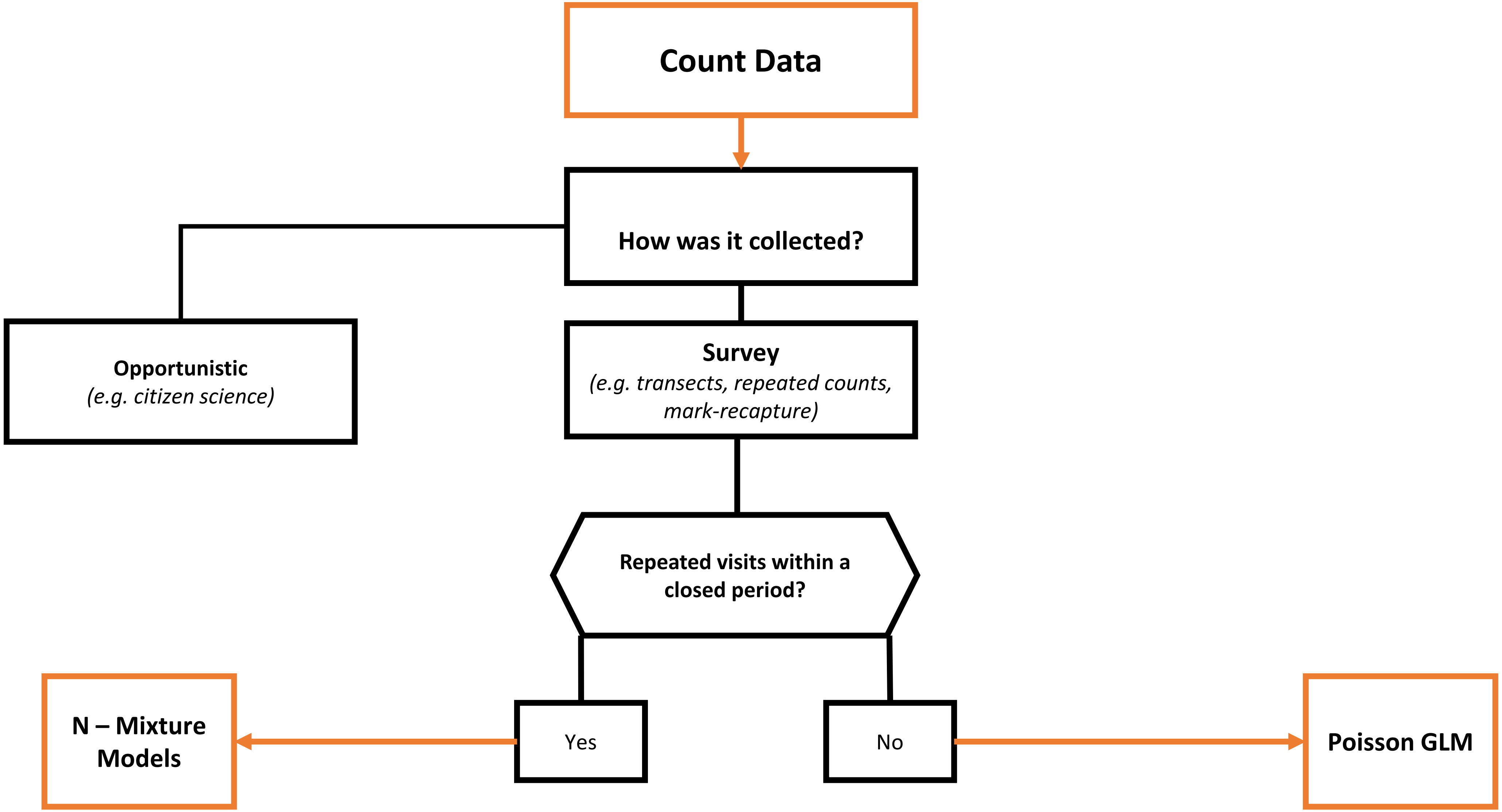
Count data

SSA 200



alabama cooperative fish and wildlife research unit

Presence/absence data vs. abundance data

- **Example 1:** A researcher conducts point counts for birds in the same forest plot each summer. During point counts, they make an effort to detect and record every species and count the individuals within the survey window.
- **Example 2:** A National Park office often gets reports from people who have seen wolverines in the park. They keep track of the locations and dates of these sights and they send two temp interns to set-up and camera traps randomly throughout the park to capture wolverine on tape.
- **Example 3:** A natural history museum has plant specimens that contain information about the location and date of collection, which can be used to map historical occurrence of the species. Associated field notebooks record the number of total unique plants along the same paths each year for 25 years.

