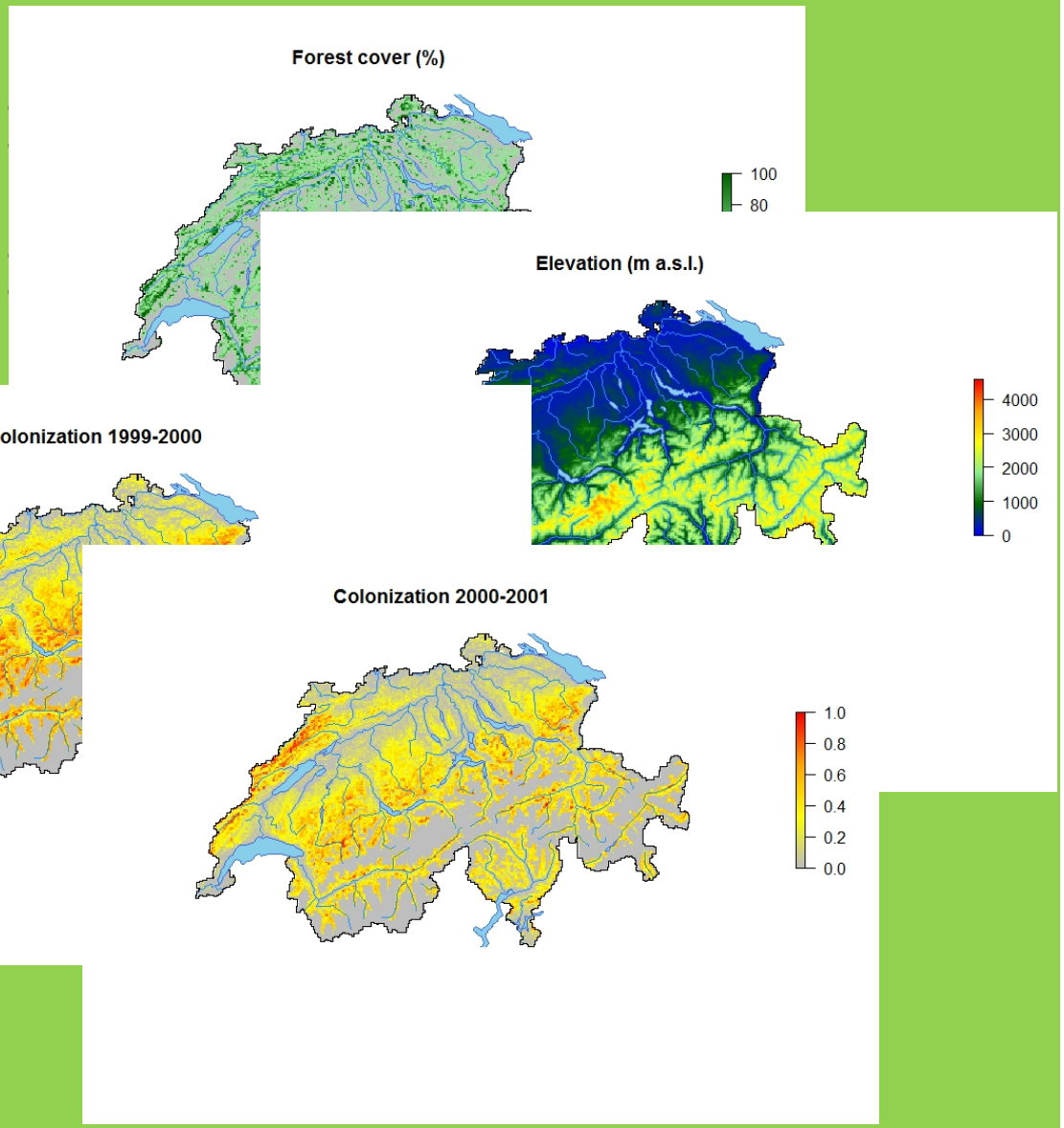






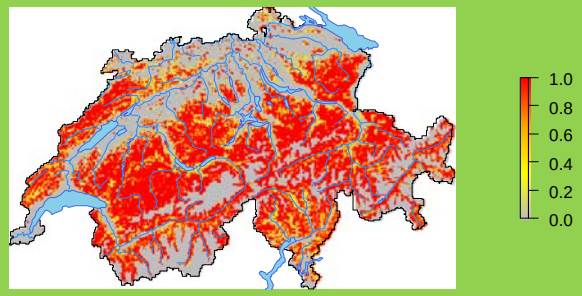
Analysis of results: Predictive maps in geographic space

```
Parameter estimates (Colonization):
                        Estimate    SE
year1999                 -0.1305 0.237
year2000                 -0.9872 0.291
...
elev                      1.0538 0.149
I(elev^2)                -0.6380 0.208
forest                   -0.3489 0.193
I(forest^2)               0.0774 0.153
elev:forest               0.0306 0.127
