Uncertainty in Bioclimatic Models: Examples from the Western Hemisphere

Josh Lawler
TNC David H. Smith Conservation Fellow
Oregon State University
Corvallis, OR
lawler.joshua@epa.gov

Collaborators

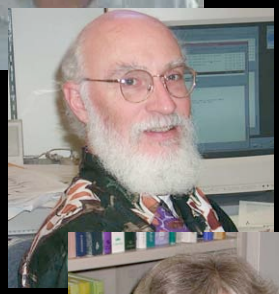
Andy Blaustein, Oregon State University

Peter Kareiva, The Nature Conservancy

Ron Neilson, US Forest Service

Sarah Shafer, US Geological Survey

Denis White, US Environmental Protection Agency



How will climate change affect species distributions?

How will range shifts affect protected lands?

How can we address climate change in the conservation planning process?

Overview

1. Background
2. Comparison of methods
3. Investigation of errors
4. Conclusions

Overview

1. Background
2. Comparison of methods
3. Investigation of errors
4. Conclusions

Background: Bioclimatic Models

Climate-envelope or niche models

Current ranges modeled as a functions of current climate

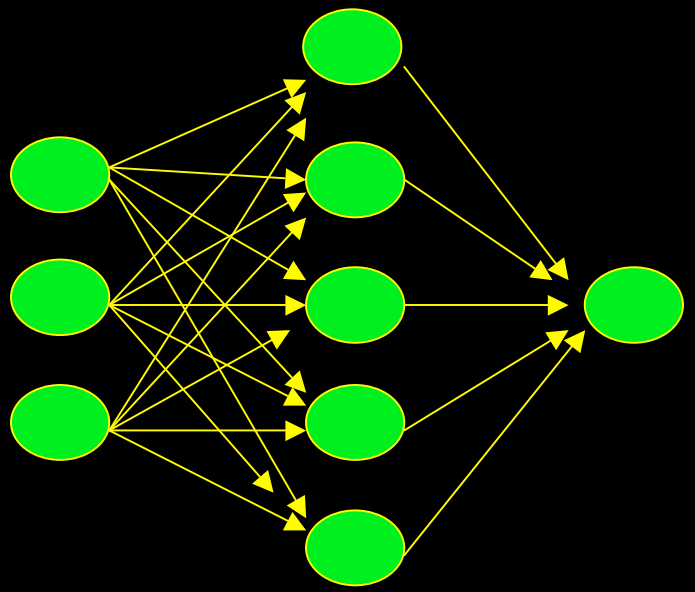
Predicted future (or past) climate used to predict future (or past) ranges

Background: Bioclimatic Models

- Statistical Models

$$presence \sim \beta_0 + \beta_1 julytemp + \beta_2 annualprecip + \beta_3 coniferfor + \varepsilon$$

- Machine-learning based Models



Background: Bioclimatic Models

Limitations

- Modeling approaches generally can not account for many processes including:
  - -Dispersal
  - -Biotic interactions
  - -Adaptation
- No one modeling approach works the best for all species.

Overview

1. Background
2. Comparison of methods
3. Investigation of errors
4. Conclusions

What are the implications of using different modeling approaches for predicted range shifts?

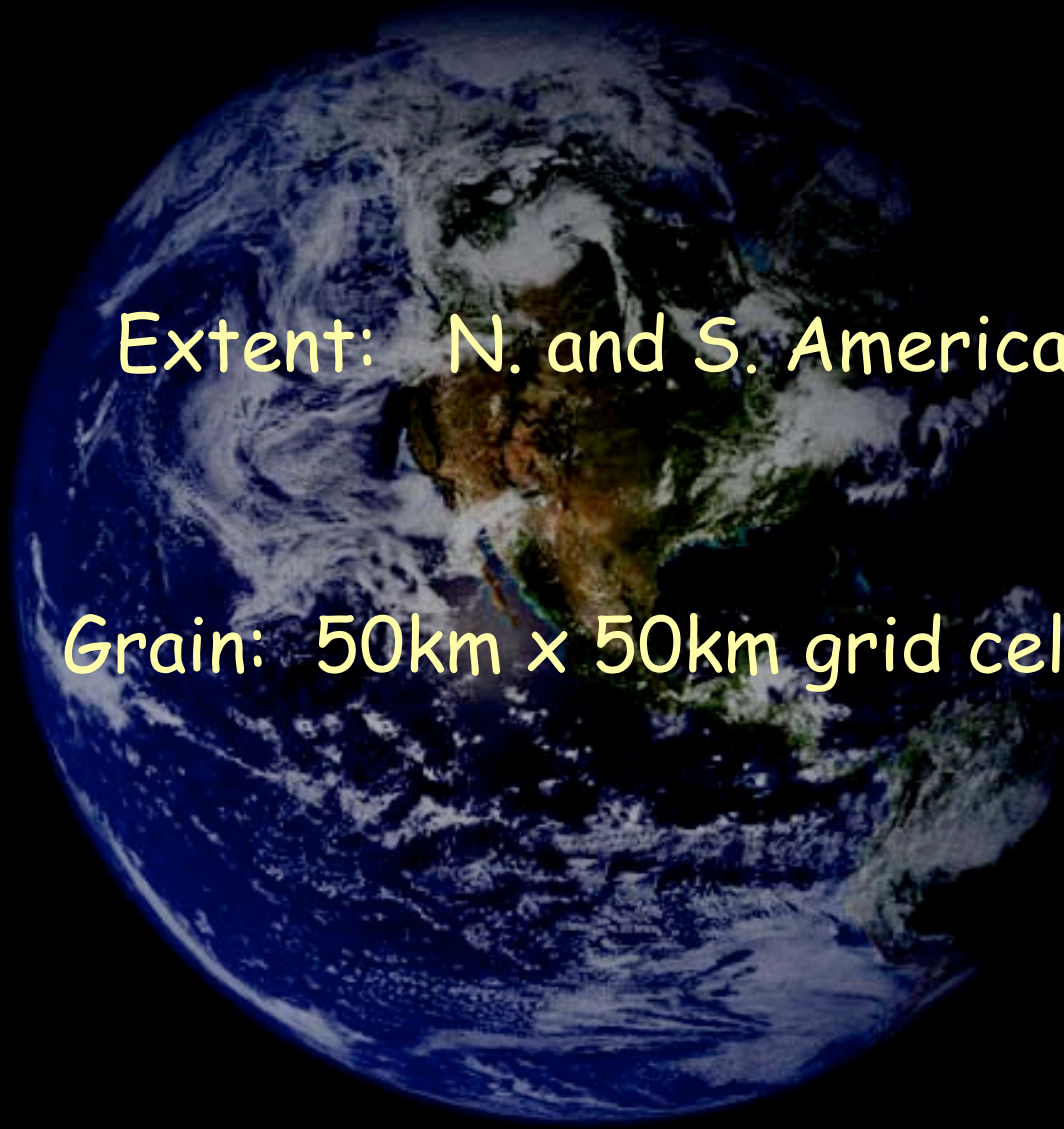
Does one modeling approach consistently outperform the others?

Objective

Compare the predictions and the accuracy of six alternative modeling approaches:

GLM, TREE, GAM, ANN, RandomForest, & GARP

Scale



Extent: N. and S. America

Grain: 50km x 50km grid cells

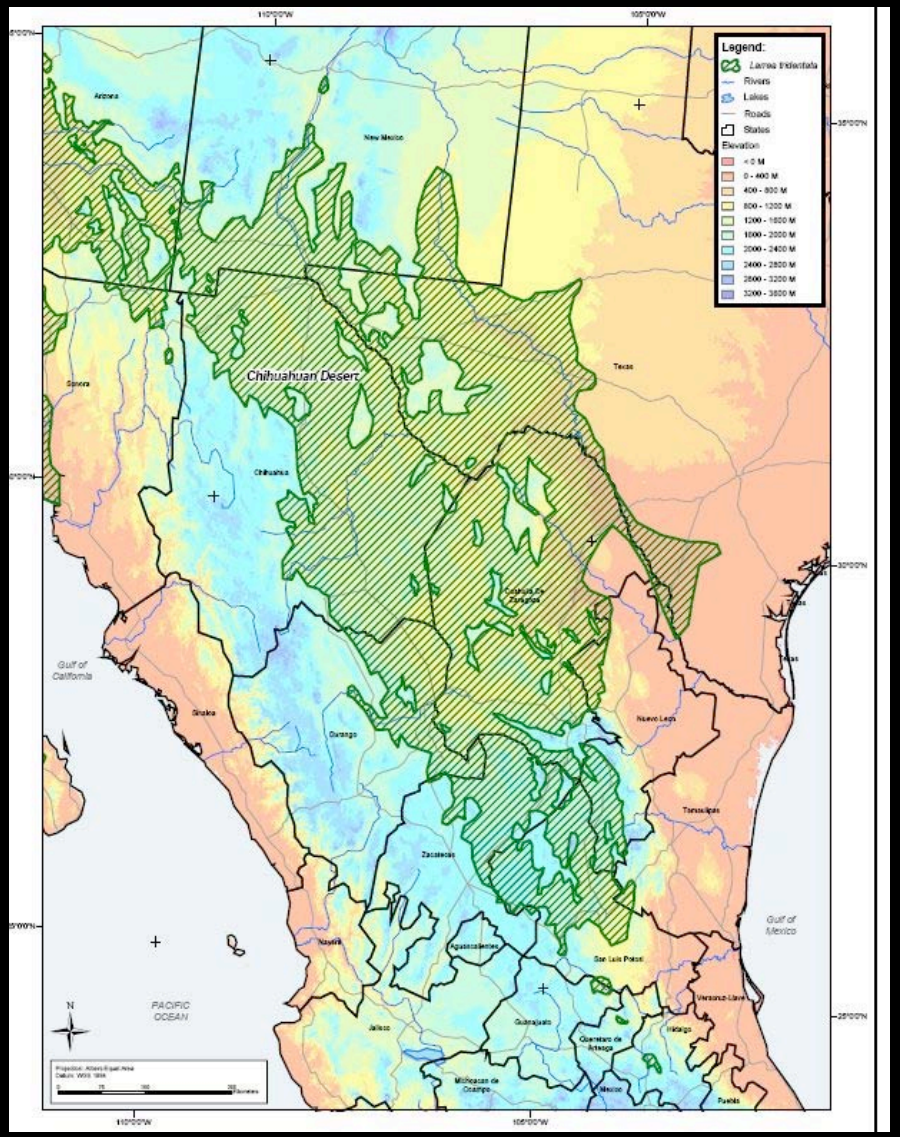
Data

- Current geographic ranges
- Current climate
- Predicted future climate
- Current land cover
- Predicted future land cover

Current Geographic Ranges

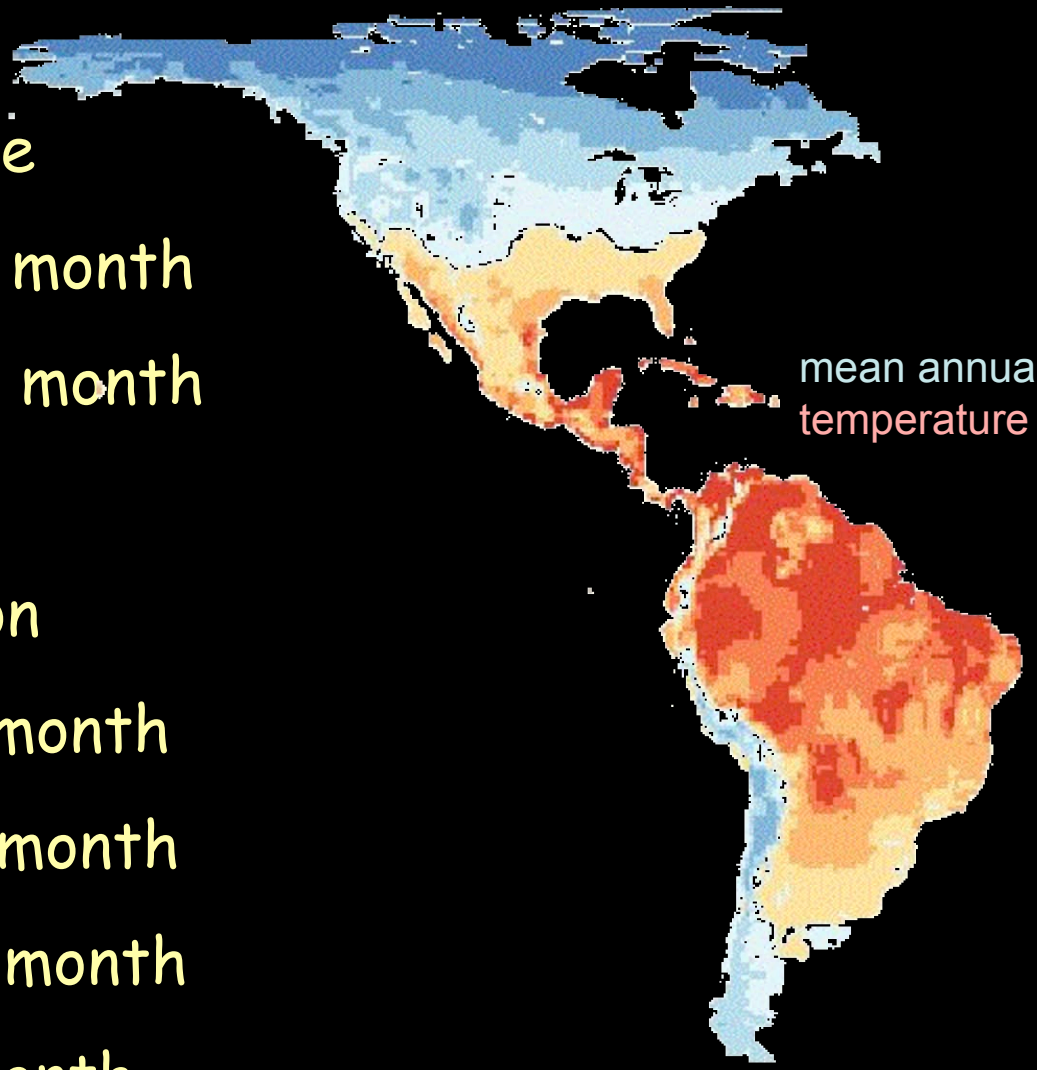
- Digitized range maps
- Presence/absence
- 100 randomly selected mammals
- Ranges > 50 cells

(Patterson et al. 2003)



Current Climate Variables

- Mean annual temperature
- Temperature of coldest month
- Temperature of hottest month
- Degree days > 5°C
- Mean annual precipitation
- Precipitation in coldest month
- Precipitation in hottest month
- Precipitation in wettest month
- Precipitation in driest month