alabama cooperative fish and wildlife research unit

U.S. FISH & WILDLIFE SERVICE

Abundance data

- When we have information at repeated places over repeated time periods
  - camera traps
  - hair/eDNA traps (unique individuals over time)
  - transect surveys
  - distance sampling

- *What else can you think of?*

Common analytical approaches

- Poisson Generalized Linear Model (GLM)
  - For non-repeated counts
- N-mixture models
  - For repeated counts

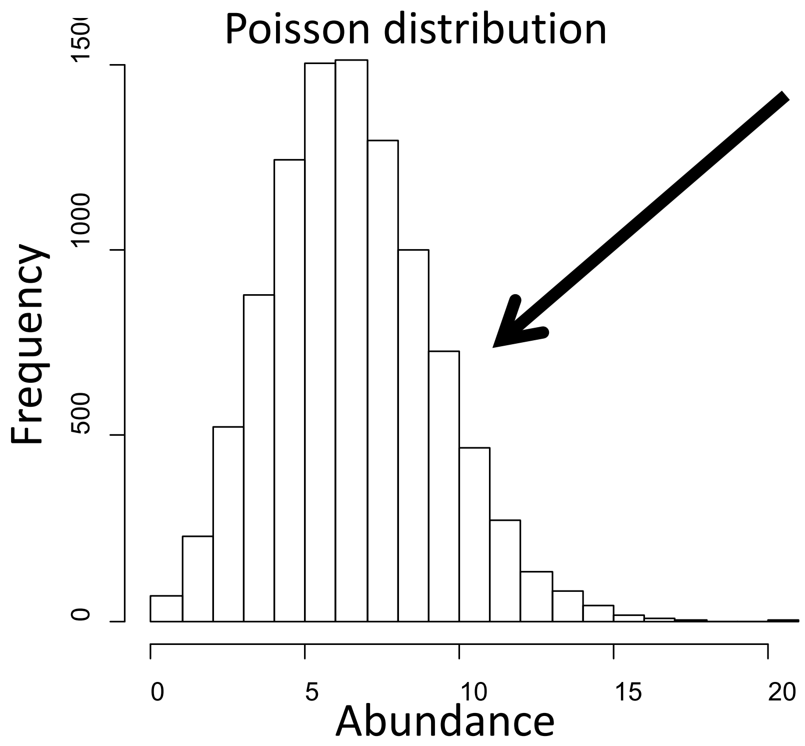
Poisson GLM

- Poisson Generalized Linear Model
    - Measure of relative abundance
- Poisson Distribution: Discrete, positive integers (0, 1, 2, ..)

- GLM: describes how explanatory variables (environmental covariates) affect an observed response (data point) through a link function

Poisson GLM

- Generalized linear model (GLM)



$$N = \alpha + \beta X$$

Environmental/habitat covariates

GLM Extensions

- Can account for lots of 0 counts in data with different statistical distributions
    - Negative binomial
    - Zero-inflated Poisson

GLM Assumptions

- Count data reasonably follow the chosen distribution
- Perfect detection of individuals
    - Counts = censuses of all animals present