elev:I(forest^2)
I(elev^2):forest
I(elev^2):I(forest^2)
....
```

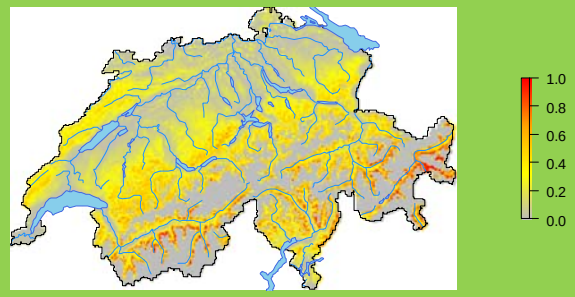


Predictive maps of occupancy probability (years 1–3)

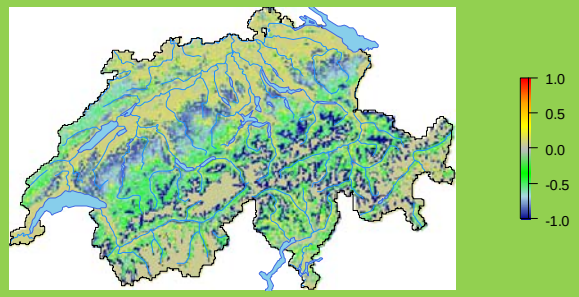
Occupancy 1999 (Dynocc)



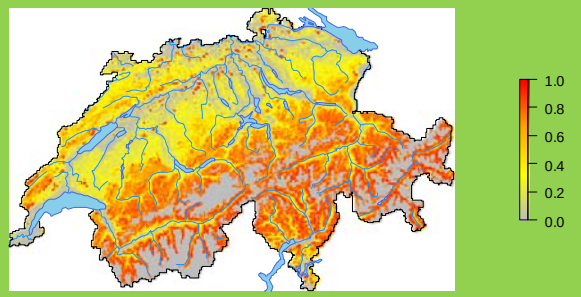
Occupancy 1999 (Trad. Metapop.)



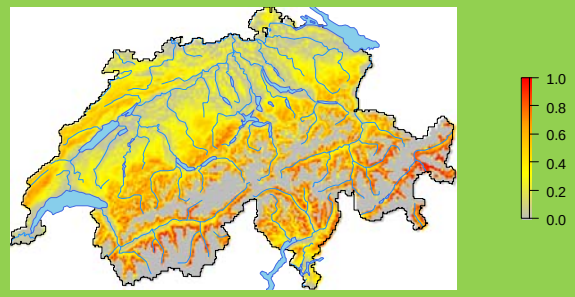
Difference Occupancy 1999



Occupancy 2000 (Dynocc)

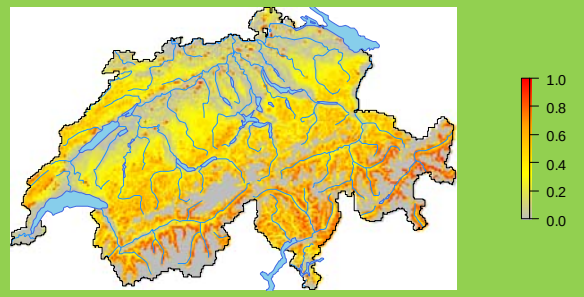


Occupancy 2000 (Trad. Metapop.)