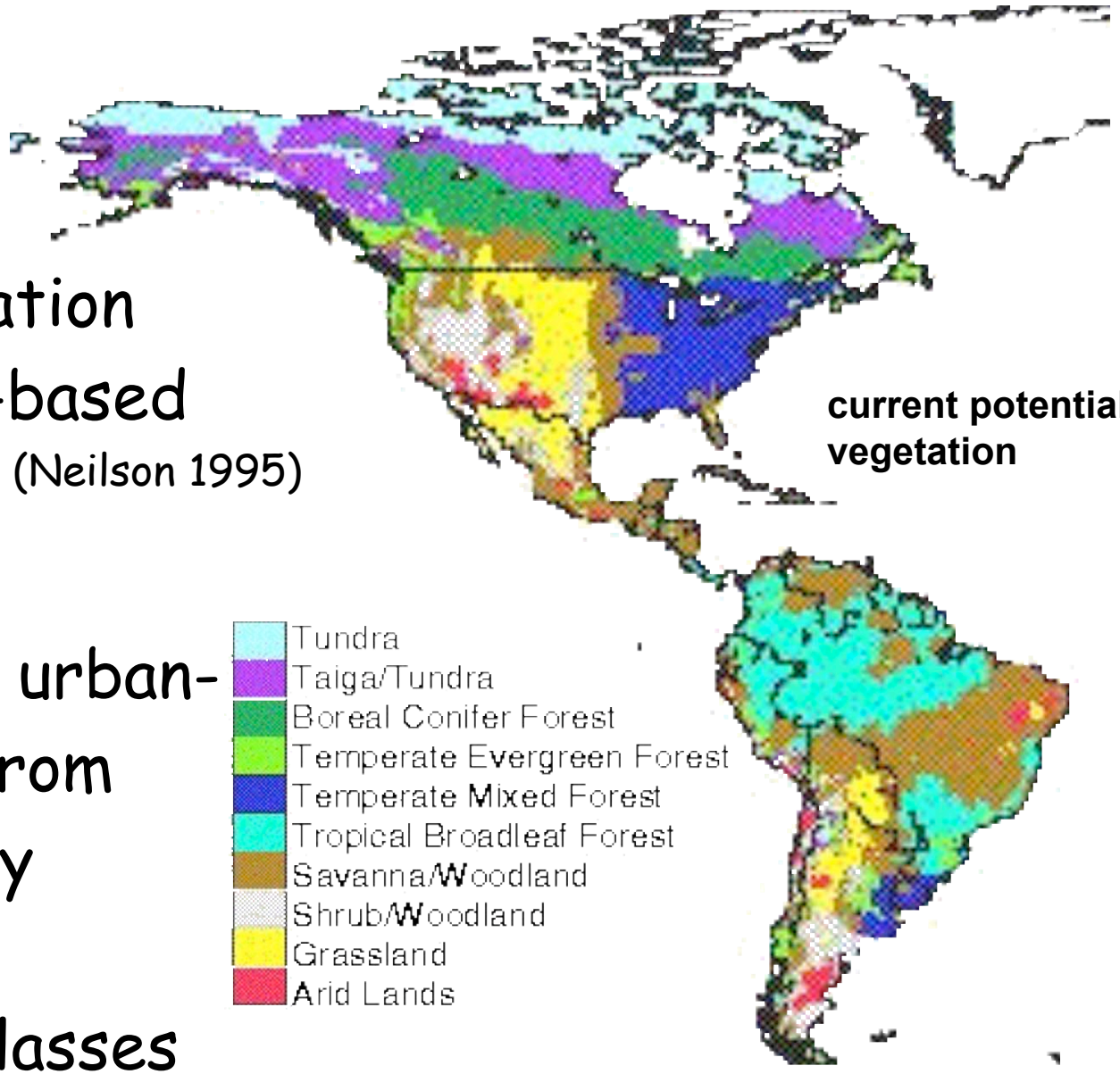
(Leemans and Cramer 1991)

Current Land-Cover Data

Potential vegetation from a process-based model (MAPSS) (Neilson 1995)

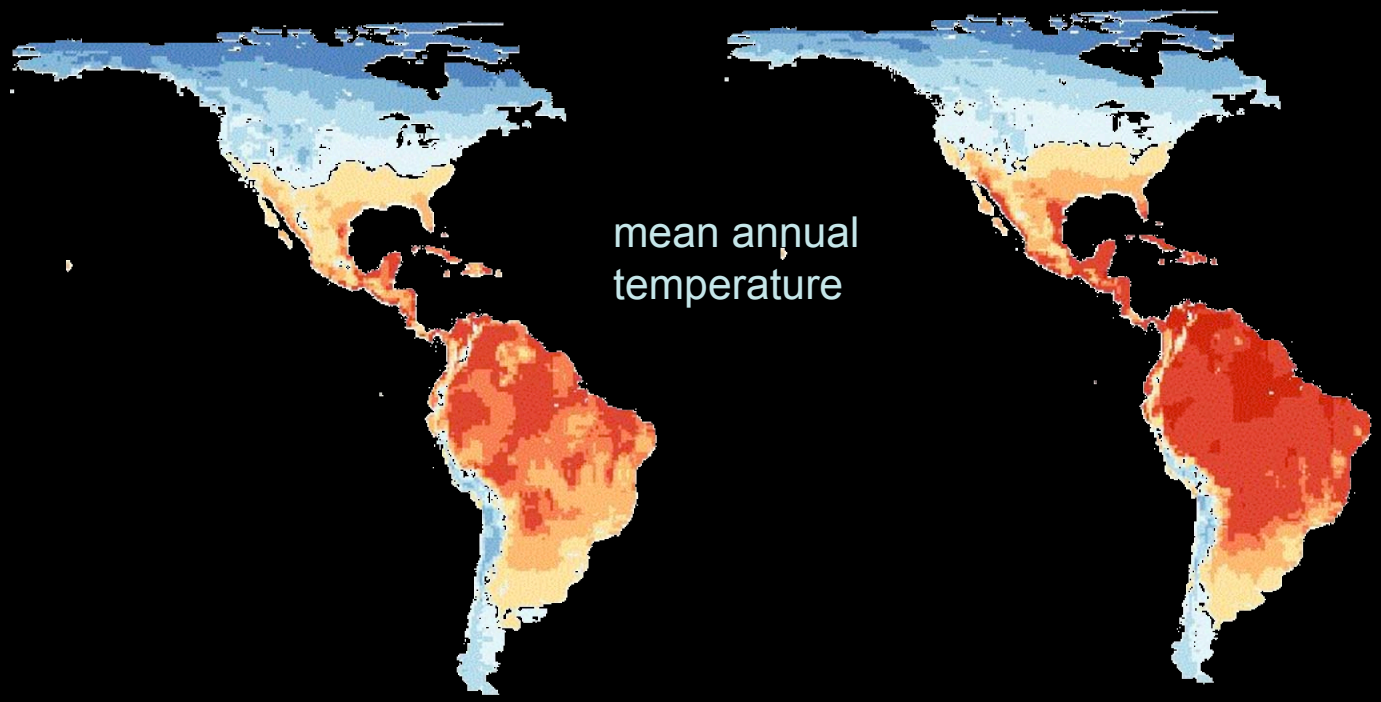
Agriculture and urban-suburban land from satellite imagery

50 land-cover classes



Predicted Future Climate

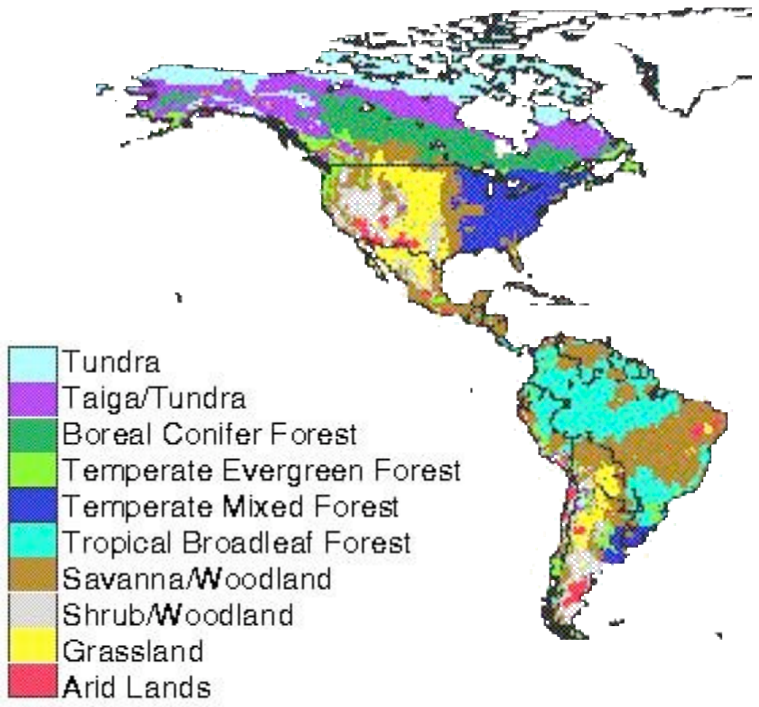
- HADCM2
- IS62a
- 2061-2090



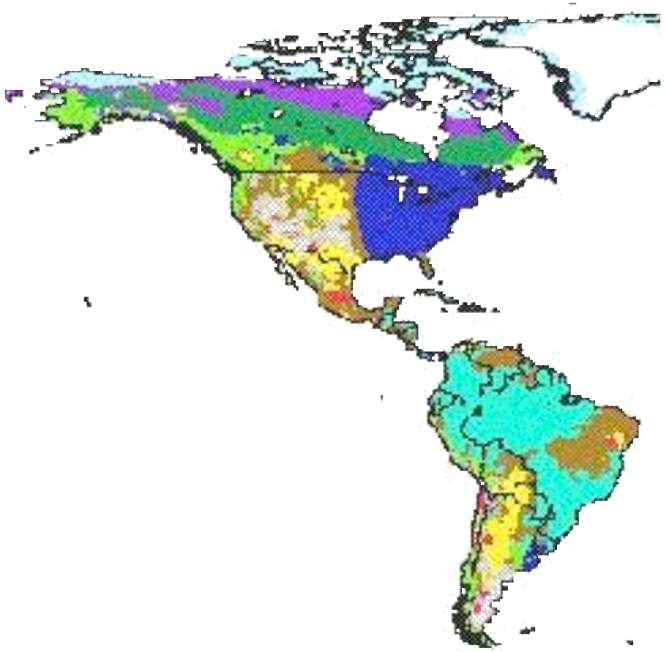
mean annual temperature

Future Land Cover Data

- Predicted potential vegetation (MAPSS) (Neilson 1995)



- Agriculture and urban-suburban land held constant



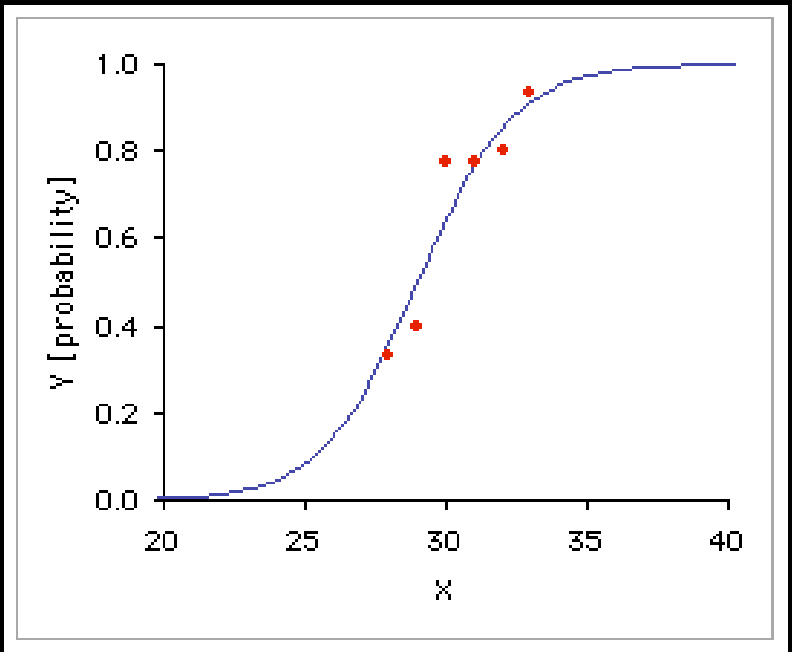
Comparison of six approaches

- Generalized Linear Models (GLM)
- Classification Trees (TREE)
- Generalized Additive Models (GAM)
- Random Forests (RF)
- Artificial Neural Networks (ANN)
- Genetic Algorithms for Rule-set Prediction (GARP)

Generalized Linear Models

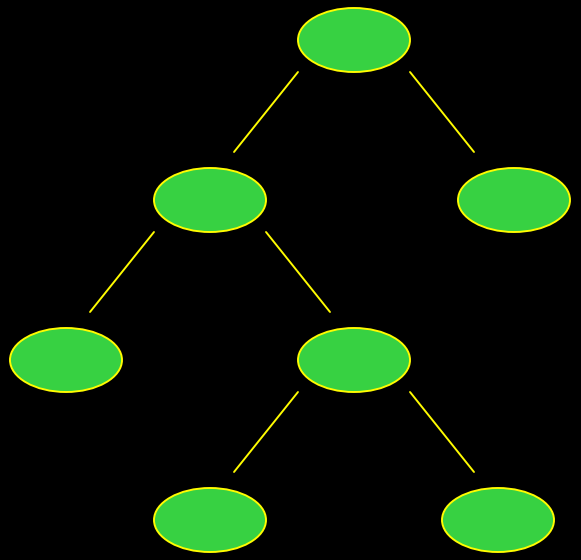
- Logistic regression models with linear and 2nd order polynomial terms

- Stepwise model selection (forward and backward selection with AIC)



Classification Tree Models

- Binary recursive partitioning
- Cross-validated pruning using a 1-SE selection method.



