A statistical explanation of MaxEnt for ecologists

Presentation by David J. Harris
paper by Jane Elith et al. 2010

Altitude data from Worldclim
Blue Jay occurrence data from Breeding Bird Survey

Maxent

Maxent

- Widely used

Maxent

- Widely used
- Good performance

Maxent

- Widely used
- Good performance
- Point-and-click workflow

Maxent

- Widely used
- Good performance
- Point-and-click workflow

- Totally foreign jargon/notation/approach

Maxent

- Widely used
- Good performance
- Point-and-click workflow

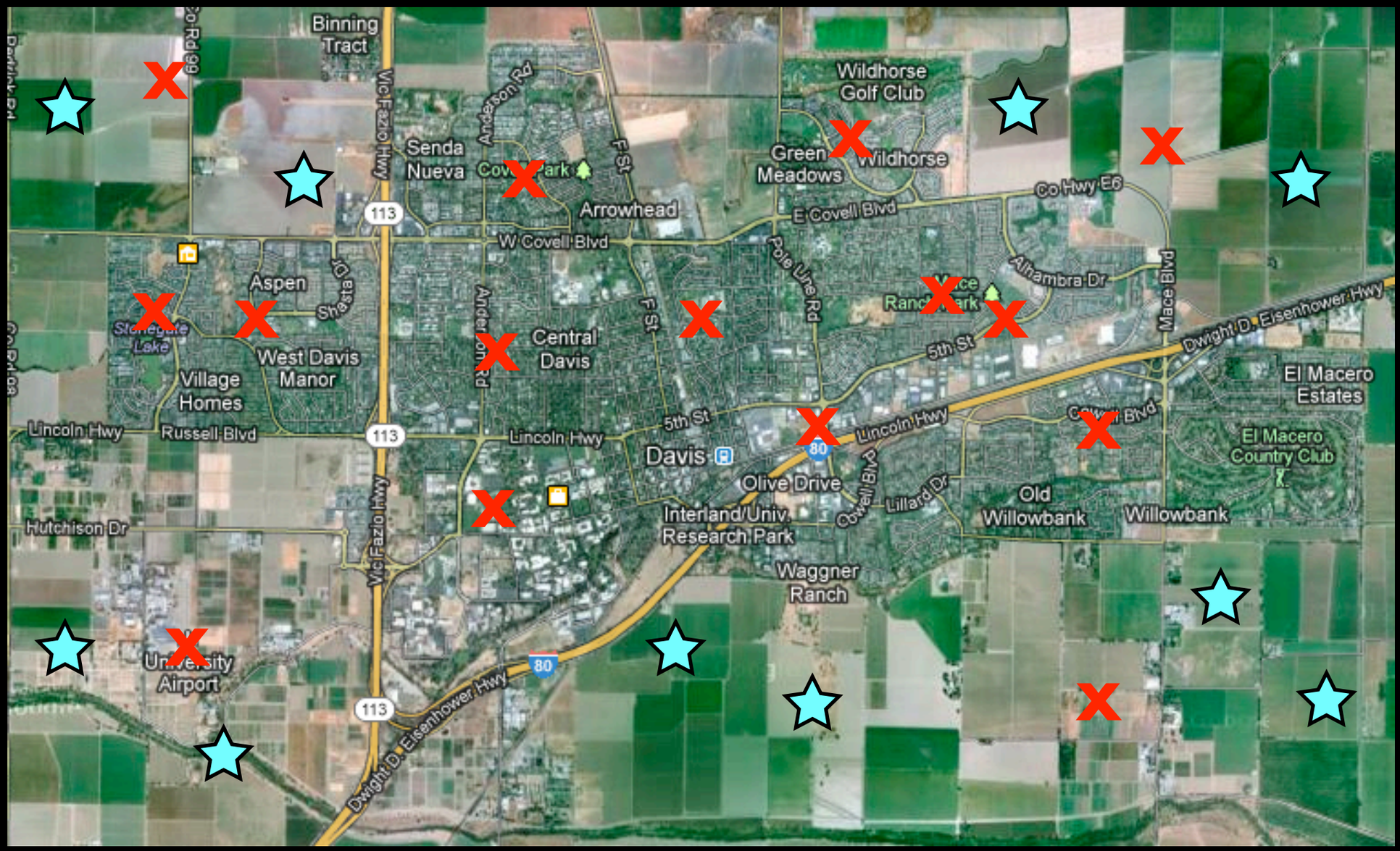
- Totally foreign jargon/notation/approach
- Elith et al. 2010 to the rescue!