Issues with estimating abundance

1. Detectability
2. Spatial variation

Issues with estimating abundance

1. Detectability

Issues with estimating abundance

1. Detectability

6,925 Retweets

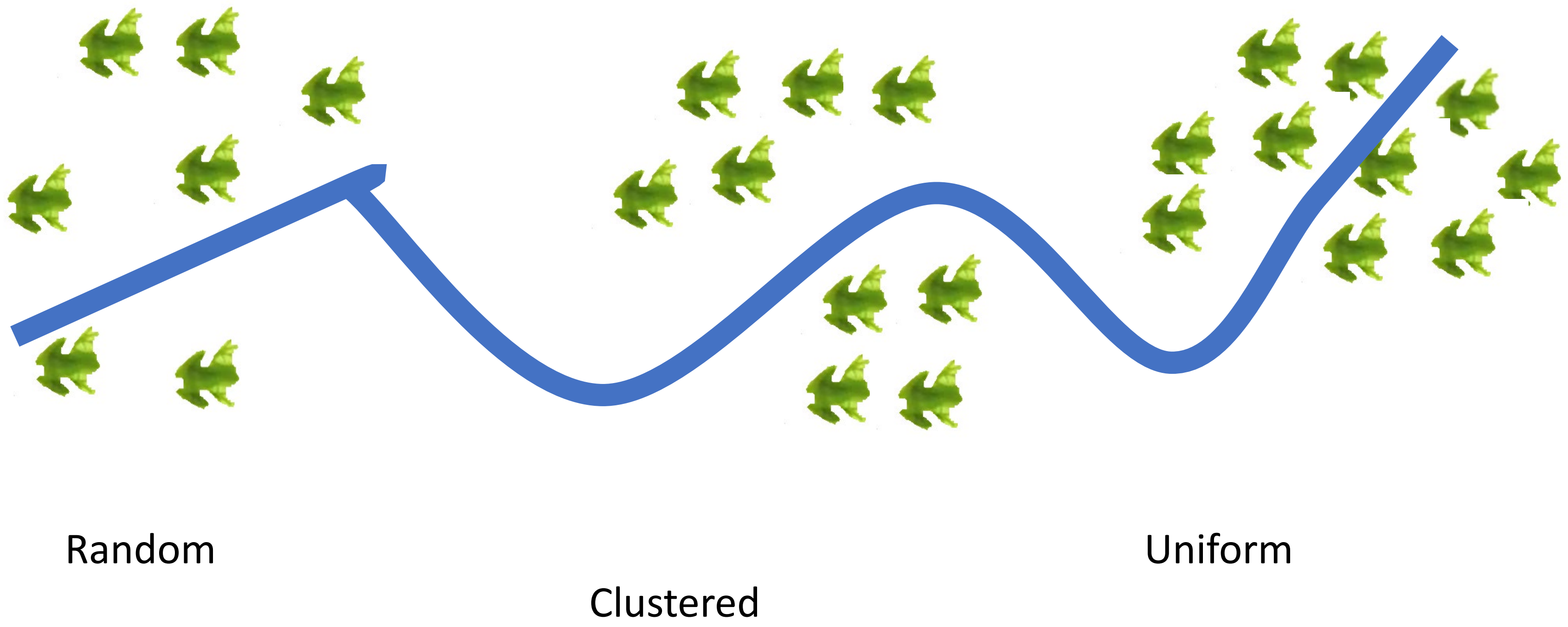
23,876 Likes

Issue 1: Detectability

- Availability
- Detection
- Intraspecific variation

Mackenzie et al 2002. Estimating site occupancy rates when detection probabilities are less than one. *Ecology*

Issue 2: Spatial variation



N-Mixture Models address both issues

- Use *repeated counts* at *several sites* to estimate detection probability directly
- Can include covariates associated with either abundance or detection
  - Explicitly model spatial and temporal variation
- Called “mixture” because it combines two GLMs
  - Poisson GLM – abundance
  - Binomial GLM (Logistic regression) – detection