Breiman et al. 1984

Generalized Additive Models

- Penalized regression splines

- Variable selection based on smoothness penalties and shrinkage parameter

Wood 2000

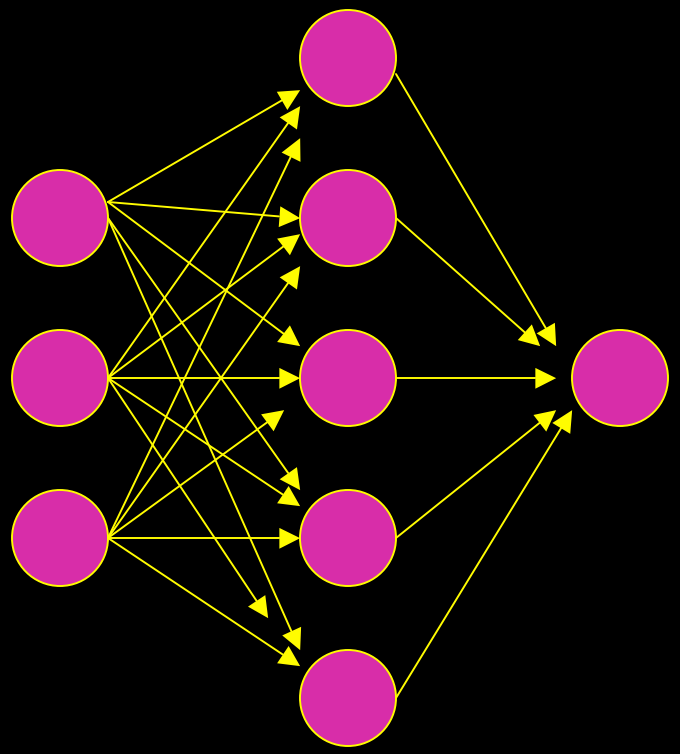
Random Forest Classifiers

- Model averaging
- Classification tree models
- Random subsets of observations and predictor variables
- 500-tree forests

Breiman 2001

Artificial Neural Networks

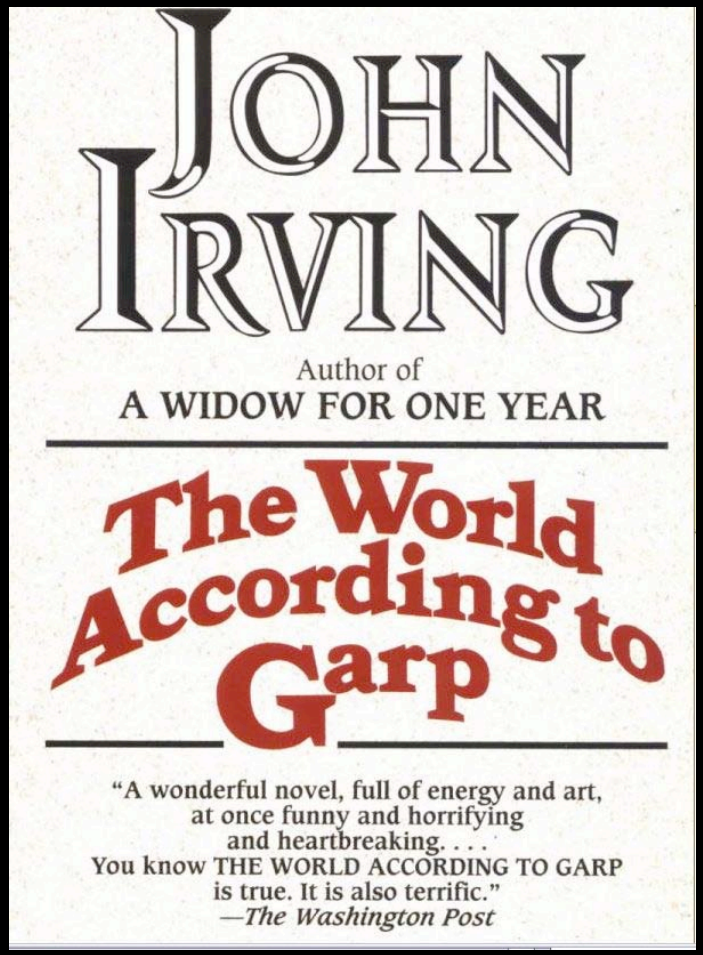
- Feed-forward
- Single hidden layer of 8 nodes
- Logistic activation function
- Averaged the predictions of 10 networks



Ripley 1993, 1996

Genetic Algorithm for Rule-Set Prediction

- Rules:
  - atomic
  - BIOCLIM
  - range
  - logit
- Genetic algorithm to minimize miss-classification
- 500 GARP models
- Averaged the predictions of the 10 best models



Stockwell and Noble 1992

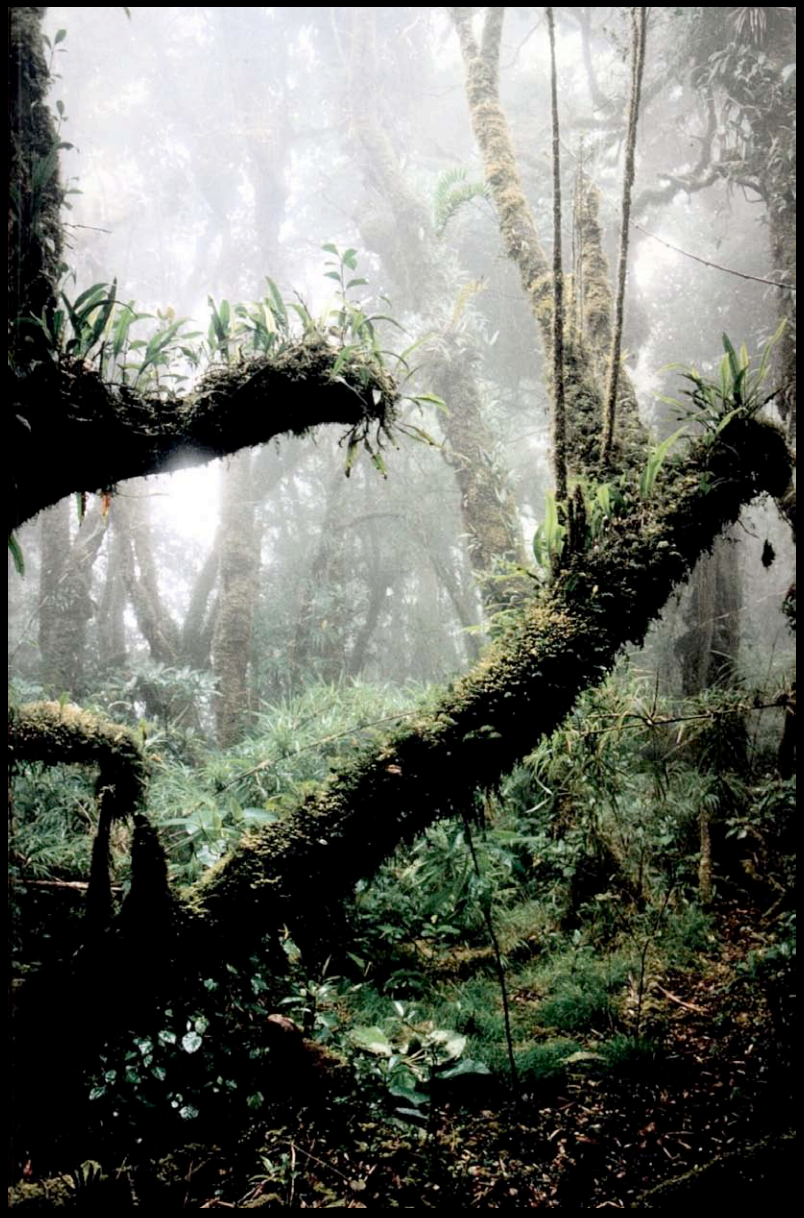
Model Building and Testing

Training data sets:

- 80% of presences
- 80% of absences

Testing data sets:

- 20% of presences
- 20% of absences



Results:
Future Predictions

Pole-ward range shift

(random forest model)