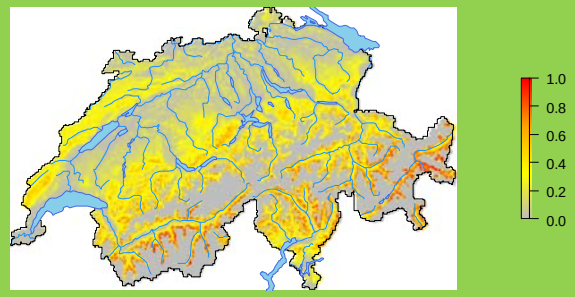


Difference Occupancy 2000

Occupancy 2001 (Dynocc)



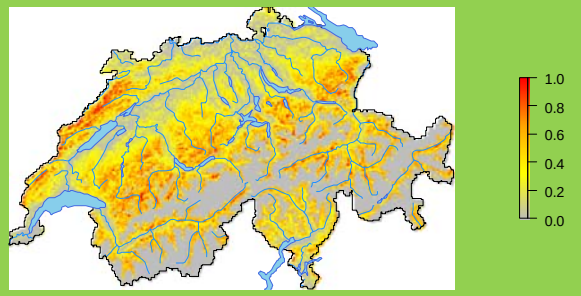
Occupancy 2001 (Trad. Metapop.)



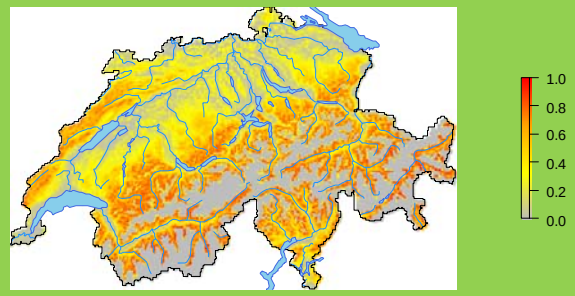
Difference Occupancy 2001

Predictive maps of colonization probability (intervals 1–3)

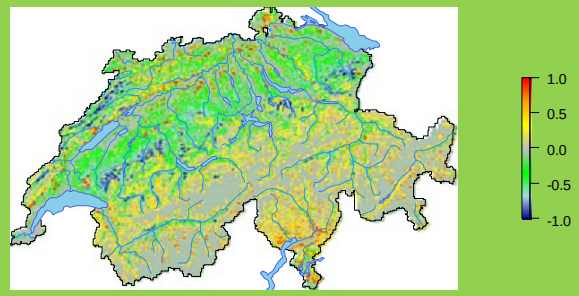
Colonization 1999-2000 (Dynocc)



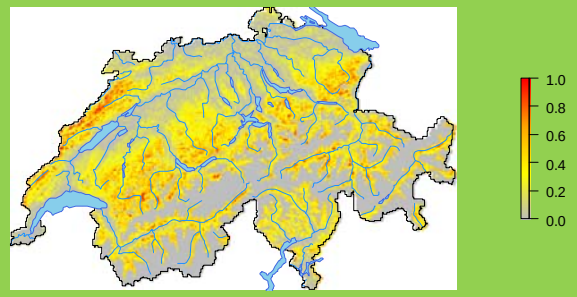
Colonization 1999-2000 (Trad. Metapop.)



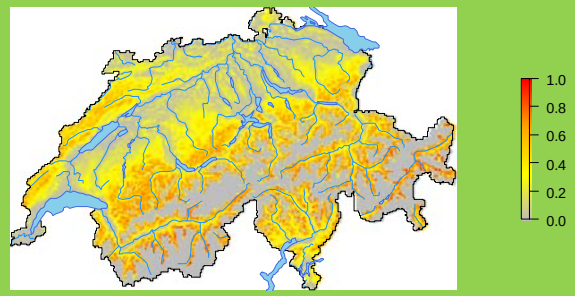
Difference Colonization 1999-2000



Colonization 2000-2001 (Dynocc)