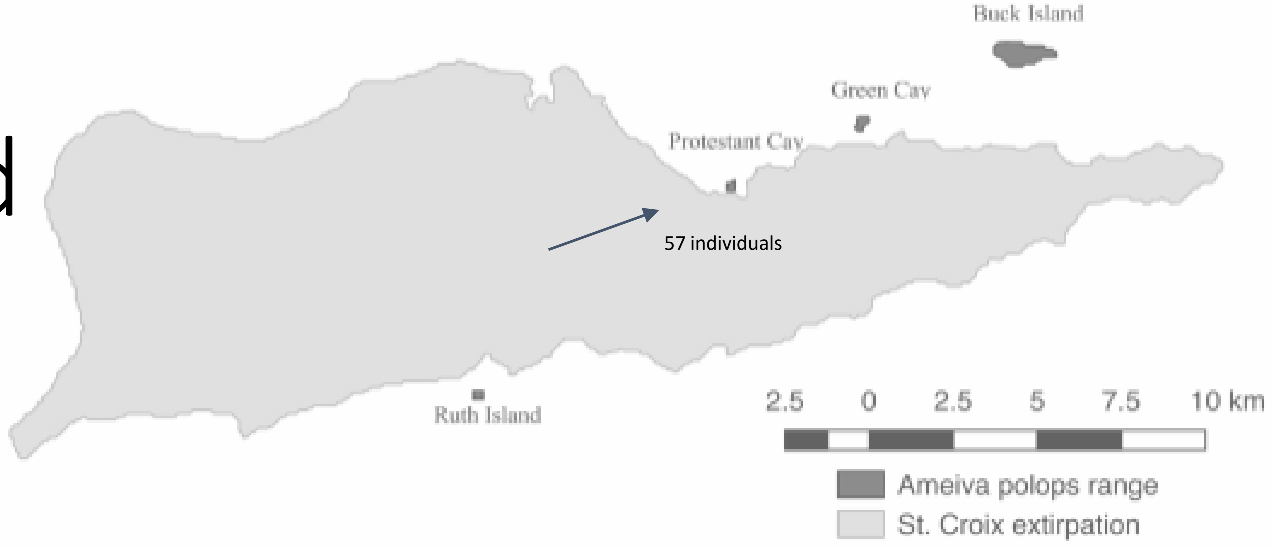
alabama cooperative fish and wildlife research unit

N-Mixture Model assumptions

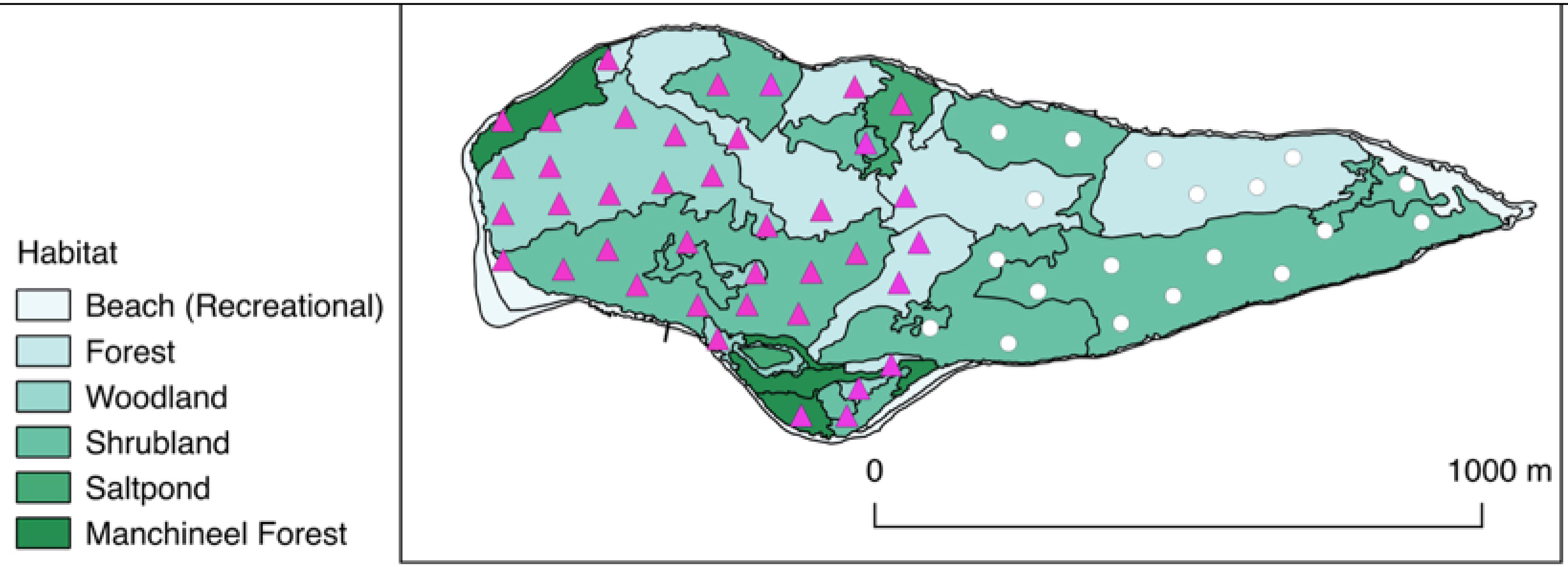
- Sites closed to immigration/emigration between surveys
- Detection process is independent at each site, can vary among sites