Absent

Stable

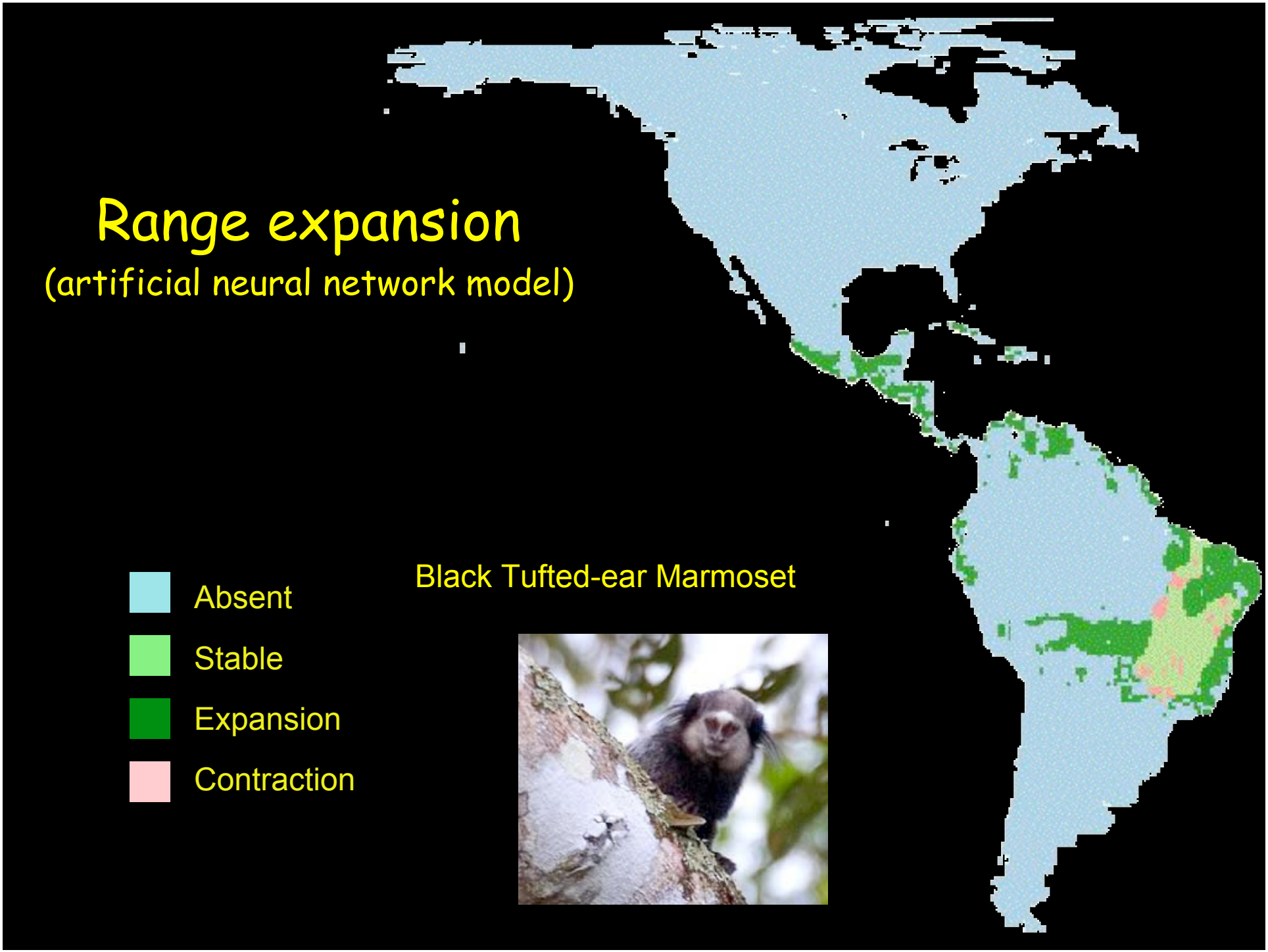
Expansion

Contraction

Fisher, *Martes pennanti*

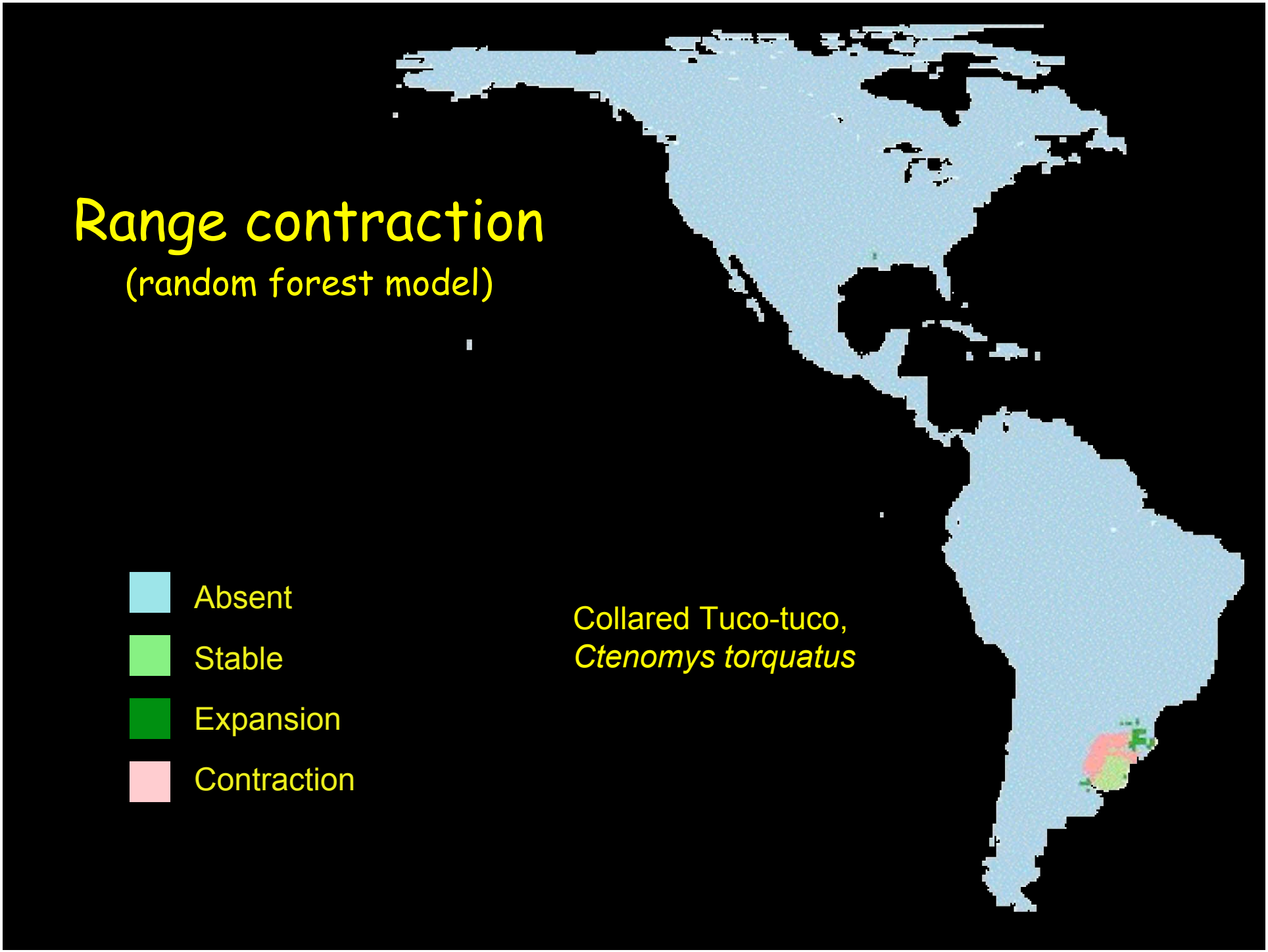
Range expansion

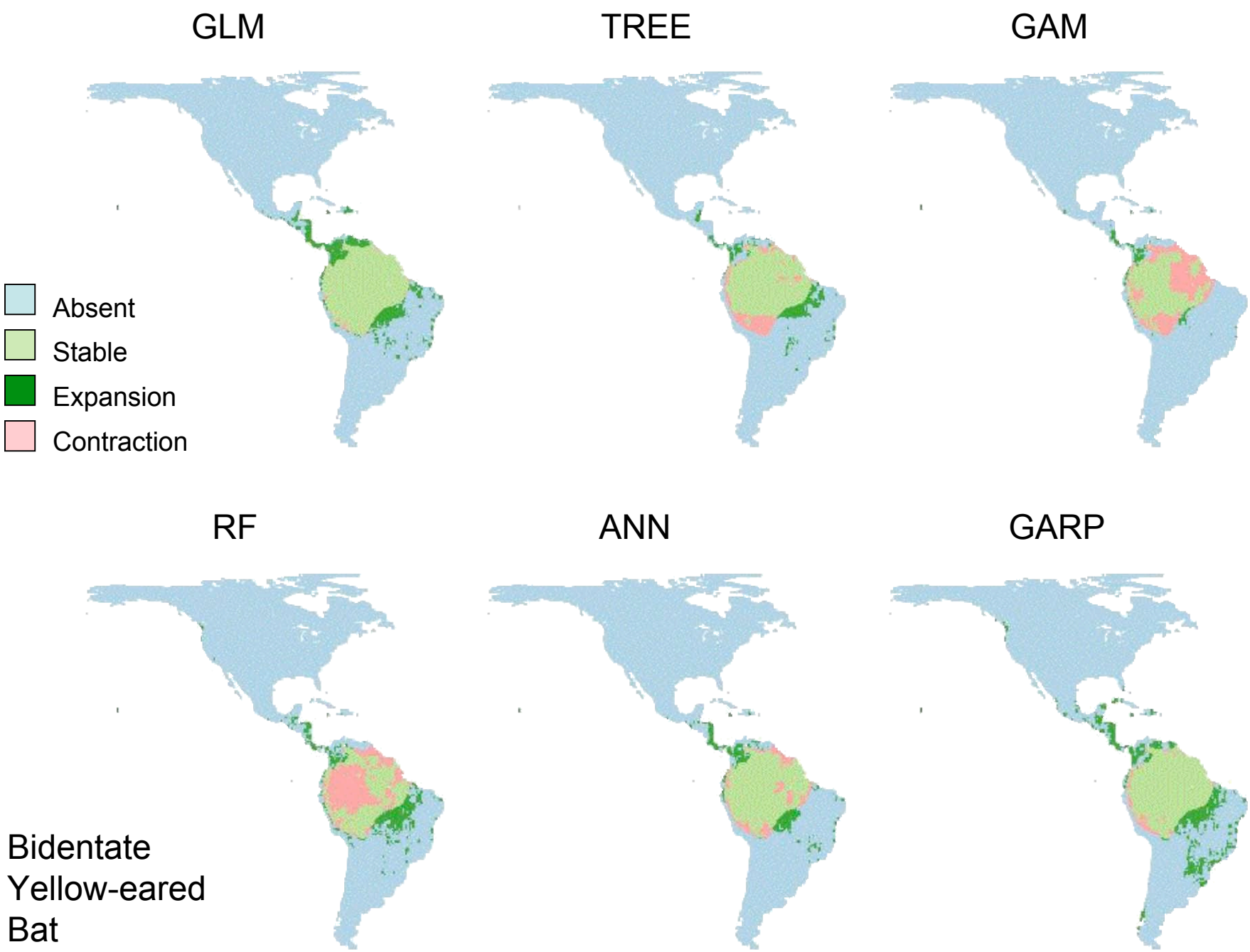
(artificial neural network model)

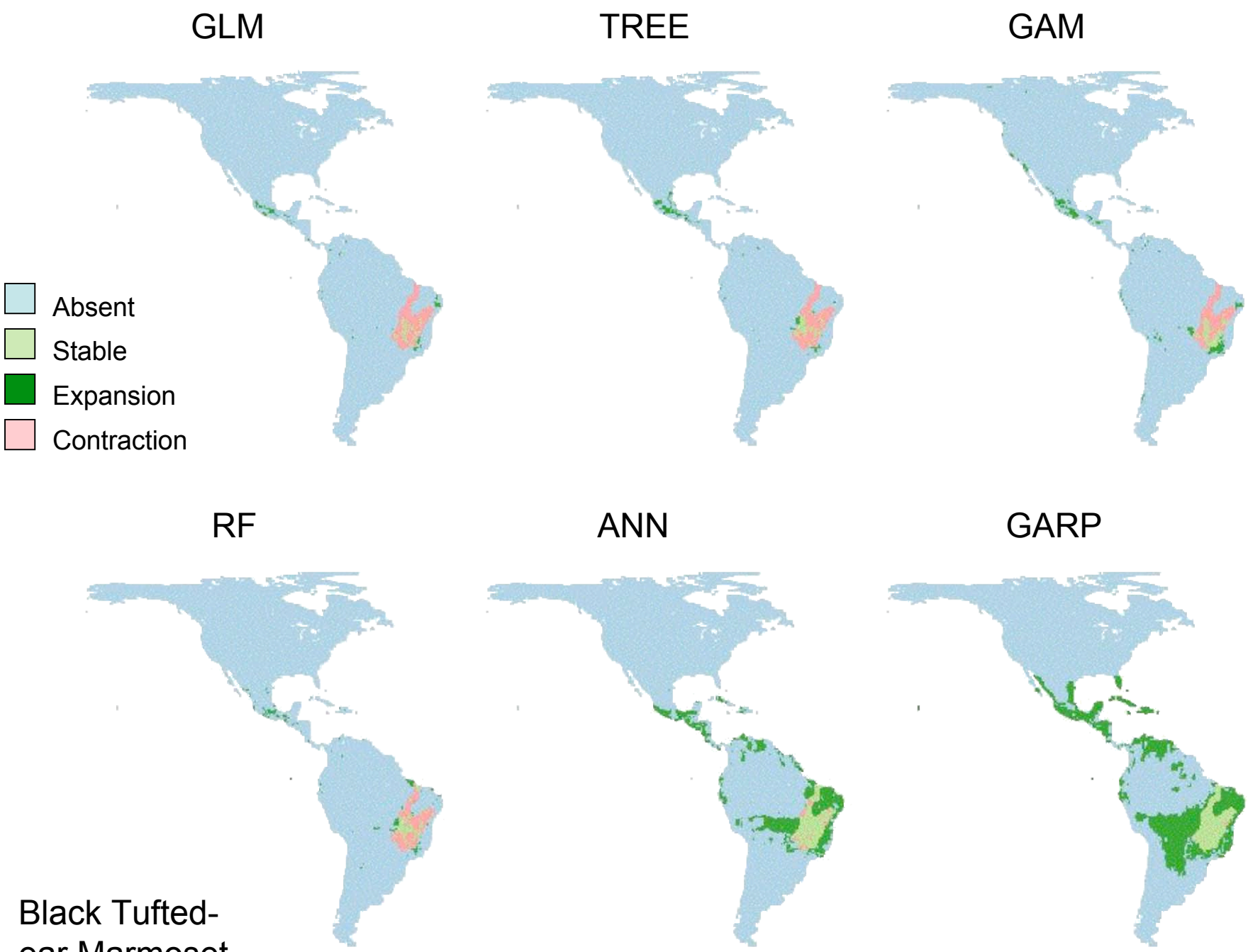


Range contraction

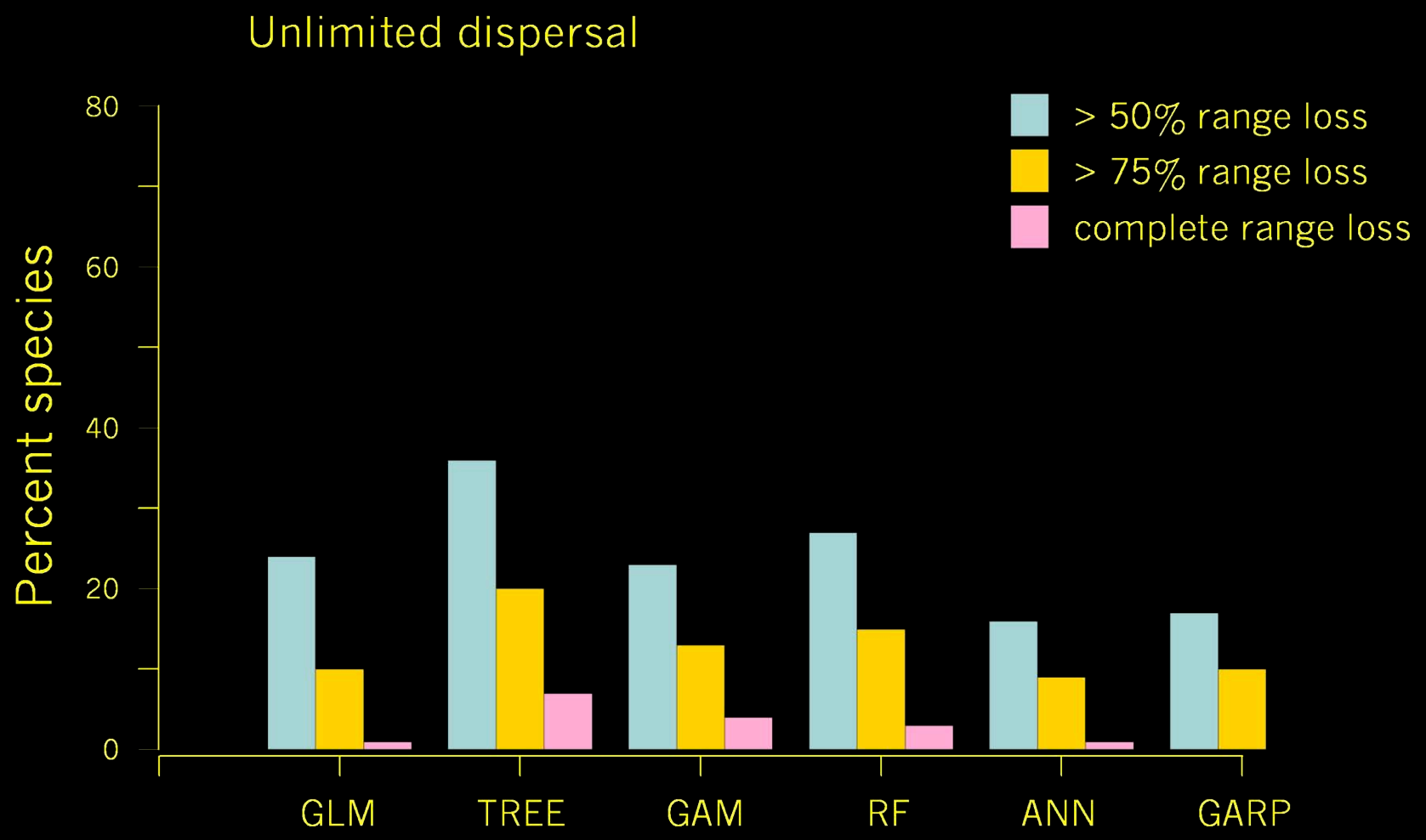
(random forest model)



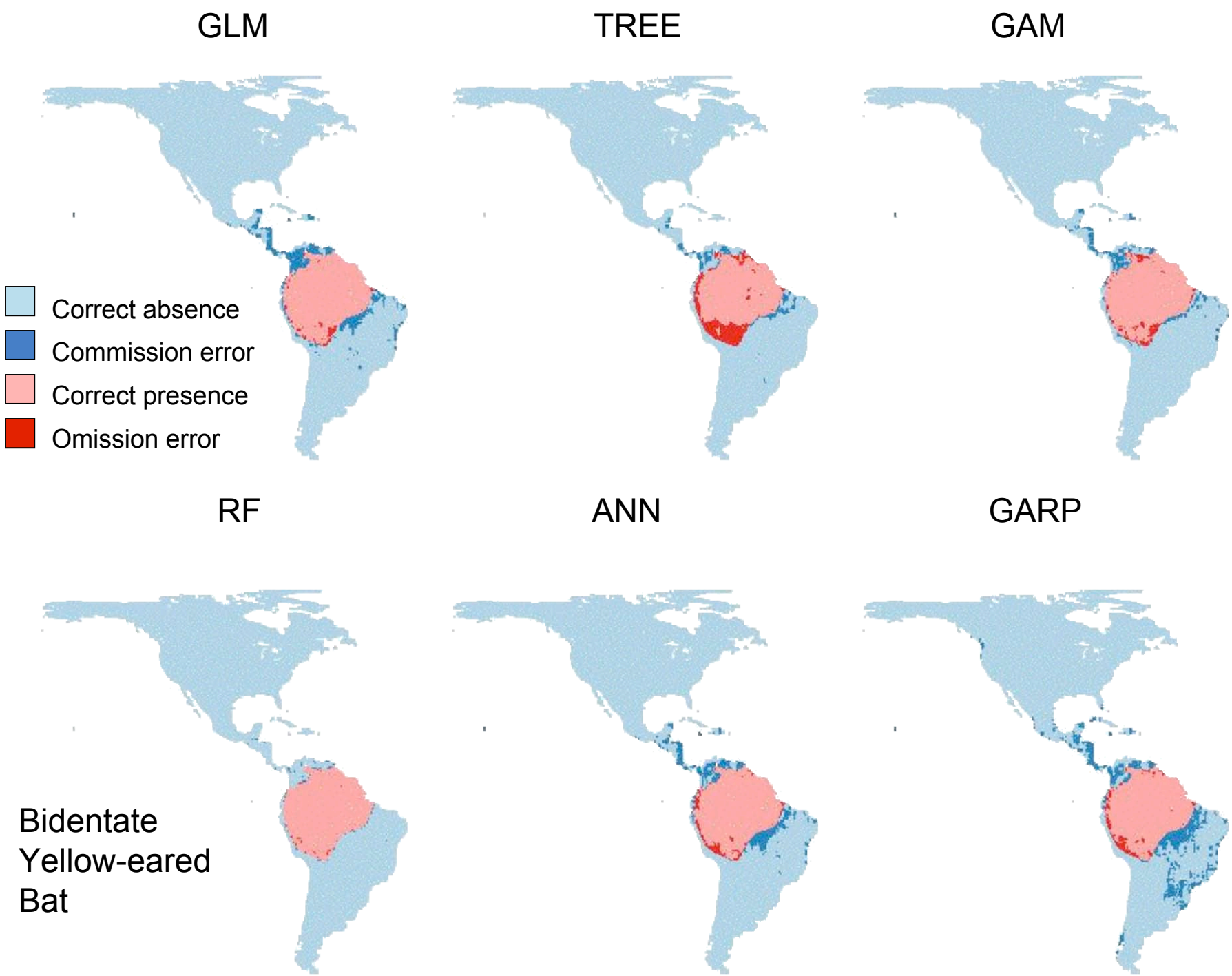

GLM
TREE
GAM
Absent
Stable
Expansion
Contraction
RF
ANN
GARP
Bidentate
Yellow-eared
Bat