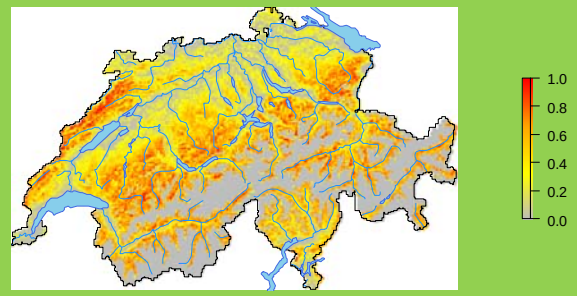


Colonization 2000-2001 (Trad. Metapop.)

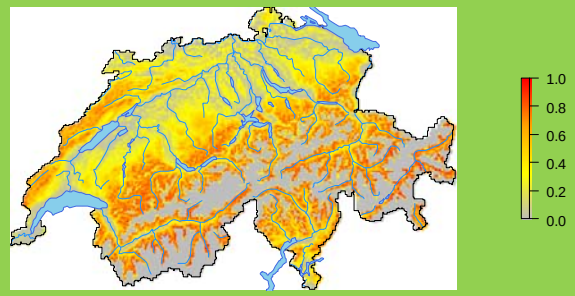


Difference Colonization 2000-2001

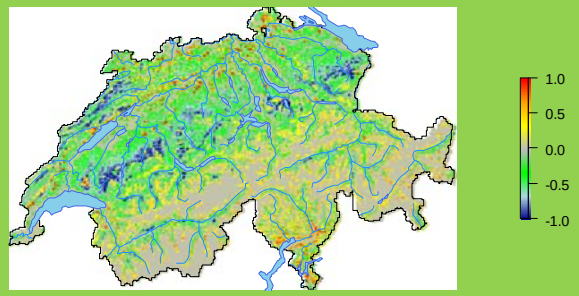
Colonization 2001-2002 (Dynocc)



Colonization 2001-2002 (Trad. Metapop.)

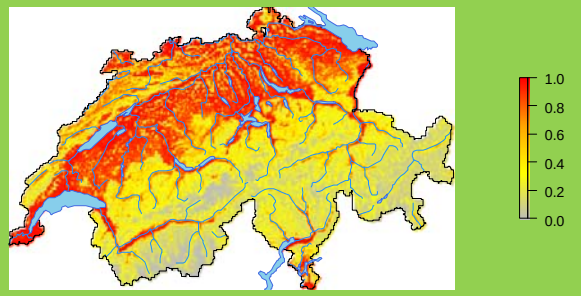


Difference Colonization 2001-2002

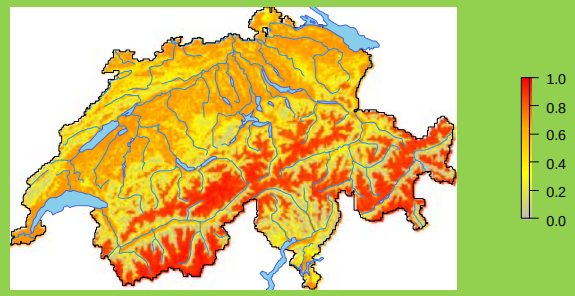


Predictive maps of extinction probability (intervals 1–3)

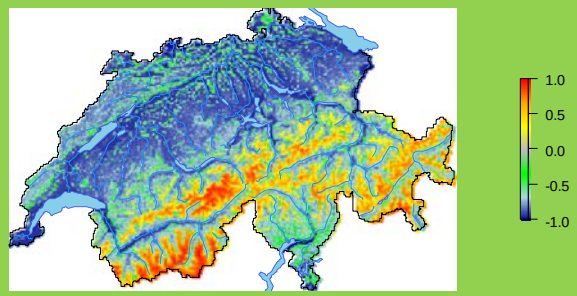
Extinction 1999-2000 (Dynocc)



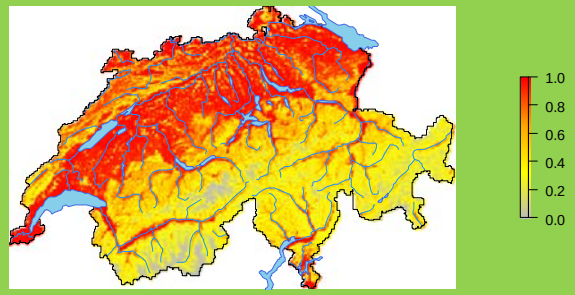
Extinction 1999-2000 (Trad. Metapop.)