Maxent *without all that "entropy" stuff*

- Widely used
- Good performance
- Point-and-click workflow

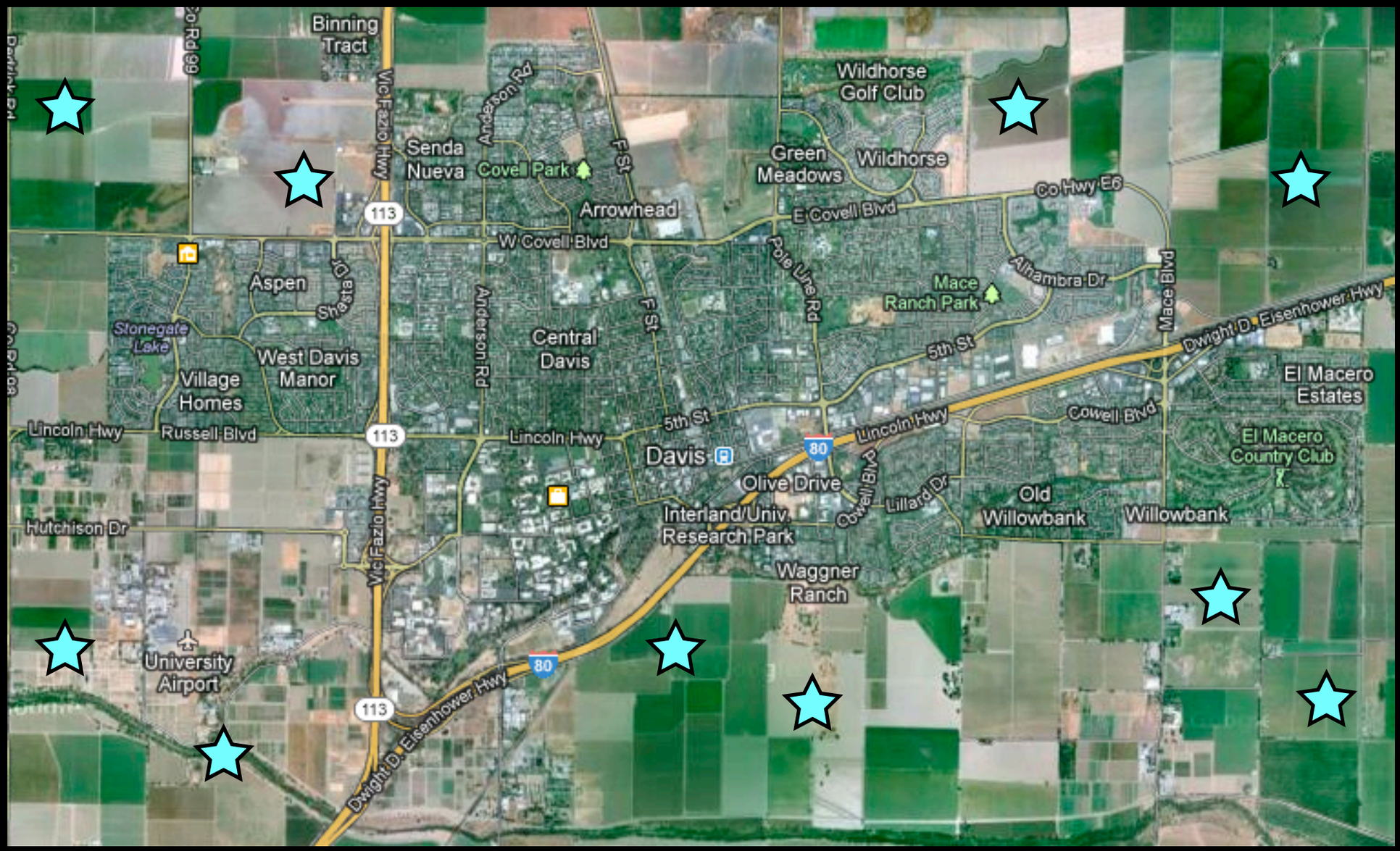
- Totally foreign jargon/notation/approach
- Elith et al. 2010 to the rescue!

Presence-absence modeling



Google Maps

Presence-only modeling



Presence-only modeling

No idea:

- Where sampling took place
- How many samples were taken

Presence-only modeling

No idea:

- Where sampling took place
- How many samples were taken

Need to assume samples came independently from a known distribution

Presence-only modeling

Presence-only modeling

P(observation | environment)