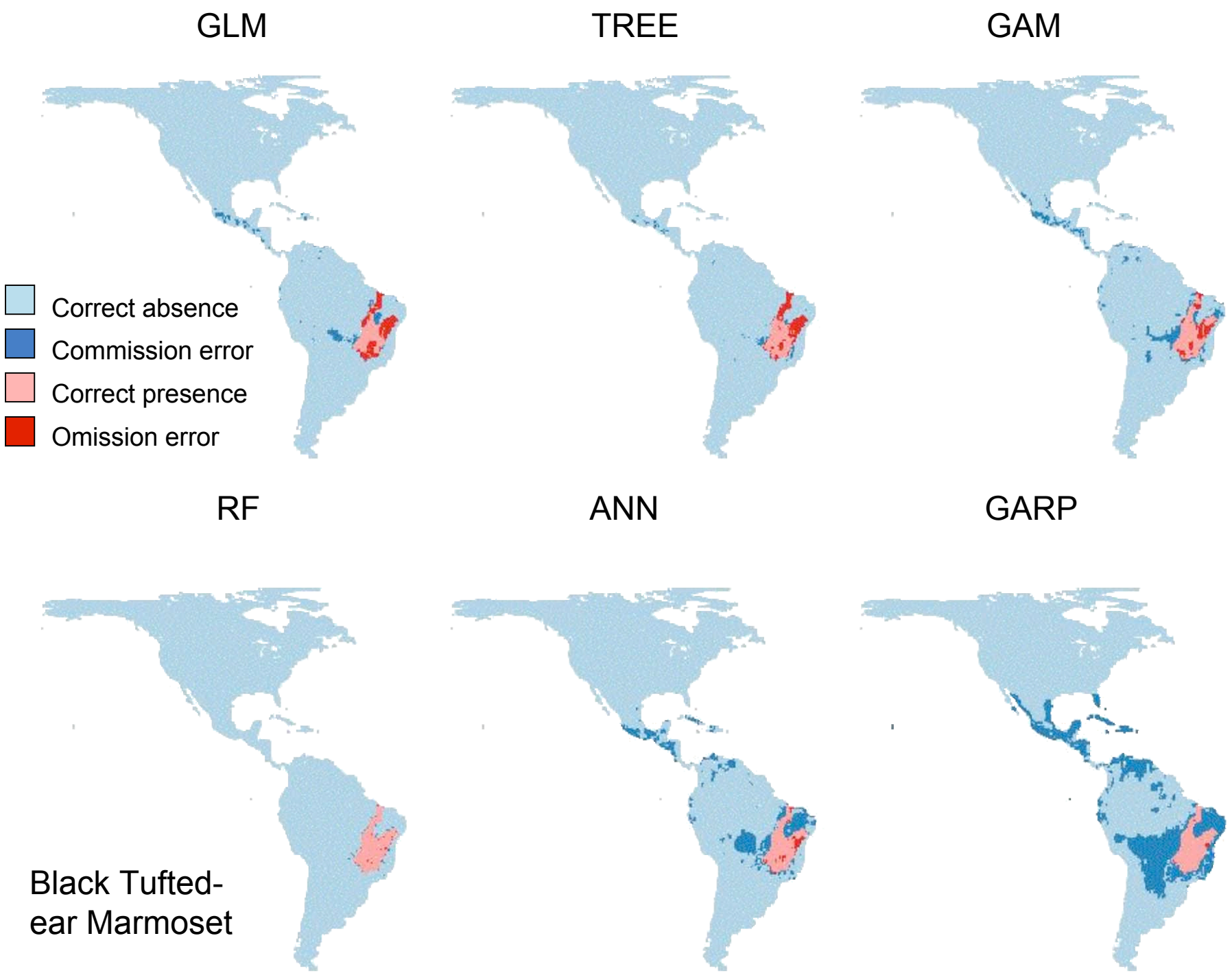

GLM
TREE
GAM
Absent
Stable
Expansion
Contraction
RF
ANN
GARP
Black Tufted-ear Marmoset

Range Contractions



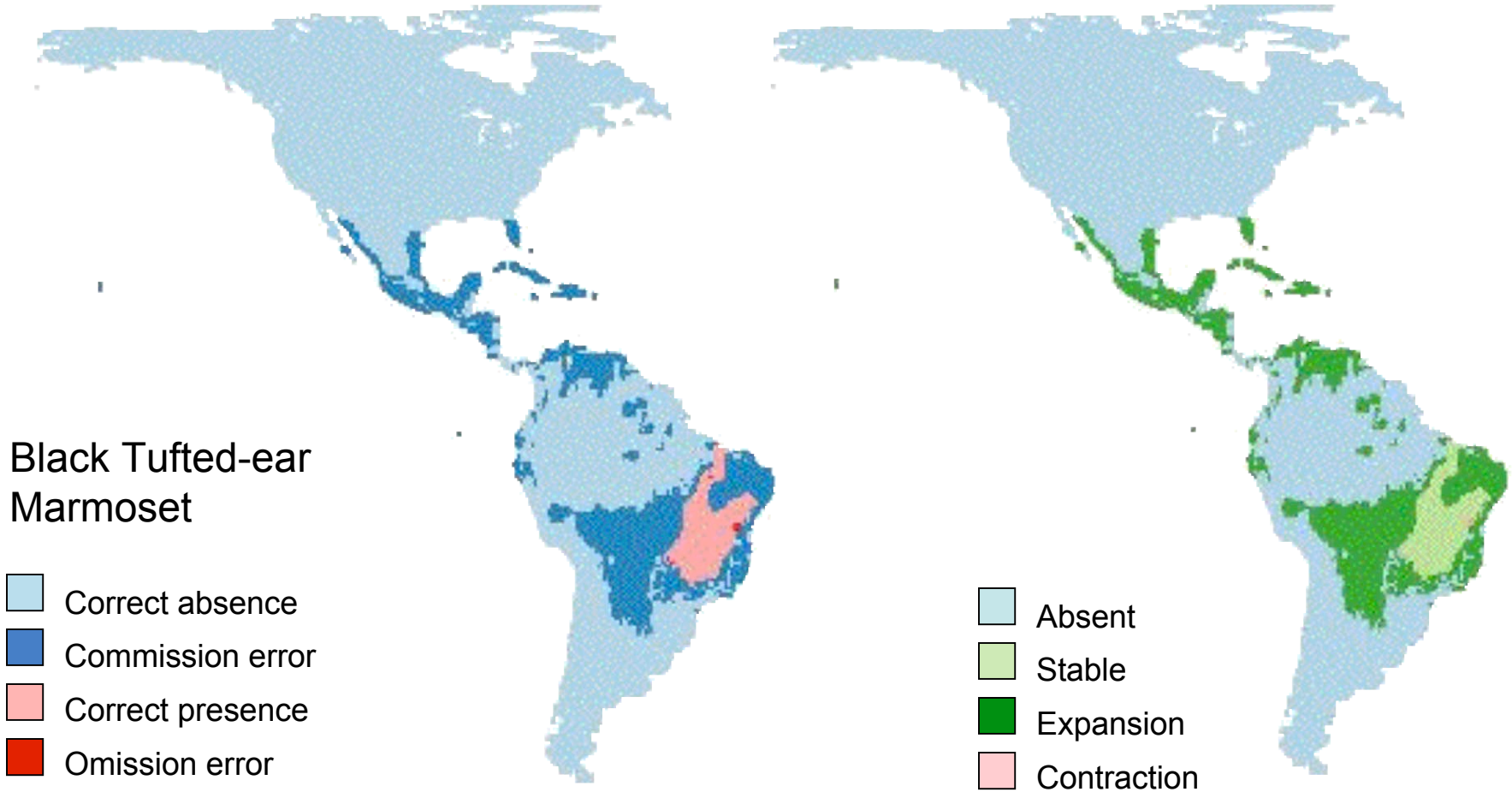
Results:
model accuracy


GLM
TREE
GAM
Correct absence
Commission error
Correct presence
Omission error
RF
ANN
GARP
Bidentate
Yellow-eared
Bat


GLM
TREE
GAM
Correct absence
Commission error
Correct presence
Omission error
RF
ANN
GARP
Black Tufted-ear Marmoset