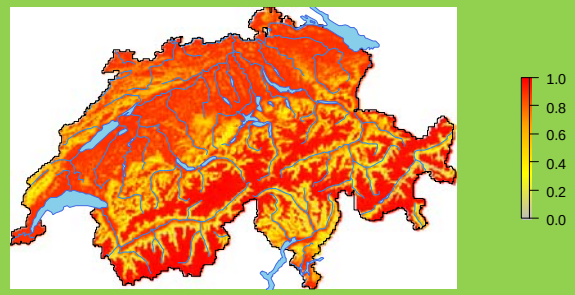
Difference Extinction 1999-2000



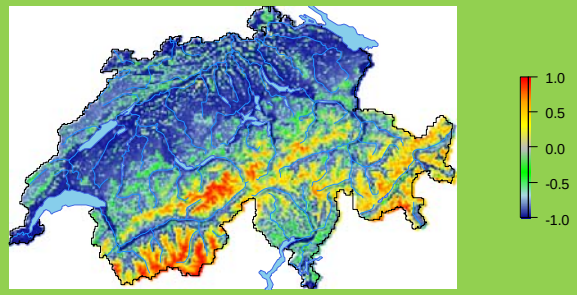
Extinction 2000-2001 (Dynocc)



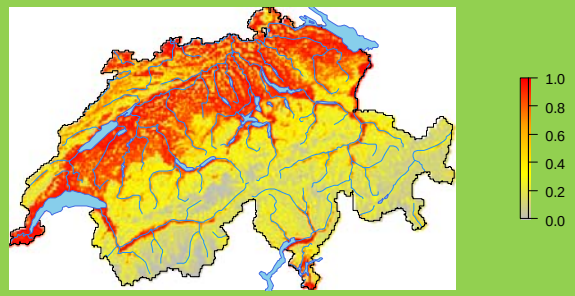
Extinction 2000-2001 (Trad. Metapop.)



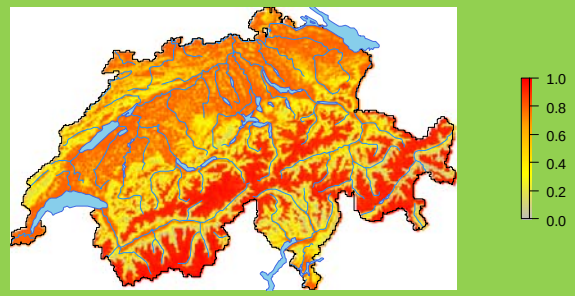
Difference Extinction 2000-2001



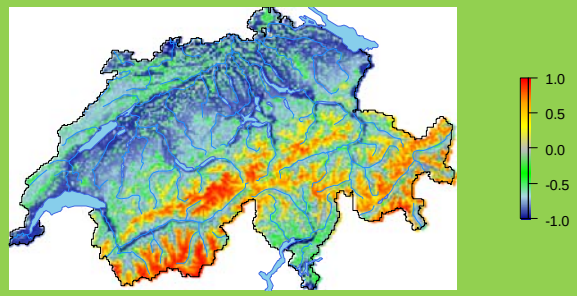
Extinction 2001-2002 (Dynocc)



Extinction 2001-2002 (Trad. Metapop.)

