INTEGRATING OCCURRENCE DATA AND EXPERT MAPS TO REFINE SPECIES RANGE PREDICTIONS

Cory Merow
Yale University
Combining data sources…

… is our only chance for general predictions

- Small sampling
- Biased samples
- Missing processes

Space and Time
Some nice advances

- Presence Only + Abundance
  - Pagel et al. 2014 MEE
- Presence Only + Presence Absence:
  - Fithian et al. 2014 MEE
  - Dorazio 2014 GEB
- Joint SDMs
  - Harris 2015 MEE
  - Clark et al. 2014 Ecol. App
  - Fithian et al. 2014 MEE
- Presence-Only + Map
  - Merow et al. 2016 GEB
  - Merow et al. 2017 GEB
- Demography + Occurrence:
  - Pagel and Schurr 2012 GEB
Maximum entropy

- Machine learning -> flexible models for exploration
- Stay as close to the null model as possible
- Update null model based on data

Phillips et al. 2006
Maximum entropy

- Machine learning -> flexible models for exploration
- Stay as close to the null model as possible
- Update null model based on data
Maximum entropy

- Machine learning -> flexible models for exploration
- Stay as close to the null model as possible
- Update null model based on data
Maximum entropy

- Machine learning -> flexible models for exploration
- Stay as close to the null model as possible
- Update null model based on data
Maximum entropy

Spatially explicit prior information

Uniform Prior

Best guess?

Data

Prediction

Expert Dispersal
Native range
Related species
Priors based on other models

Uniform Prior

Some other model

Data

Data

Prediction

Prediction

relative occurrence rate
What is an offset?

• When sampling abundance (which entails a Poisson model), twice as much sampling effort should yield twice as many individuals

• Sampling effort = Offset

• A way to say ‘more samples are expected here’
Expert Maps

How can we translate this to a prior/offset?

How much more likely is presence inside vs outside?
By combining their powers….

- High resolution
- Links to Environment
- Suitability
- Large data bases
  BUT...
- Biased sampling

- Field experience
- Low bias
- Range edges
- Non-climate processes
  BUT...
- Coarse resolution
- No explicit environment
- False presence
- Not updated
- Not reproducible
- Binary
Other useful Priors

- **Expert maps**
- **Sampling Bias**
- **Native range info**
  - Realized vs
  - Potential Distributions
- **Phylogenetic constraints**
- **Combining different types of occurrence data**

Merow et al. GEB 2016
Questions?

Thanks:

Matt Aiello-Lammens
Jenica Allen
Walter Jetz
John Silander
Adam Wilson