Presence-only modeling

~~*P(observation | environment)*~~

Presence-only modeling

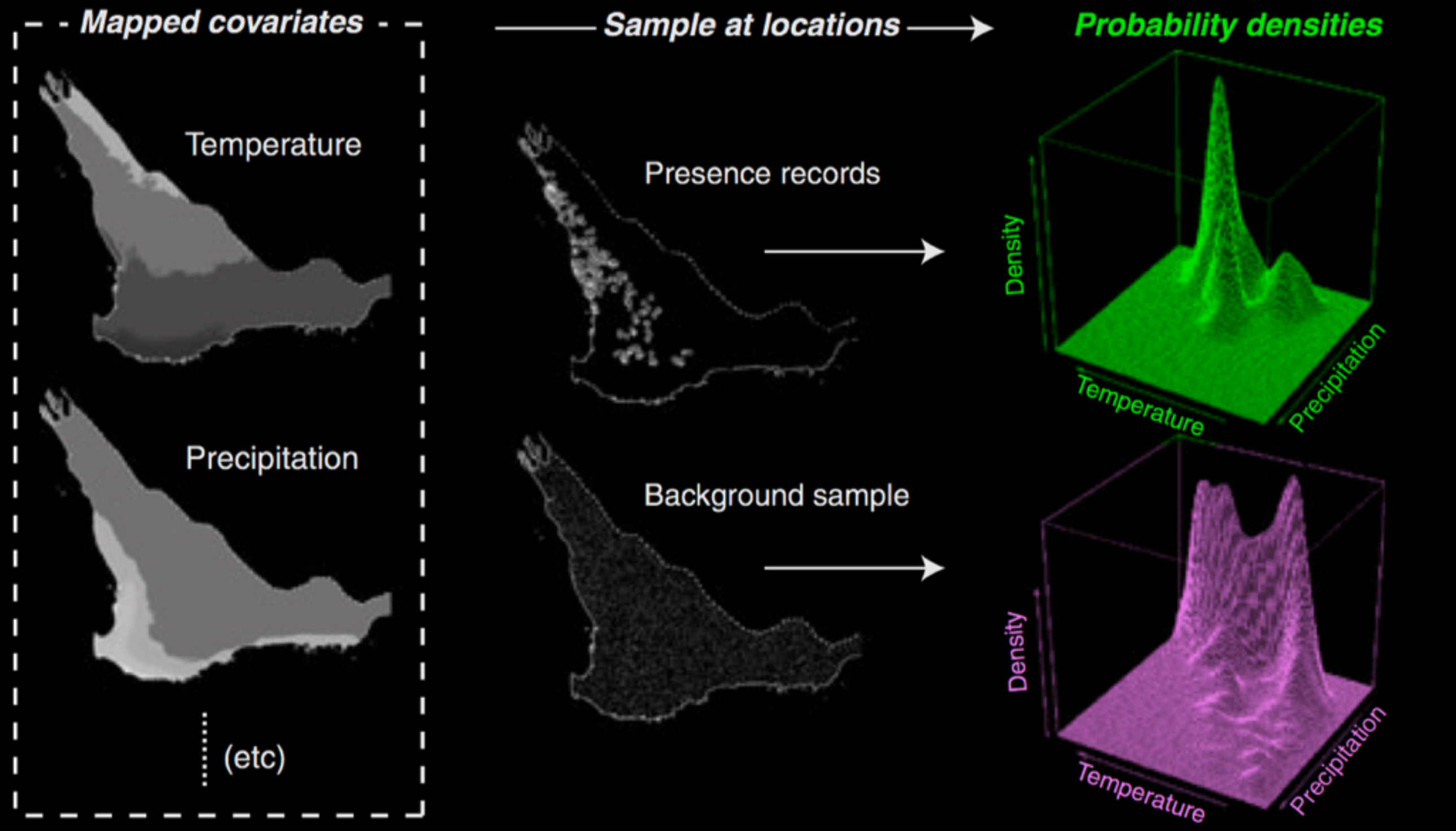
~~*P(observation | environment)*~~

P(environment | observation)

Presence-only modeling

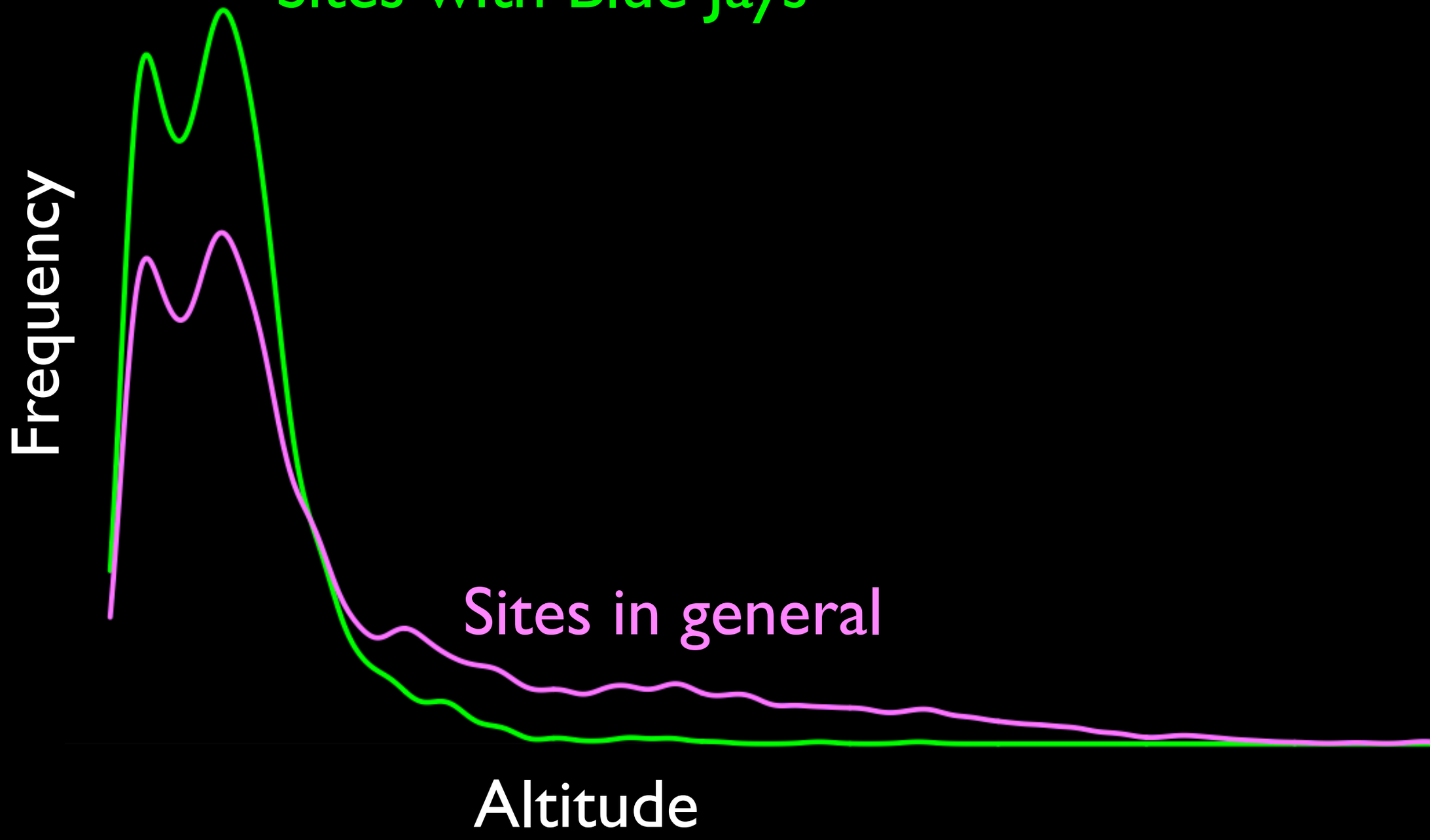
~~*P(observation | environment)*~~

$$\frac{P(environment \mid observation)}{P(environment)}$$



Elith et al. 2010

Sites with Blue Jays



Sites with Blue Jays

p(Altitude | observed)

p(Altitude)