- **These assumptions should be addressed in the sampling design**

St. Croix ground lizard



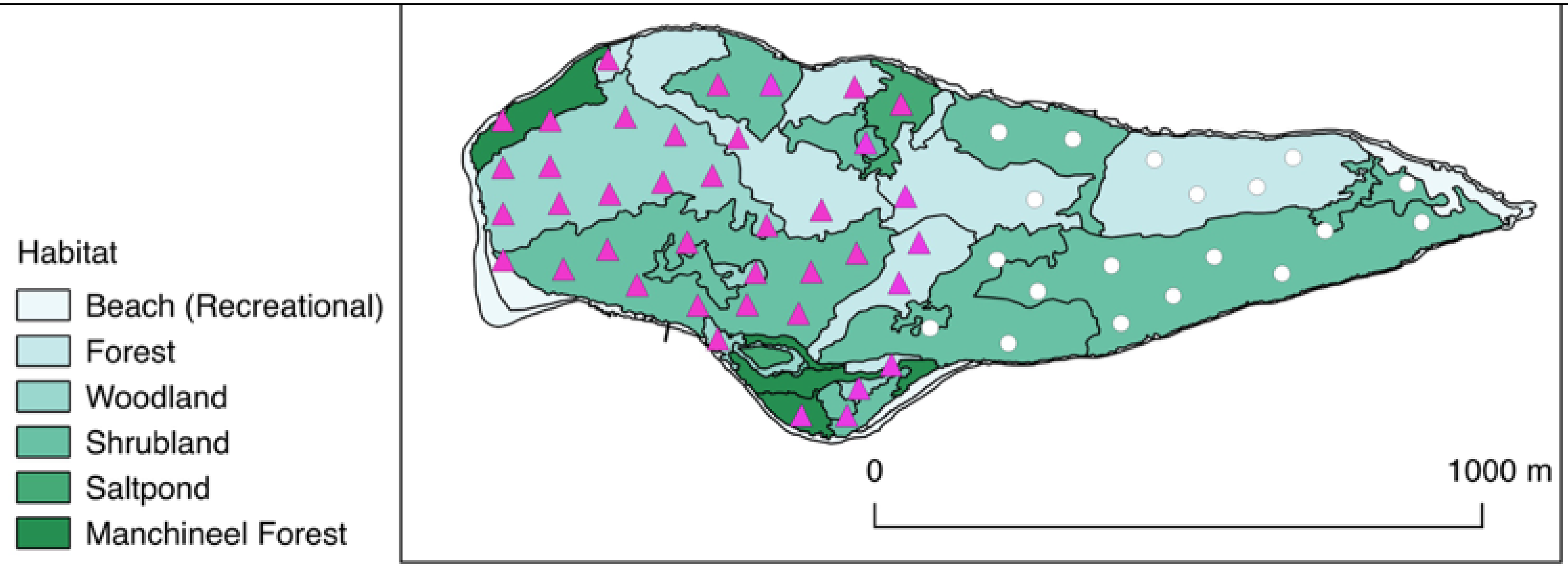
Assumption: Site independence



Site (i) – 42 sites
Survey (j) – 6 surveys
$Y_{i,j}$ –12,900 observations

- Each site was >80 m apart
- Sites randomly stratified by habitat

Assumption: Site closure



Site (i) – 42 sites
Survey (j) – 6 surveys
$Y_{i,j}$ –12,900 observations

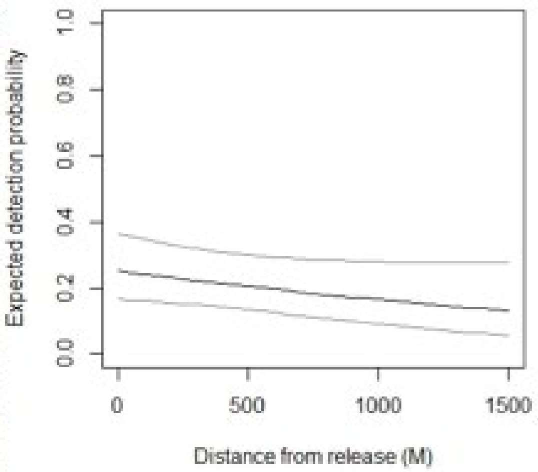
- Repeated surveys within 4 days

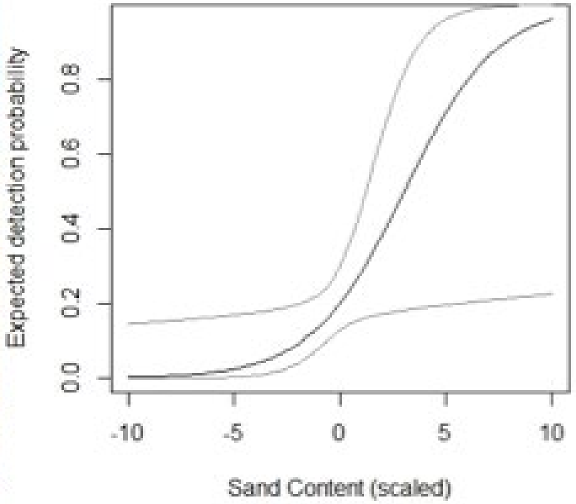