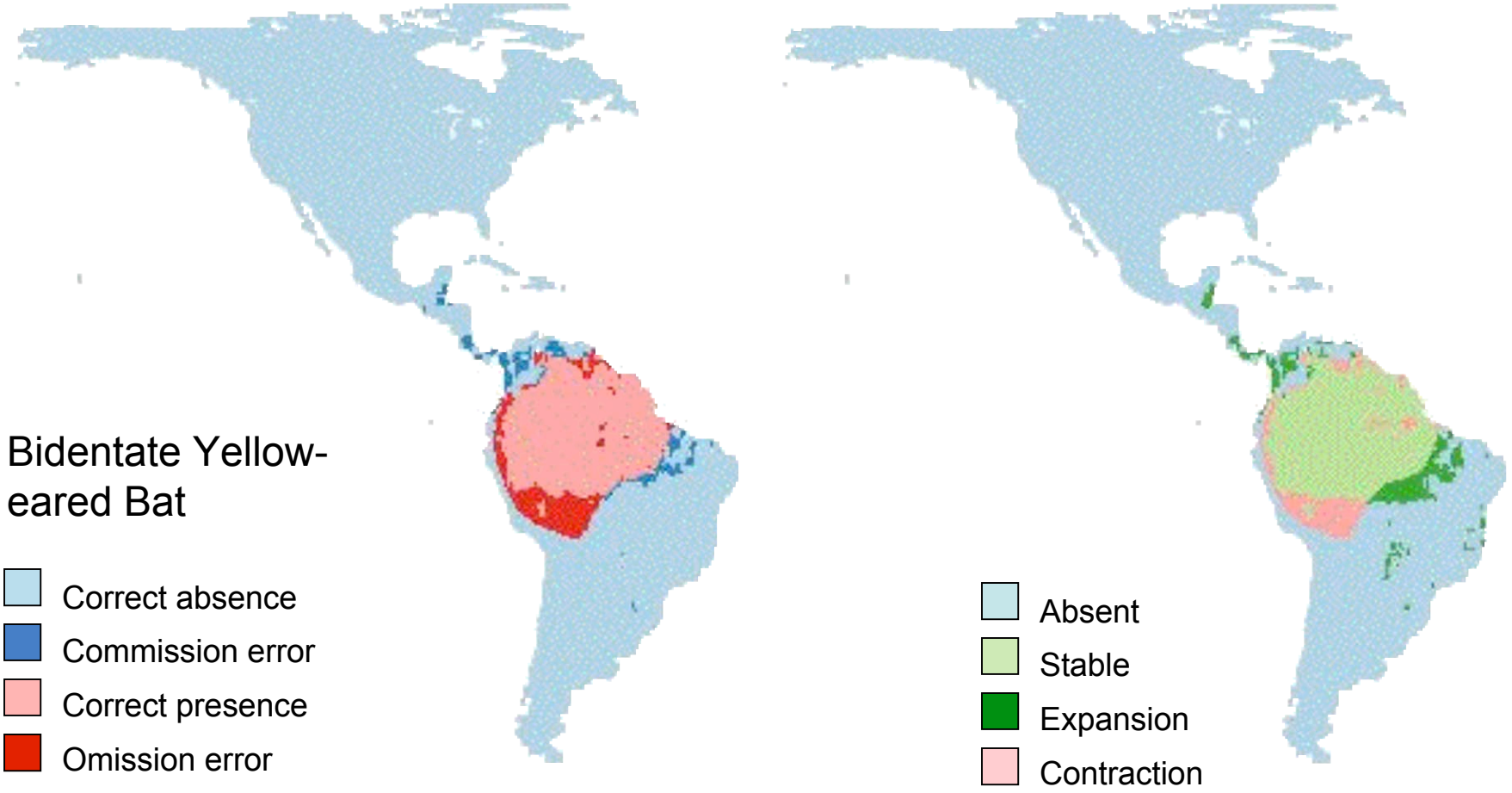
Errors and Predictions

GARP MODELS

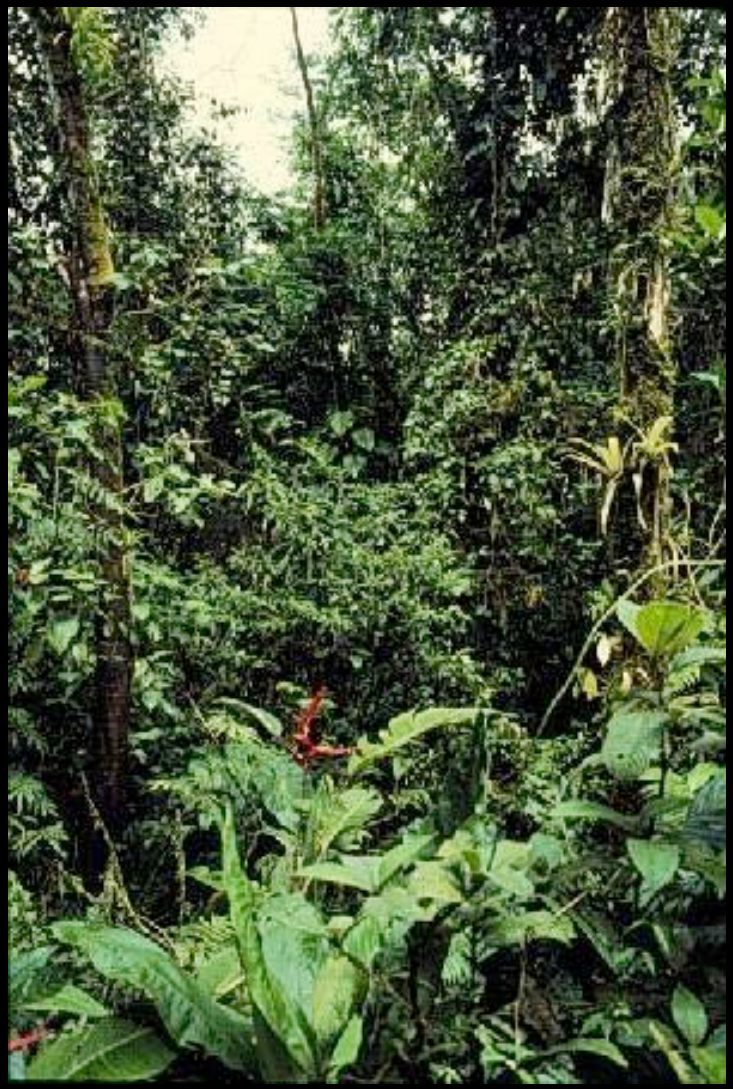


Errors and Predictions

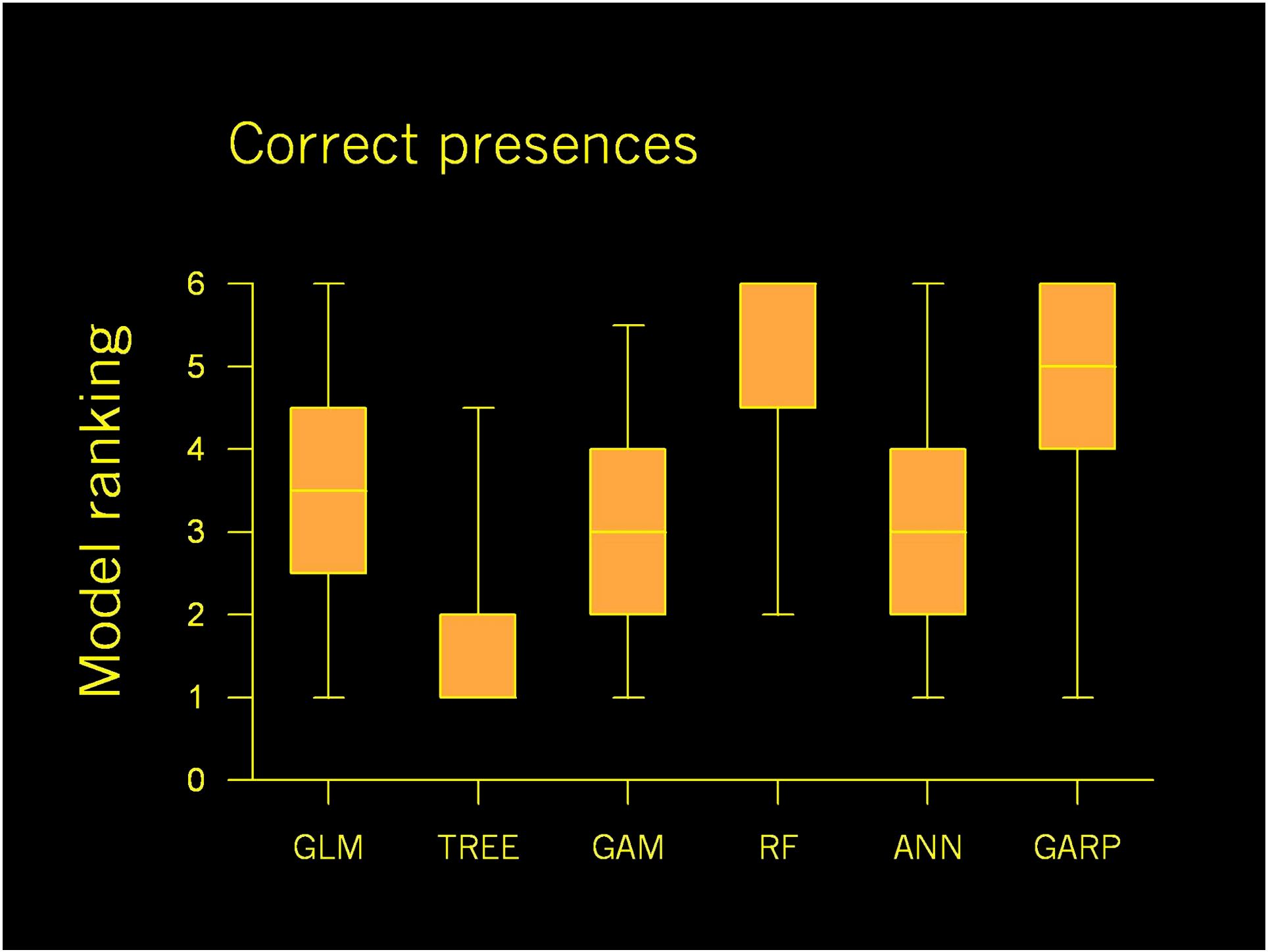
TREE MODELS

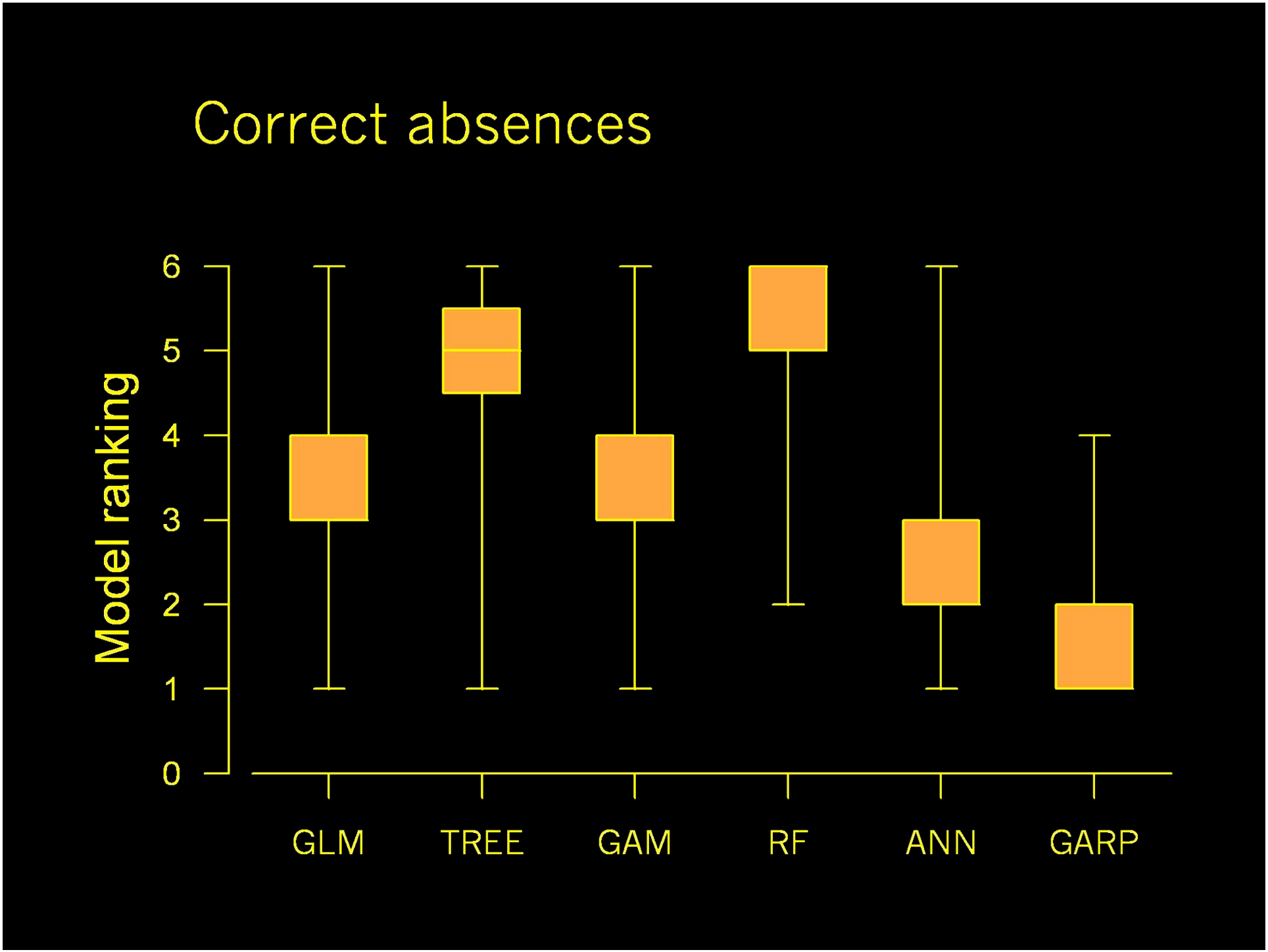


Model Comparisons



- % correct presences
- % correct absences
- AUC: Area under the receiver-operating characteristic (ROC) curve
- Kappa


Correct presences
Model ranking
6
5
4
3
2
1
0
GLM
TREE
GAM
RF
ANN
GARP


Correct absences
Model ranking
6
5
4
3
2
1
0
GLM
TREE
GAM
RF
ANN
GARP

Overview

1. Background
2. Comparison of methods
3. Investigation of errors
4. Conclusions

Do model errors vary systematically by geography or species?

What is the spatial distribution of model errors?

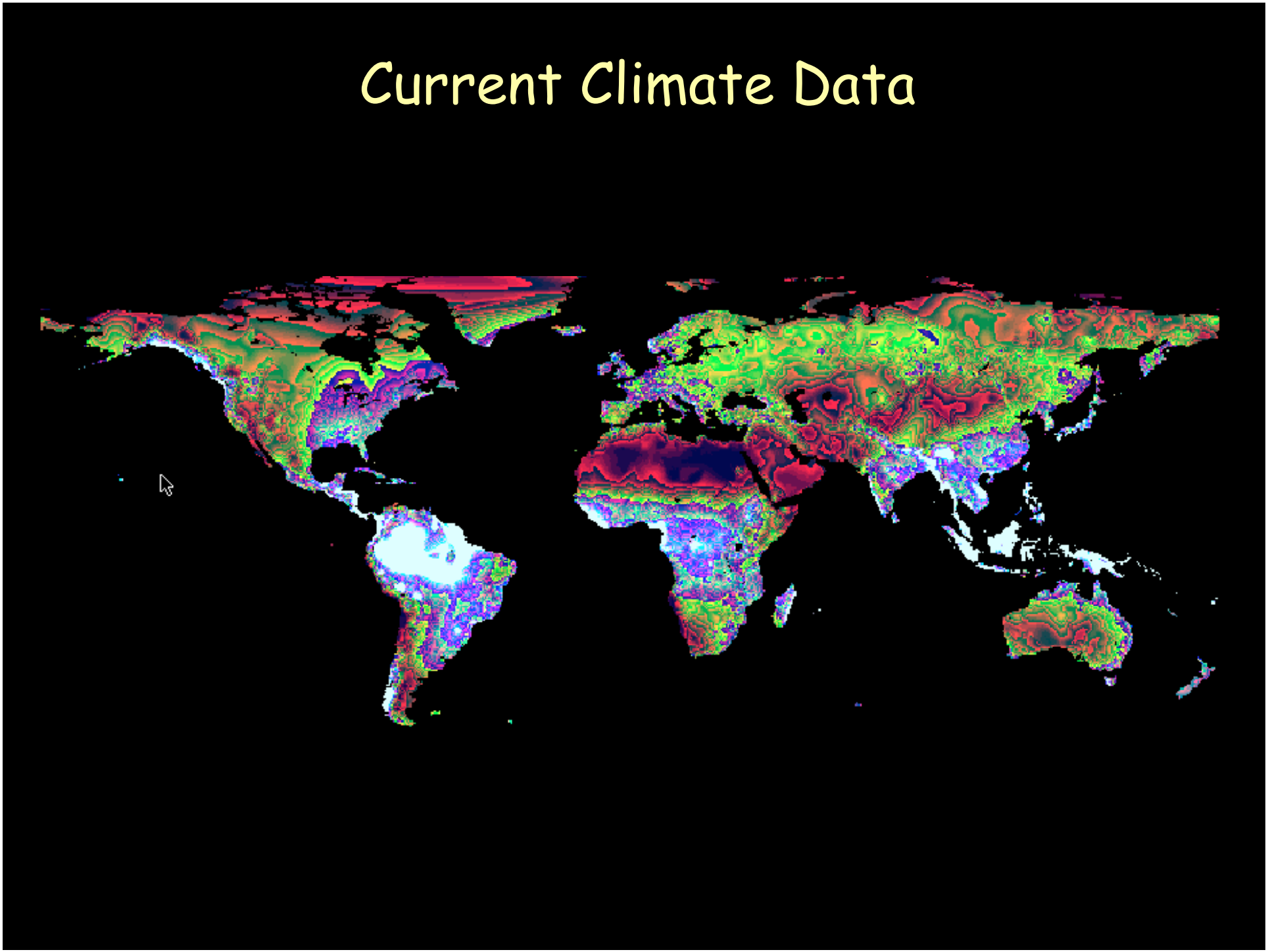
How do taxon and range size affect model accuracy?

Map and compare model errors from Random Forest models built for 6,933 vertebrates