Frequency

Sites in general

Altitude

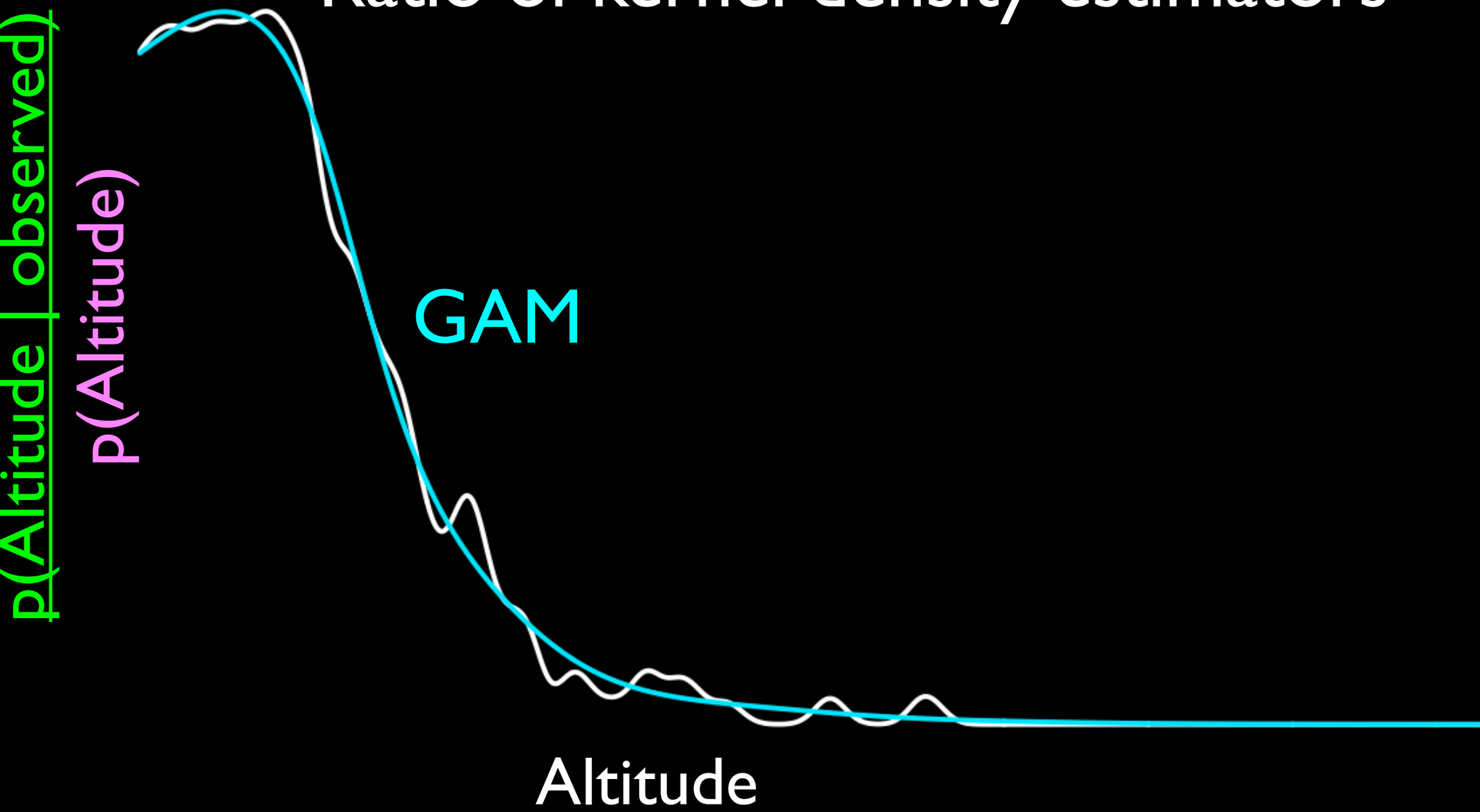
Sites with Blue Jays

$$p(\text{observed} \mid \text{Altitude}) \propto \frac{p(\text{Altitude} \mid \text{observed})}{p(\text{Altitude})}$$