Covariates and model selection

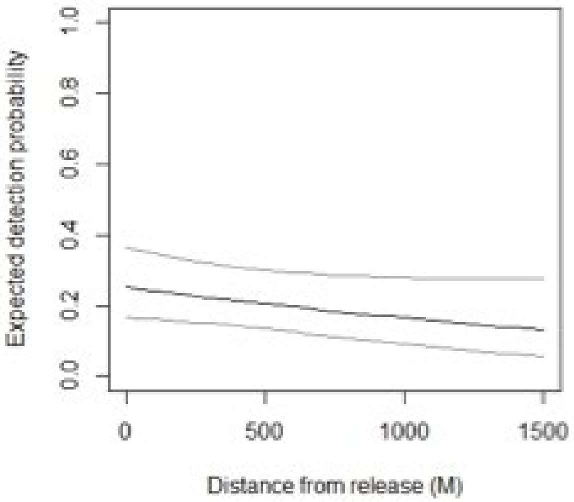
Abundance Only	# Parameters	AICc	w(AIC)
λ(Avg surface+Avg OTM+Dist+LLd+CWD)σ(.)α	7	943.3	0.281
λ(Avg OTM+Dist_Encl+LLd+Woody)σ(.)α	6	943.5	0.251
λ(Avg surface+Avg OTM +Dist+LLd+Max OTM+Woody)σ(.)α	8	945.3	0.103

Detection Only	# Parameters	AICc	w(AIC)
λ(.)σ(Avg OTM+%Cov+Dist+Hab+Max surface+Max OTM)α	9	1165.1	0.337
λ(.)σ(Avg OTM+%Cov+Dist+Hab+LLD+Max surface+Max OTM)α	10	1165.9	0.226
λ(.)σ(Avg OTM+%Cov+Dist+Hab+Max surface+Max OTM+obs)α	13	1166.5	0.163