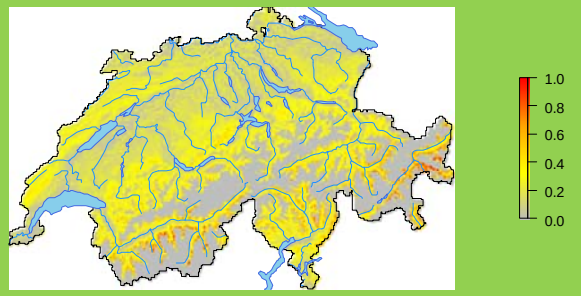


Difference Extinction 2001-2002

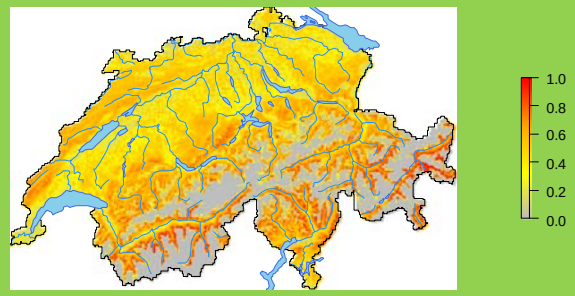


Predictive maps of detection probability

Detection 1999 (Dynocc)



Detection 2000 (Dynocc)

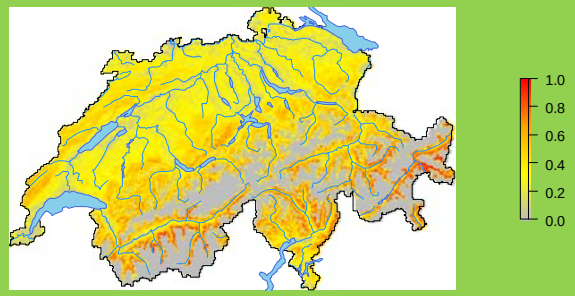


Detection 2001 (Dynocc)

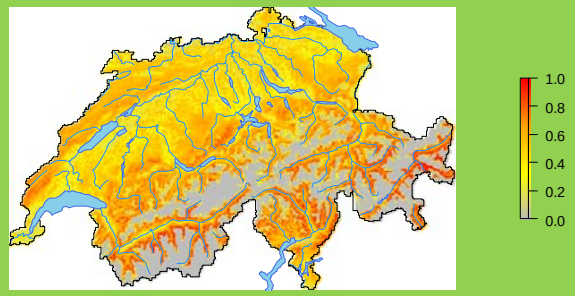
Detection 2002 (Dynocc)

Detection 2003 (Dynocc)

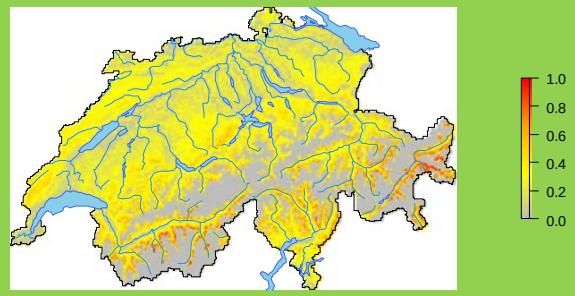
Detection 2004 (Dynocc)



Detection 2005 (Dynocc)



Detection 2006 (Dynocc)



Detection 2007 (Dynocc)

Grand summary: Dynocc models for distribution change

- Animals and plants small & discrete (“points”)
- Locations: Realization of spatial point process
- Concept of distribution requires/implies grid
- (Also, state of occurrence must be defined)
- Distribution and abundance areal-summary of spatial point process
- Distribution information-reduced summary of abundance

Grand summary: Dynocc models for distribution change

- Dynamic occupancy model: hidden Markov model, metapopulation model with accommodation of imperfect detection
- Requires within-year replicate data & “closure”
- Useful R package **unmarked**, also many others (frequentist/Bayesian)
- **Dynocc model powerful and natural framework of inference about changes in species distributions**

Grand summary: Dynocc models for distribution change

- Swiss crossbills: Dynocc model enables inference about distribution dynamics plus observation process (imperfect detection)
- Detection probability always <1
- Parameter estimates quite different whether detection estimated or not (Dynocc vs. metapopulation model)
- Predictions (maps) also quite different
- => Estimate p if possible
- Useful to map effects/predictions for spatially indexed covariates !