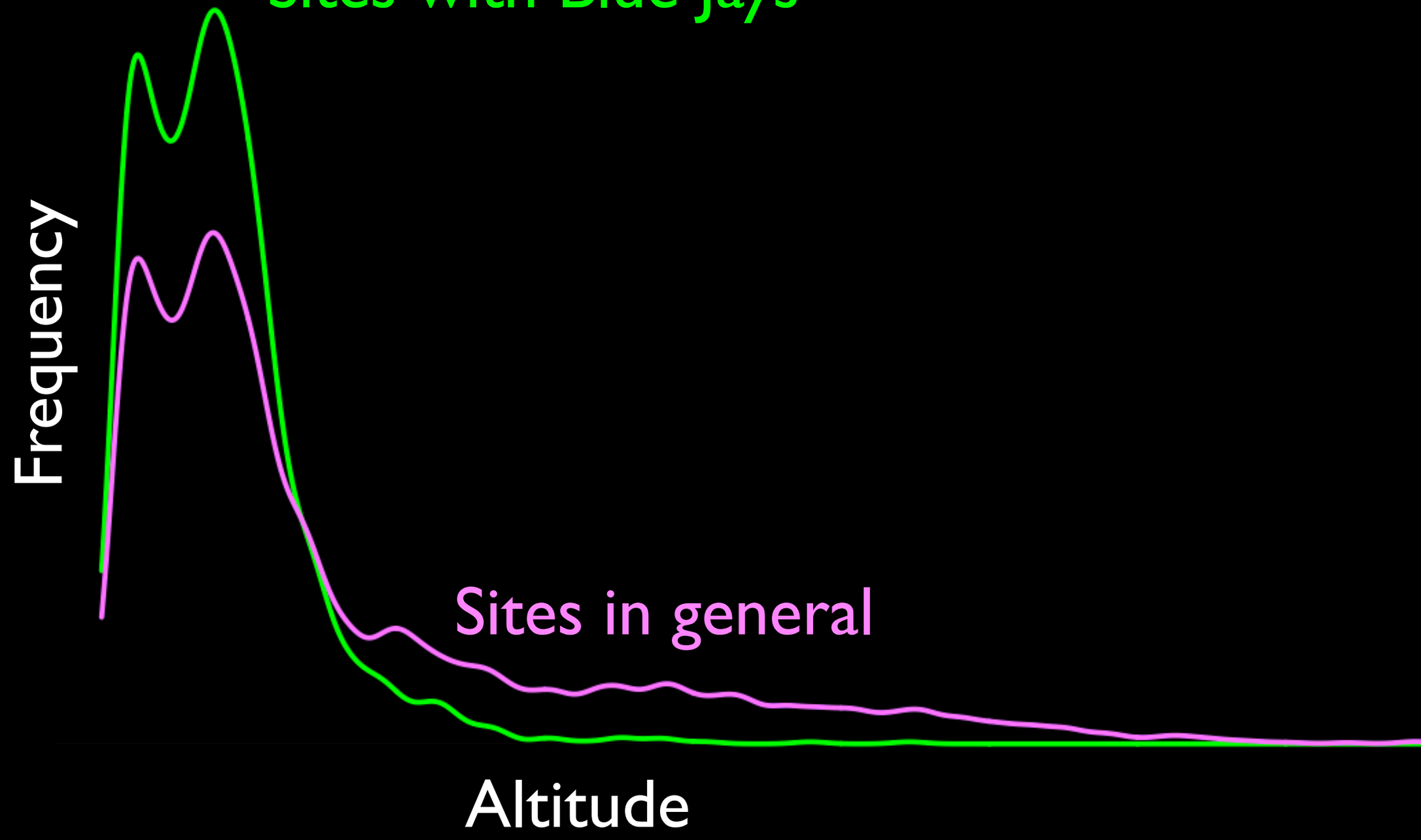
Sites in general

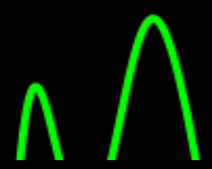
Altitude

Ratio of kernel density estimators

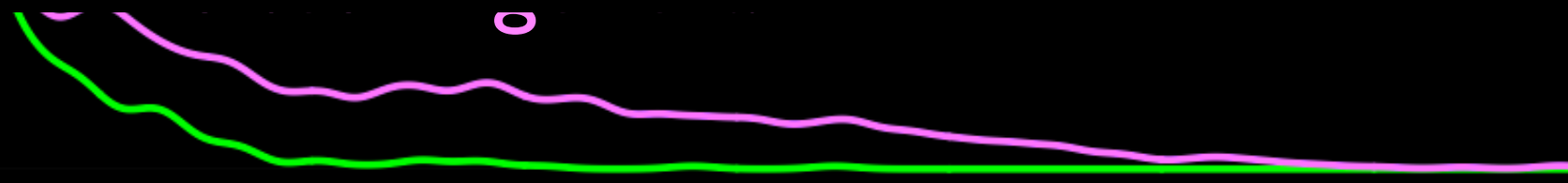


Sites with Blue Jays

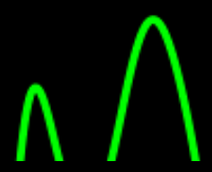




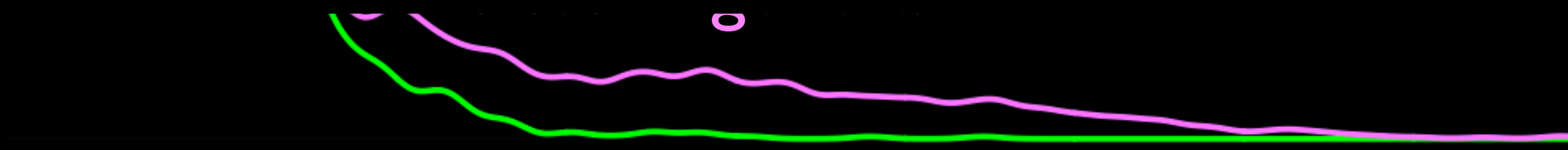
What's the best way to estimate p(Altitude | observed) when observations are limited?