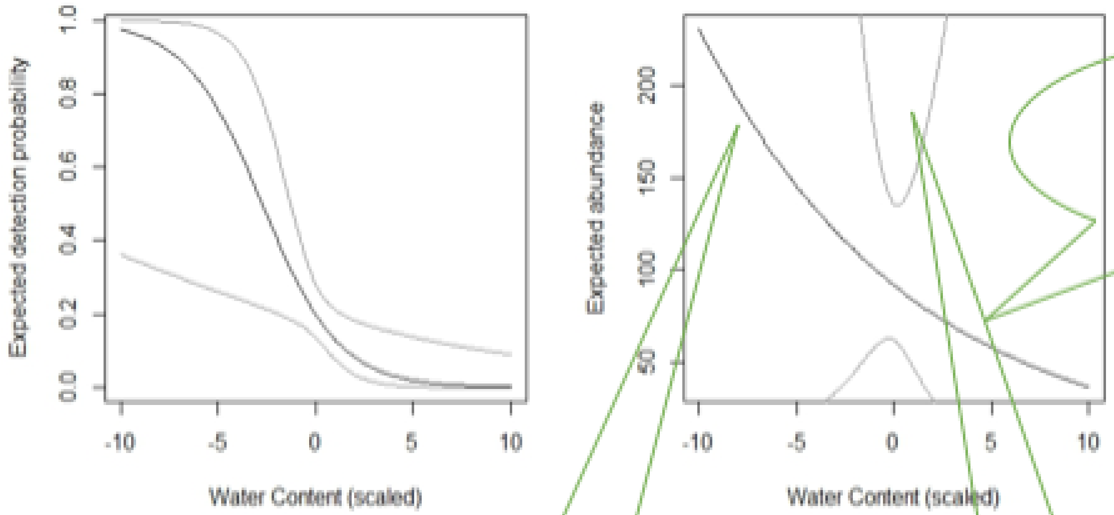
Detection varied among habitat types







Covariates



Here, more water, fewer lizards
Poor confidence

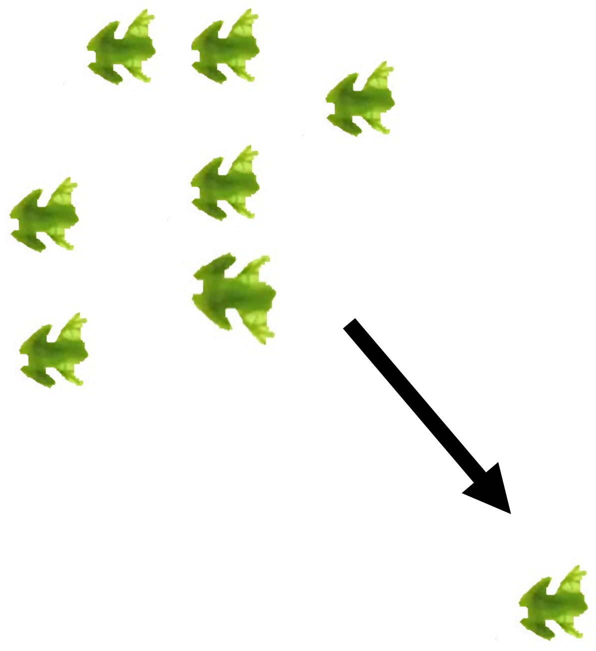
- Water content was associated with both detection probability and abundance

- What's going on with the confidence intervals?

Here, less water, more lizards
Poor confidence

Here, not too much or too little water, best abundance confidence intervals

N-Mixture Models and site-occupancy models



- Occupancy and N-mixture models have a lot in common
  - Sampling design
  - Assumptions
  - Mathematical underpinnings
- The same dataset could be used in both ways
  - Collapse all counts into 1 or 0
- Occupancy models are the poor man's abundance
- Prioritize N-mixture models if possible

Presence/absence data vs. abundance data

- **Example 1:** A researcher conducts point counts for birds in the same forest plot each summer. During point counts, they make an effort to detect and record every species and count the individuals within the survey window.

Abundance data

alabama cooperative fish and wildlife research unit

Presence/absence data vs. abundance data

- **Example 2:** A National Park office often gets reports from people who have seen wolverines in the park. They keep track of the locations and dates of these sights and they send two temp interns to set-up and camera traps randomly throughout the park to capture wolverine on tape.