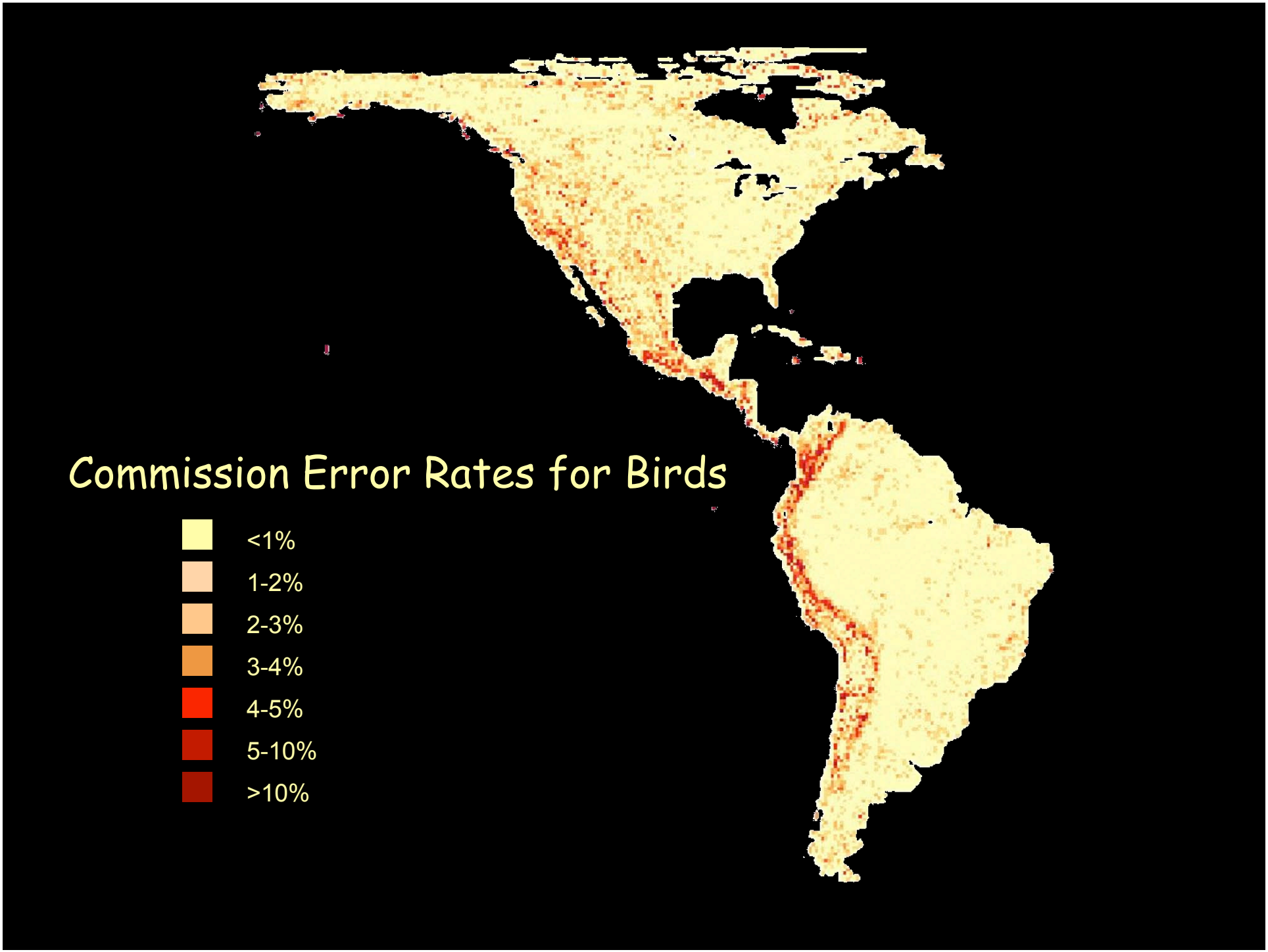
Current Geographic Ranges

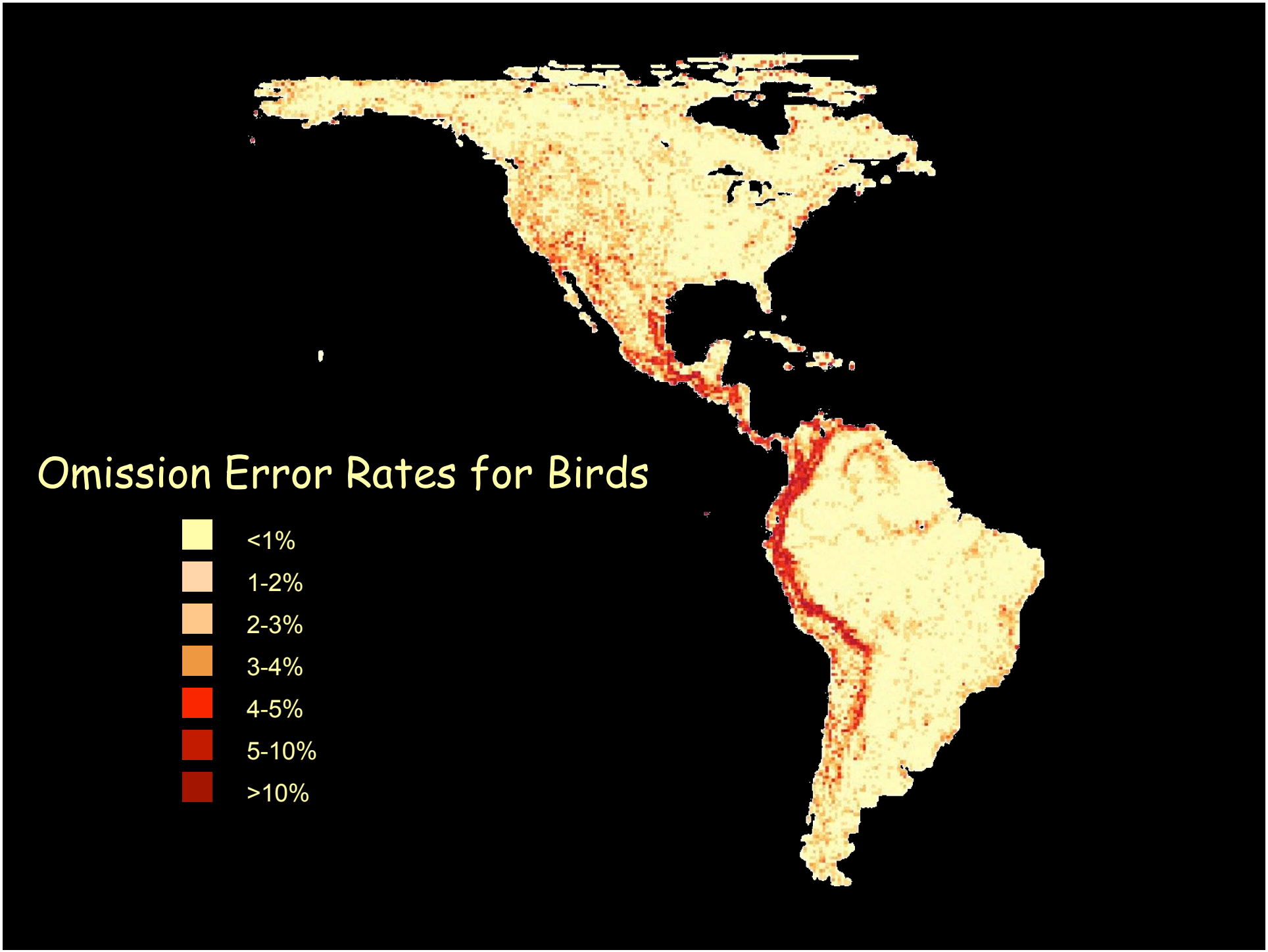
- 3,756 birds
  (Ridgely et al. 2003)
- 1,561 mammals
  (Patterson et al. 2003)
- 1,616 amphibians
  (IUCN, Conservation International, and NatureServe. 2004)