Altitude



How would we estimate it if we had *no occurrence data?*



Altitude

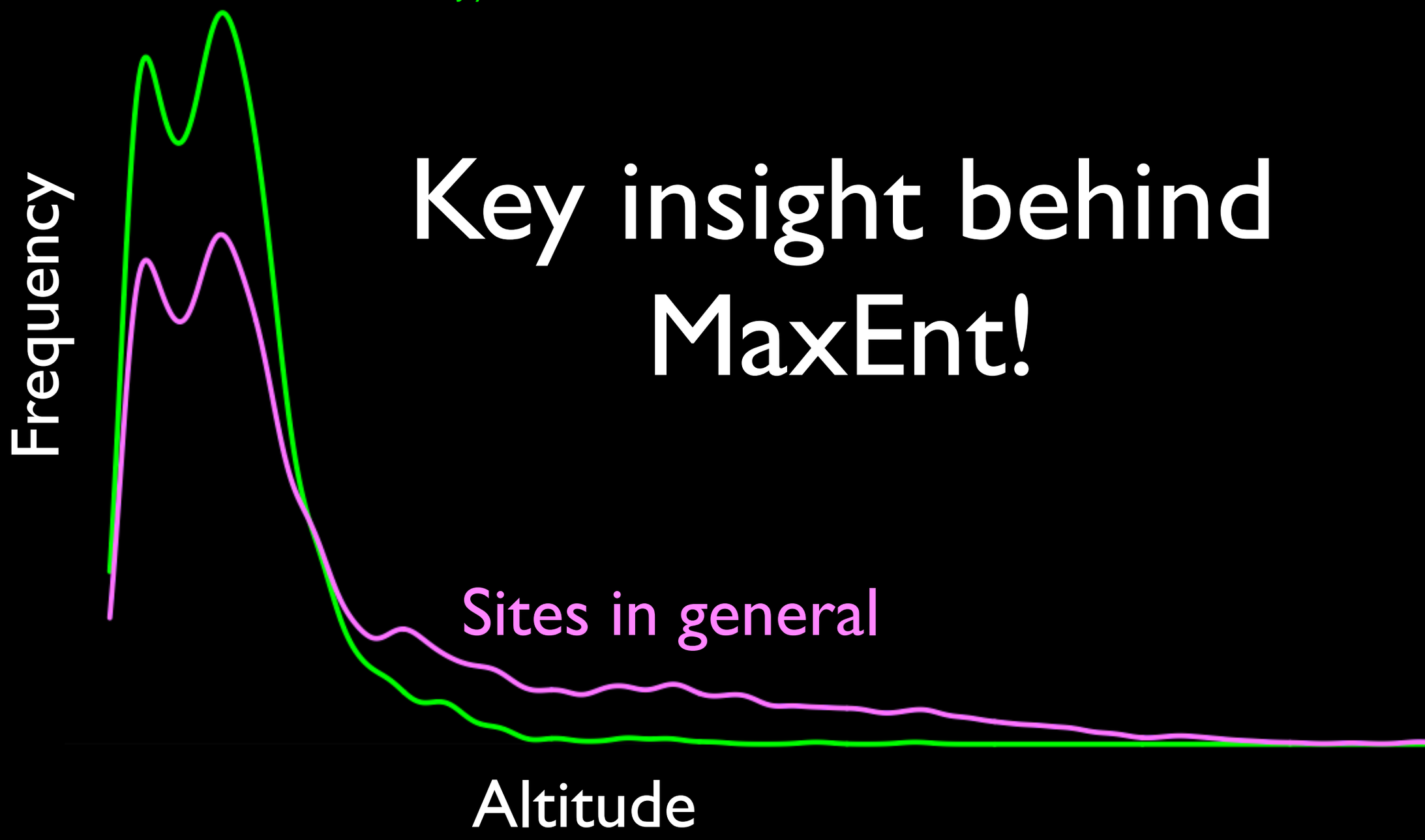
Sites with Blue Jays

Key insight behind MaxEnt!