Abundance data

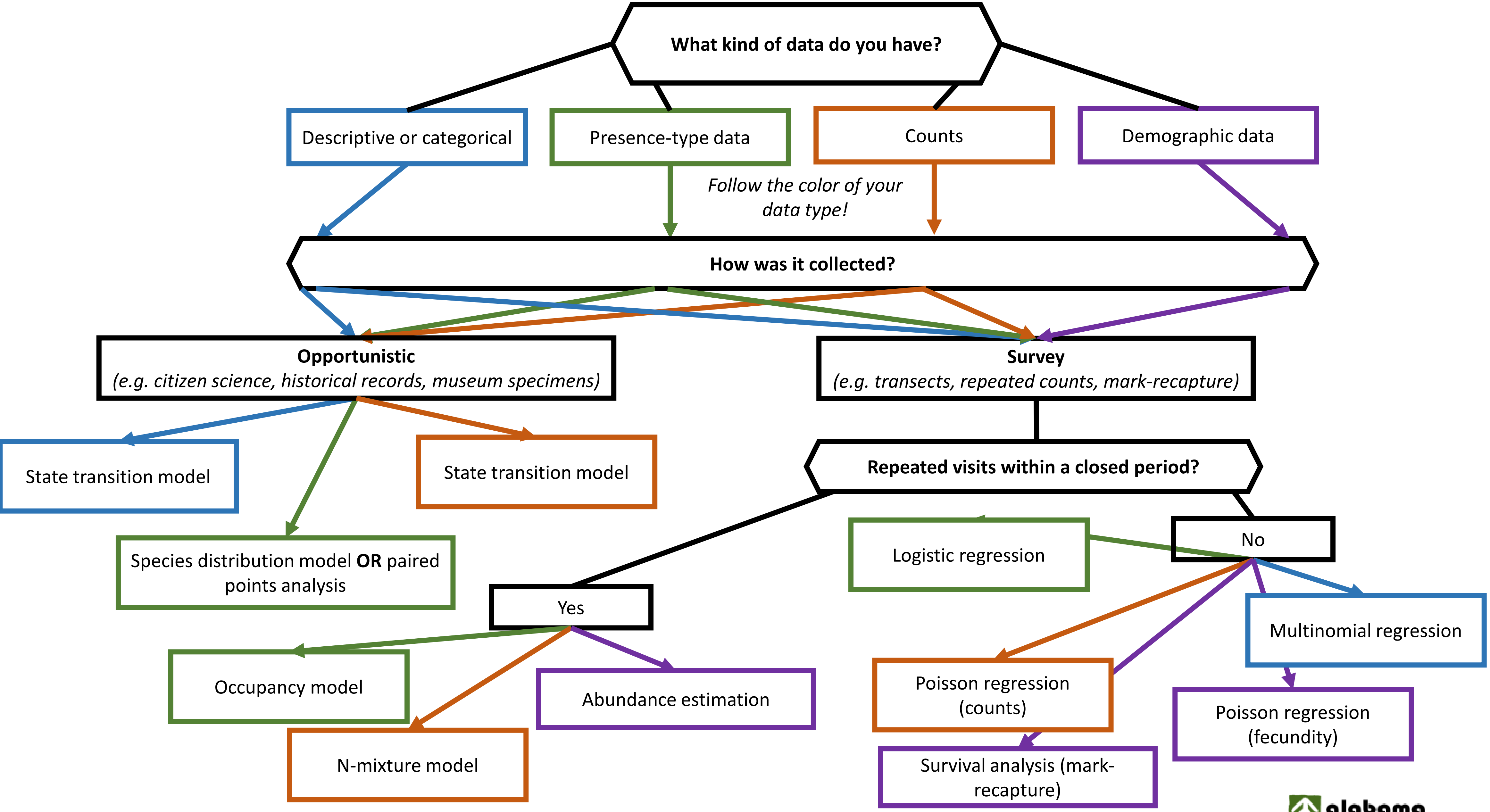
Presence/absence data vs. abundance data

- **Example 3:** A natural history museum has plant specimens that contain information about the location and date of collection, which can be used to map historical occurrence of the species. Associated field notebooks record the number of total unique plants along the same paths each year for 25 years.

Presence - absence data

alabama cooperative fish and wildlife research unit

U.S. FISH & WILDLIFE SERVICE



*This roadmap is to serve as a general guide and is not an exhaustive list of all analysis options. Also, **always check the specific assumptions of your planned modeling approach!***

alabama cooperative fish and wildlife research unit

U.S. FISH & WILDLIFE SERVICE

Questions?

Extra perspective (optional)
N-mixture models are occupancy models

Model 1: state

$z_{i,t} \sim \text{Bernoulli}(\psi_i)$

Model 2: observation

$y_{i,j,t} \sim \text{Bernoulli}(z_{i,t}, p_{i,j,t})$

MacKenzie et al. (*Ecol.*, 2002); Tyre et al. (*Ecol. App.*, 2003)

Model 1: state

$N_{i,t} \sim \text{Poisson}(\lambda_i)$

Model 2: observation

$y_{i,j,t} \sim \text{Binomial}(N_i, p_{i,t})$

Royle & Nichols, (*Ecol*, 2003)

Model 1: state

$N_{i,t} \sim \text{Poisson}(\lambda_i)$

Model 2: observation

$y_{i,j,t} \sim \text{Multinom}(N_i, f(p_{it}))$

Royle, (*Animal Bio. and Cons*, 2004)

alabama cooperative fish and wildlife research unit