Current Climate Data



Results:
Spatial distribution of errors

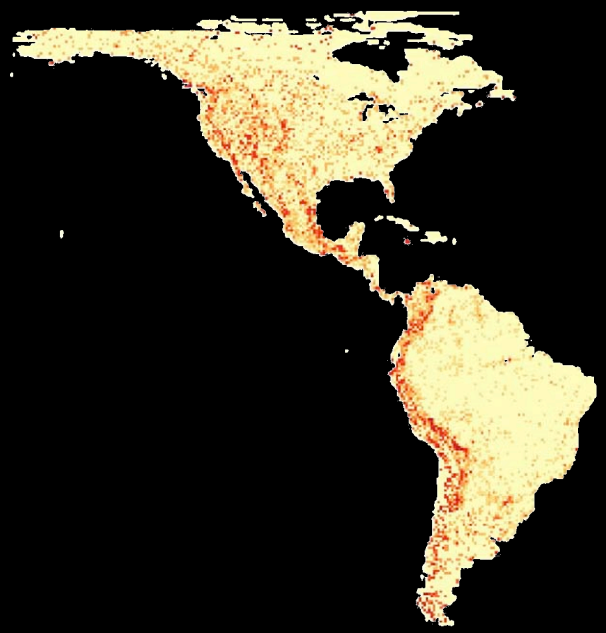

Commission Error Rates for Birds
<1%
1-2%
2-3%
3-4%
4-5%
5-10%
>10%


Omission Error Rates for Birds
<1%
1-2%
2-3%
3-4%
4-5%
5-10%
>10%

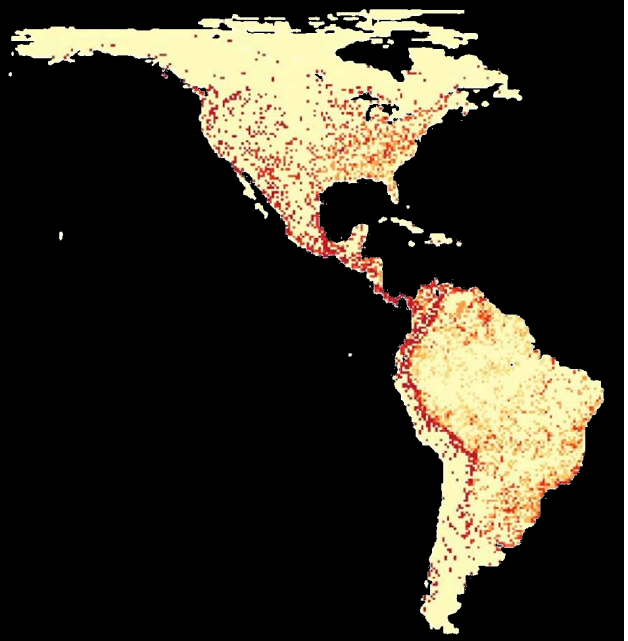
Omission Error Rates

Birds

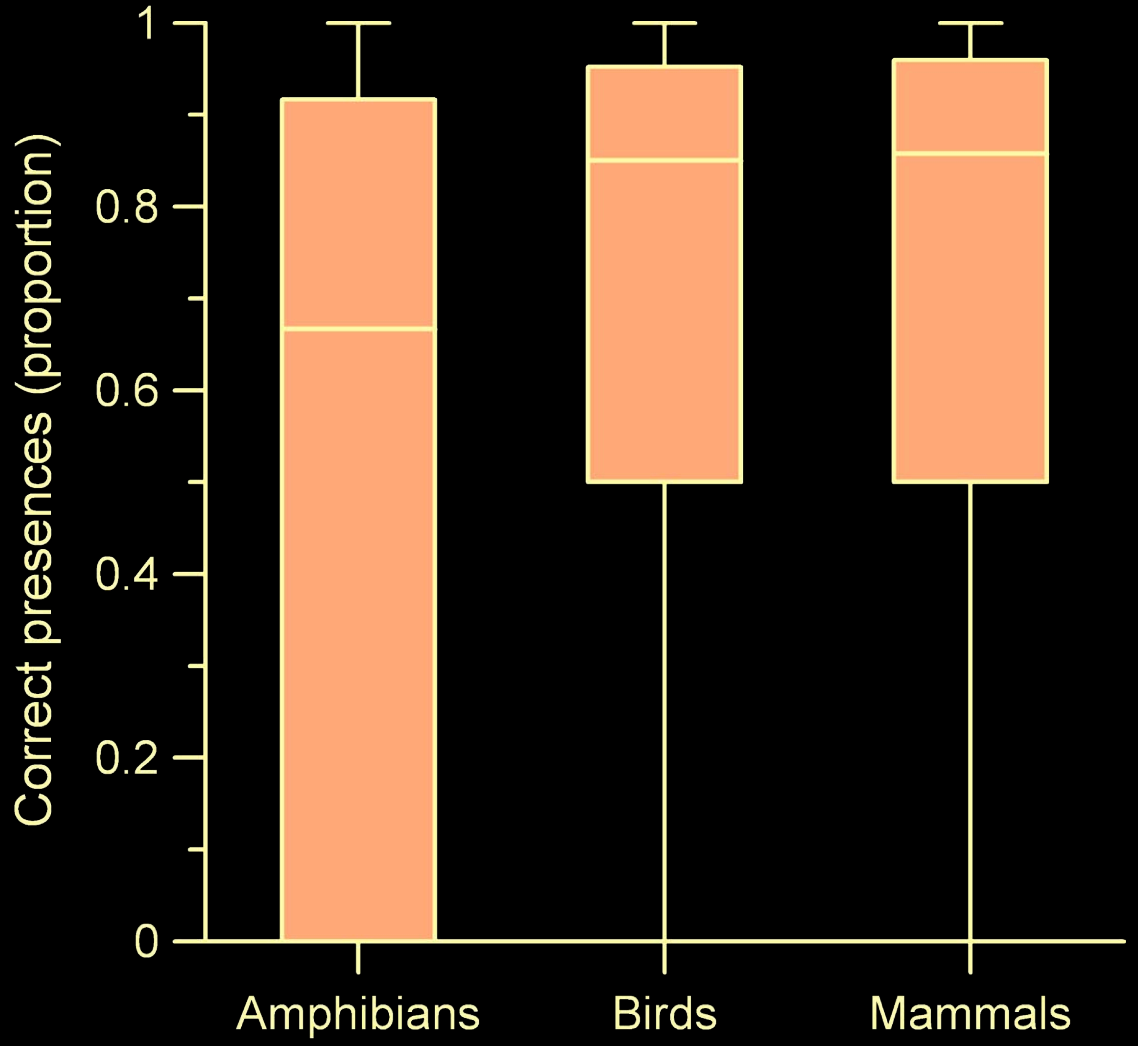
Mammals



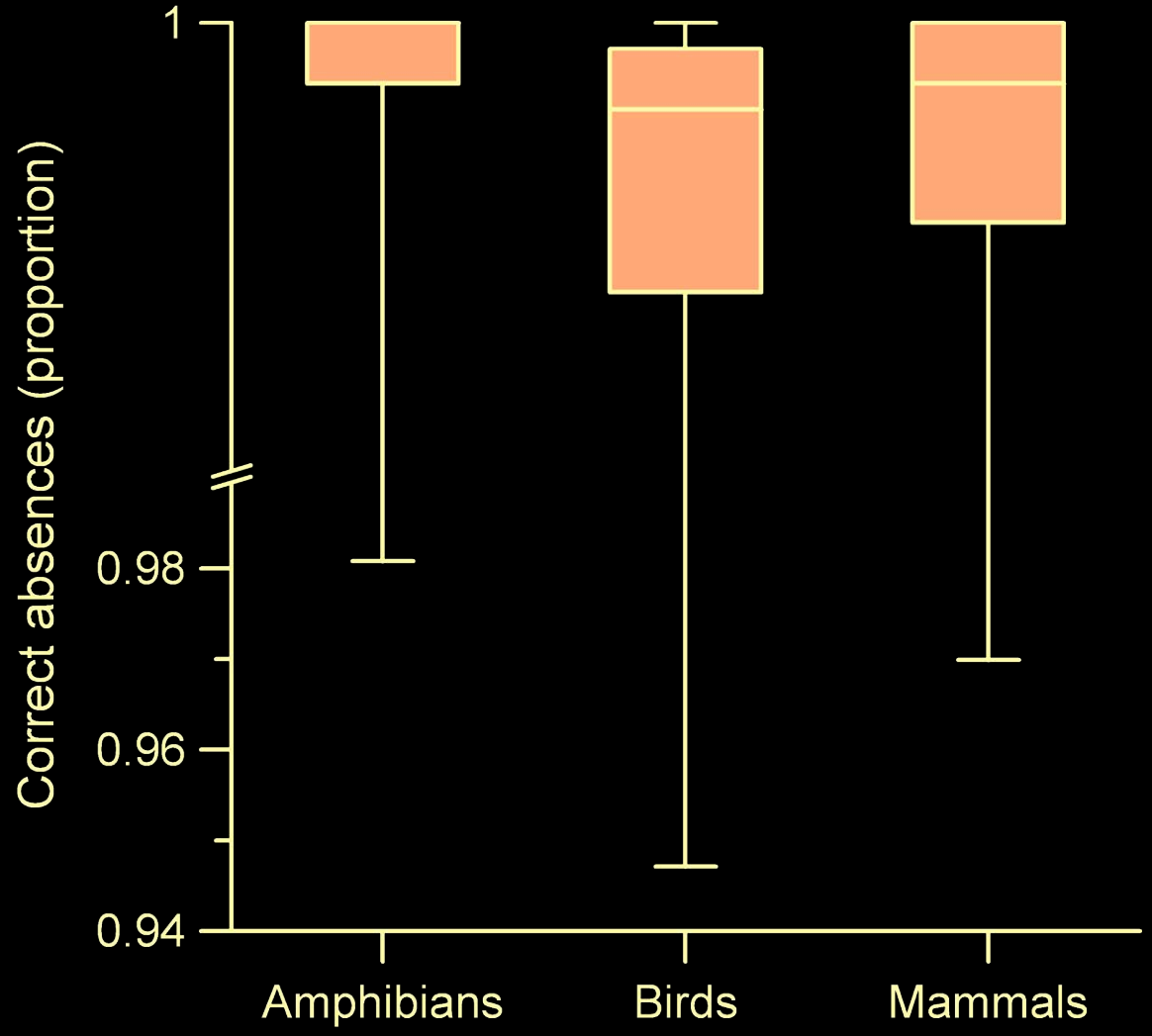
Amphibians



Correct Presences

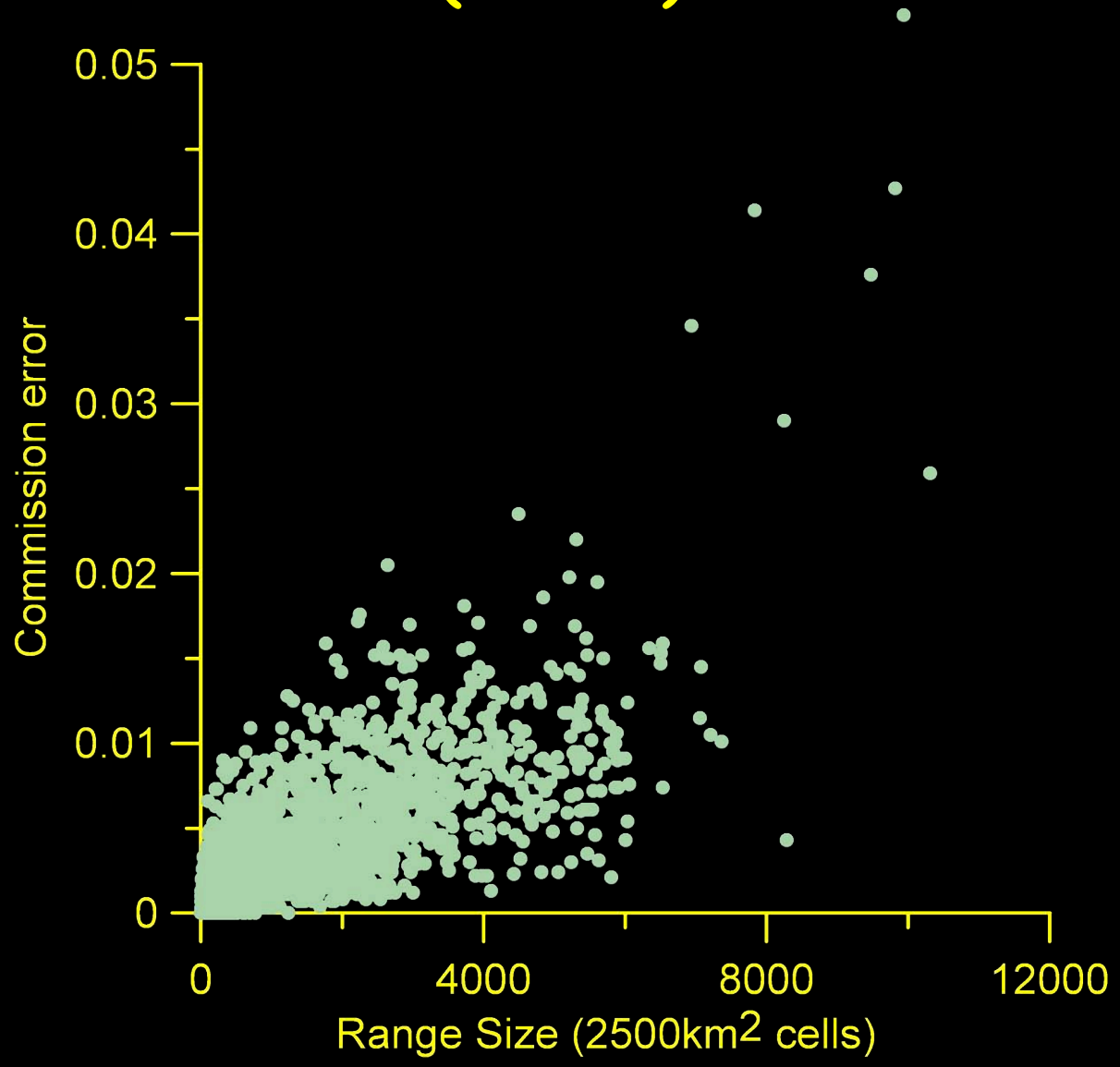


Correct Absences

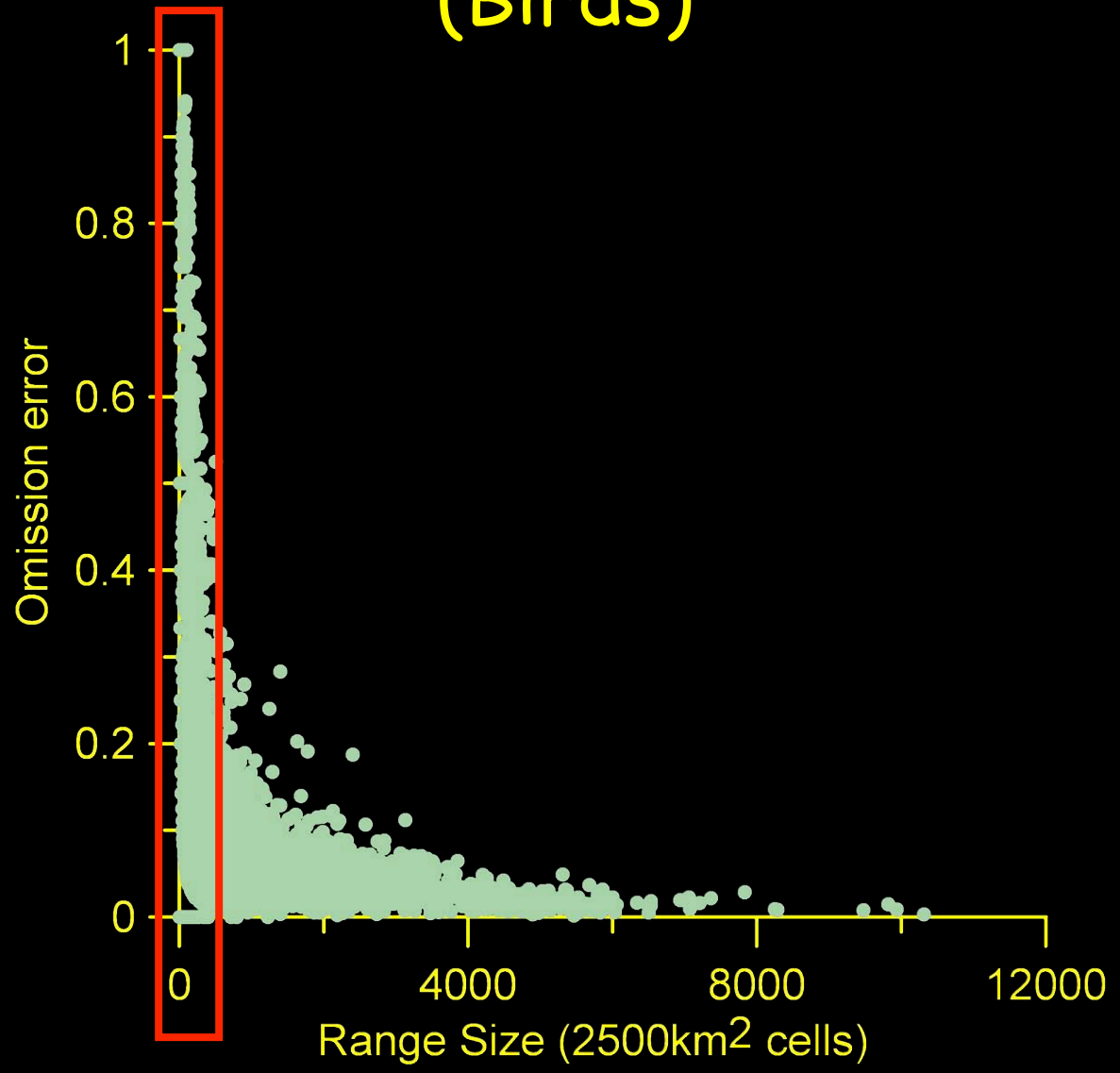


Results:
Effects of range size

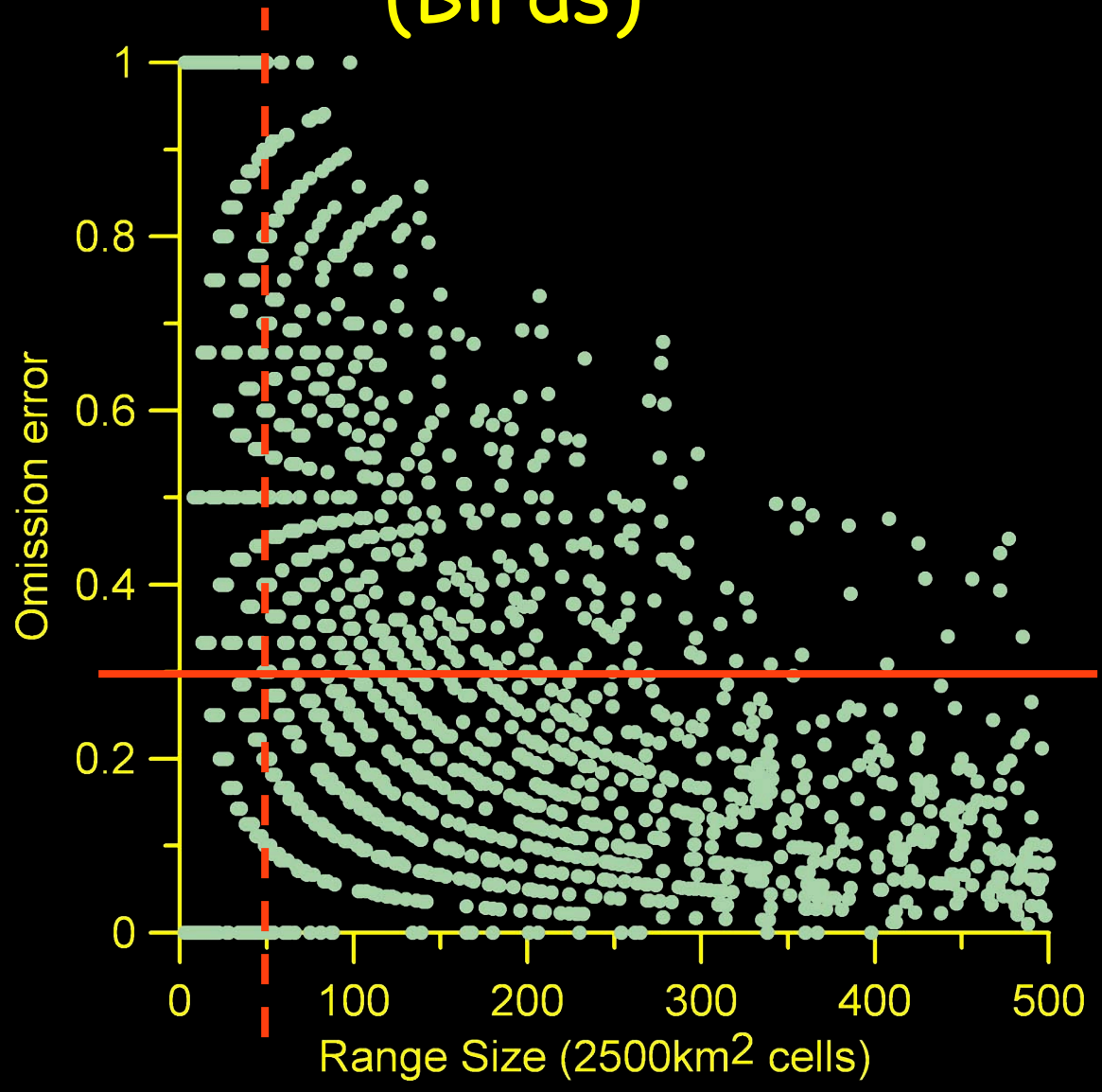
Effect of Range Size on Commission Errors (Birds)



Effect of Range Size on Omission Errors (Birds)



Effect of Range Size on Omission Errors (Birds)



Measures of Accuracy

Taxon	% spp with AUC > 0.90	% spp. with > 80% correct presences
Amphibians	74	32
Birds	95	55
Mammals	93	56

Overview

1. Background
2. Comparison of methods
3. Investigation of errors
4. Conclusions

Conclusions

1. Different models can produce drastically different predictions.

2. Model averaging may provide a highly accurate approach (e.g., Random Forest models).

3. Uncertainty varies geographically.

Conclusions

4. Effects of range size on model accuracy may generate bias.

5. Current estimates of model accuracy likely overestimate performance (AUC, Kappa).

What can we give conservation planners?

Relative vulnerability of areas to climate change.

General rules of thumb for conservation planning.

And…

Acknowledgements

This research was funded by the TNC Smith Fellows program and the EPA's Regional Vulnerability Assessment Program. I thank Sarah Shafer, Peter Kareiva, Ron Neilson, Andy Blaustein, and Denis White.

Additional Photo Credits:

Glacier National Park	Don Klosterman	Marmoset	Brandywine Zoo
Earth from space	NASA Goddard Space Flight Center	Tetons	© John Donahue, 1995
Butterflies	Jay Cossey	Sycan Marsh	Larry Olson
Yellowstone	© John Donahue, 1995	Grand Canyon	© John Donahue, 1995
Clear Lake Ridge	Eric Valentine	Spadefoot Toad	John White
Joshua Tree	© John Donahue, 1995	Red-breasted Sapsucker	© Peter La Tourrette 1992
North Cascades NP	QT Luong	Western Bluebird	© Peter La Tourrette 1995
Rainforest	David W. Hilbert, © Copyright CSIRO 'http://www.tfrc.csiro.au'		
Crater Lake	© John Donohue, 1995		