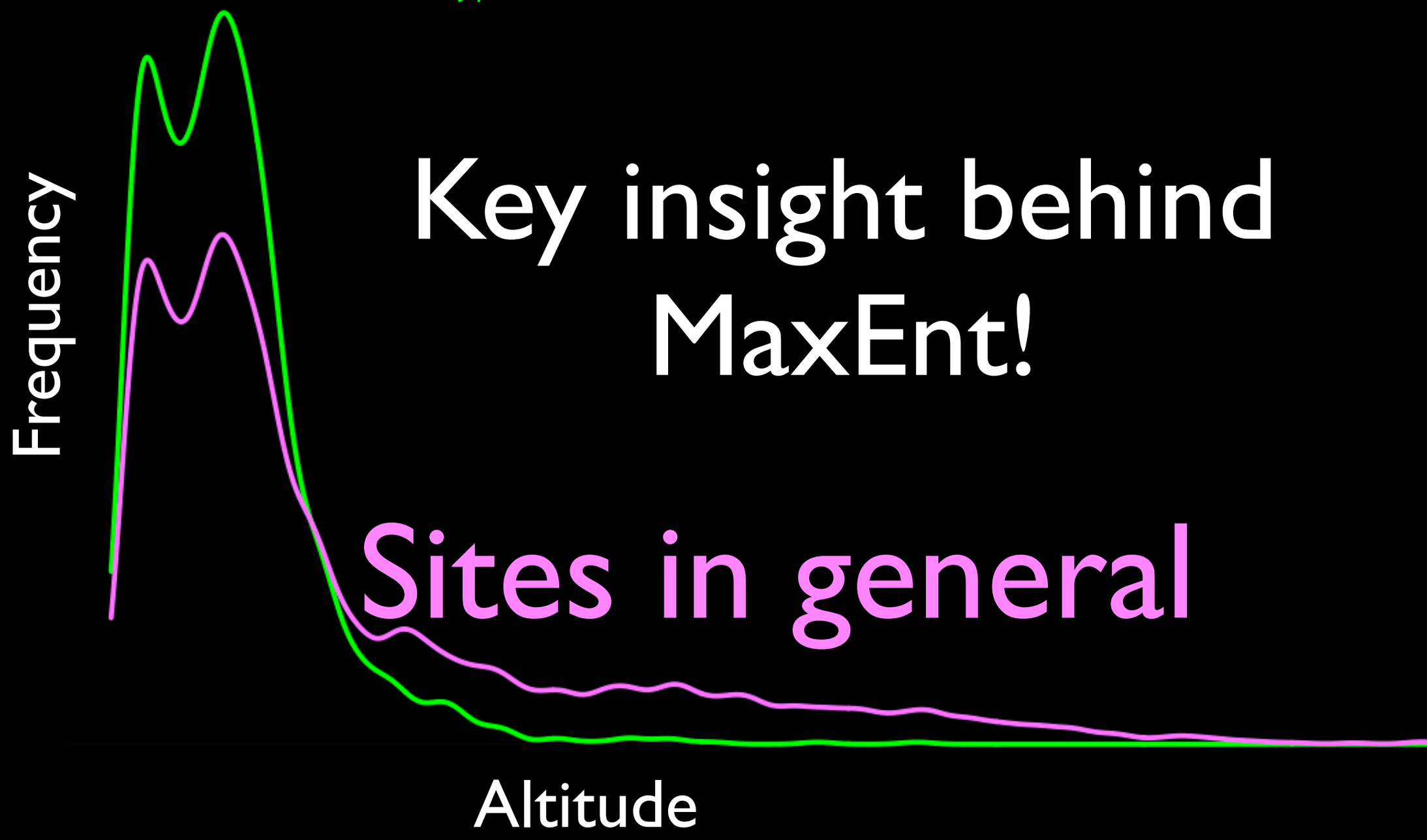
Sites in general

Altitude

Key insight behind MaxEnt!



Key insight behind MaxEnt!



1) Define the model so this null hypothesis,

p(Altitude | observed) = p(Altitude),

is where all parameters = 0

1) Define the model so this null hypothesis,

p(Altitude | observed) = p(Altitude),

is where all parameters = 0

2) Encourage parameters to have small magnitudes

MaxEnt optimizes two conflicting criteria:

MaxEnt optimizes two conflicting criteria:

- p(Altitude | observed) should be high for the altitude values where the species was observed

MaxEnt optimizes two conflicting criteria:

- p(Altitude | observed) should be high for the altitude values where the species was observed
- Maximize the log-likelihood

MaxEnt optimizes two conflicting criteria:

- p(Altitude | observed) should be high for the altitude values where the species was observed
  - Maximize the log-likelihood

- p(Altitude | observed) shouldn't be too different from p(Altitude)

MaxEnt optimizes two conflicting criteria:

- p(Altitude | observed) should be high for the altitude values where the species was observed
  - Maximize the log-likelihood

- p(Altitude | observed) shouldn't be too different from p(Altitude)
  - Minimize the absolute value of the parameters

MaxEnt optimizes two conflicting criteria:

- p(Altitude | observed) should be high for the altitude values where the species was observed
  - Maximize the log-likelihood

- p(Altitude | observed) shouldn't be too different from p(Altitude)
  - Minimize the absolute value of the parameters
  - In Bayesian terms, this is a Laplace prior

“Features”

Model output

Transformed environmental variables

Measured environmental variables

Model output

Transformed environmental variables

Trees, splines, polynomials, etc.

Measured environmental variables

Model output

Weighted sum & final transformation

Transformed environmental variables

Trees, splines, polynomials, etc.

Measured environmental variables

Gibbs distributions

- Proportional to $exp(ax + bx^2 + cxy \ldots)$

Gibbs distributions

- Proportional to $exp(ax + bx^2 + cxy \ldots)$
- Linear stuff in the exponent controls mean

Gibbs distributions

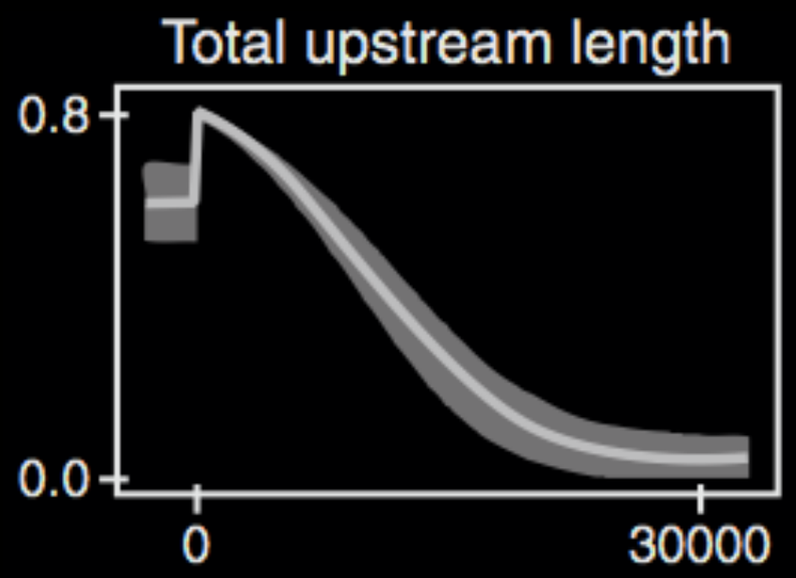
- Proportional to $exp(ax + bx^2 + cxy ....)$
- Linear stuff in the exponent controls mean
- Squared stuff in the exponent controls variance

Gibbs distributions

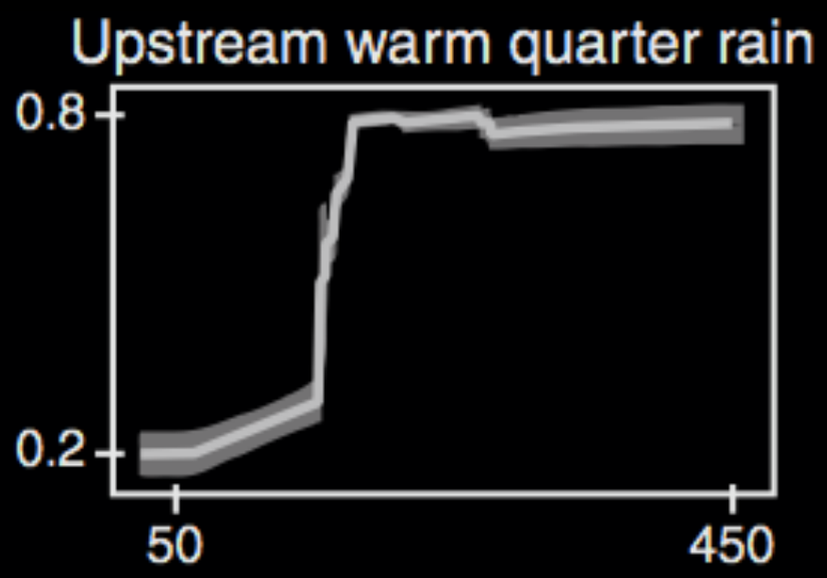
- Proportional to $exp(ax + bx^2 + cxy \ldots)$
- Linear stuff in the exponent controls mean
- Squared stuff in the exponent controls variance
- Products in the exponent control covariance

Gibbs distributions

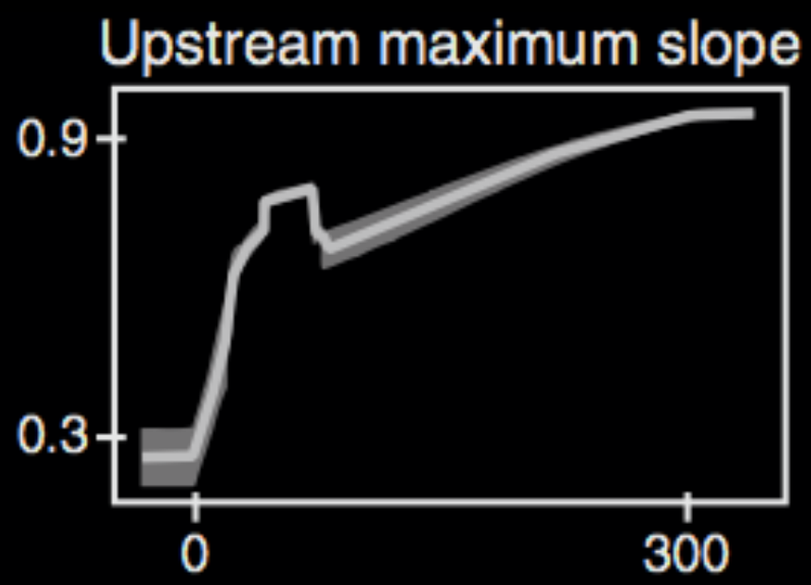
- Proportional to $exp(ax + bx^2 + cxy ....)$
- Linear stuff in the exponent controls mean
- Squared stuff in the exponent controls variance
- Products in the exponent control covariance
- Other stuff can make fancier shapes



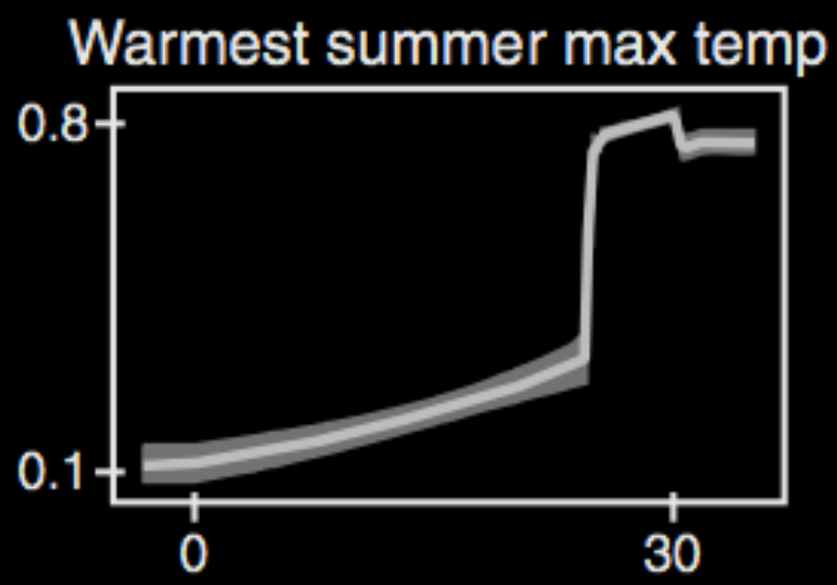
45%



18%



8%



6%

Elith et al. 2010

Interpreting the outputs

Interpreting the outputs

- Raw output: proportional to p(present | environment)
- Don’t know the multiplicative factor

Interpreting the outputs

- Raw output: proportional to p(present | environment)
    - Don’t know the multiplicative factor
- Logistic output: sorta like p(present | environment), *assuming* p(present) is 0.5 at “typical” sites
    - Not proportional to anything in particular, but ranges from 0 to 1

Questions?