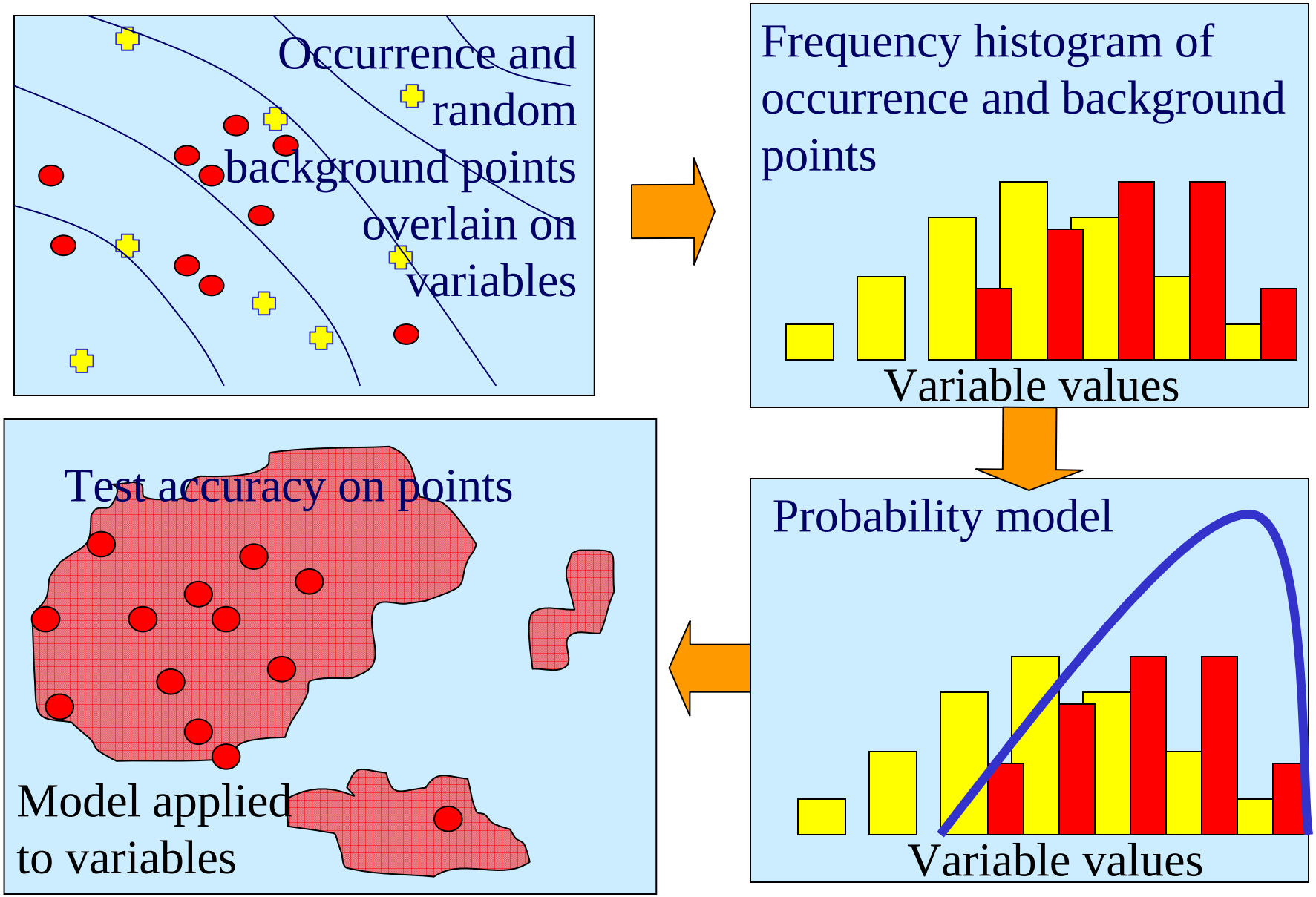
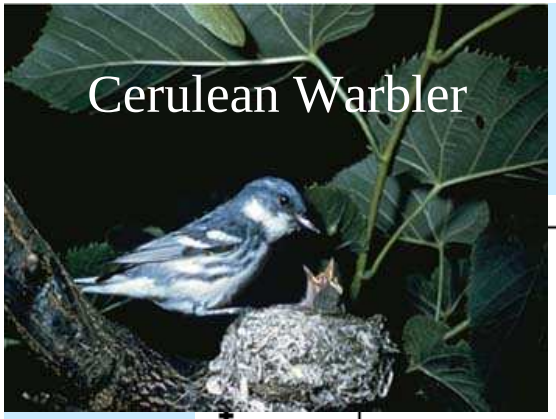
Advances in Ecological Niche Modeling

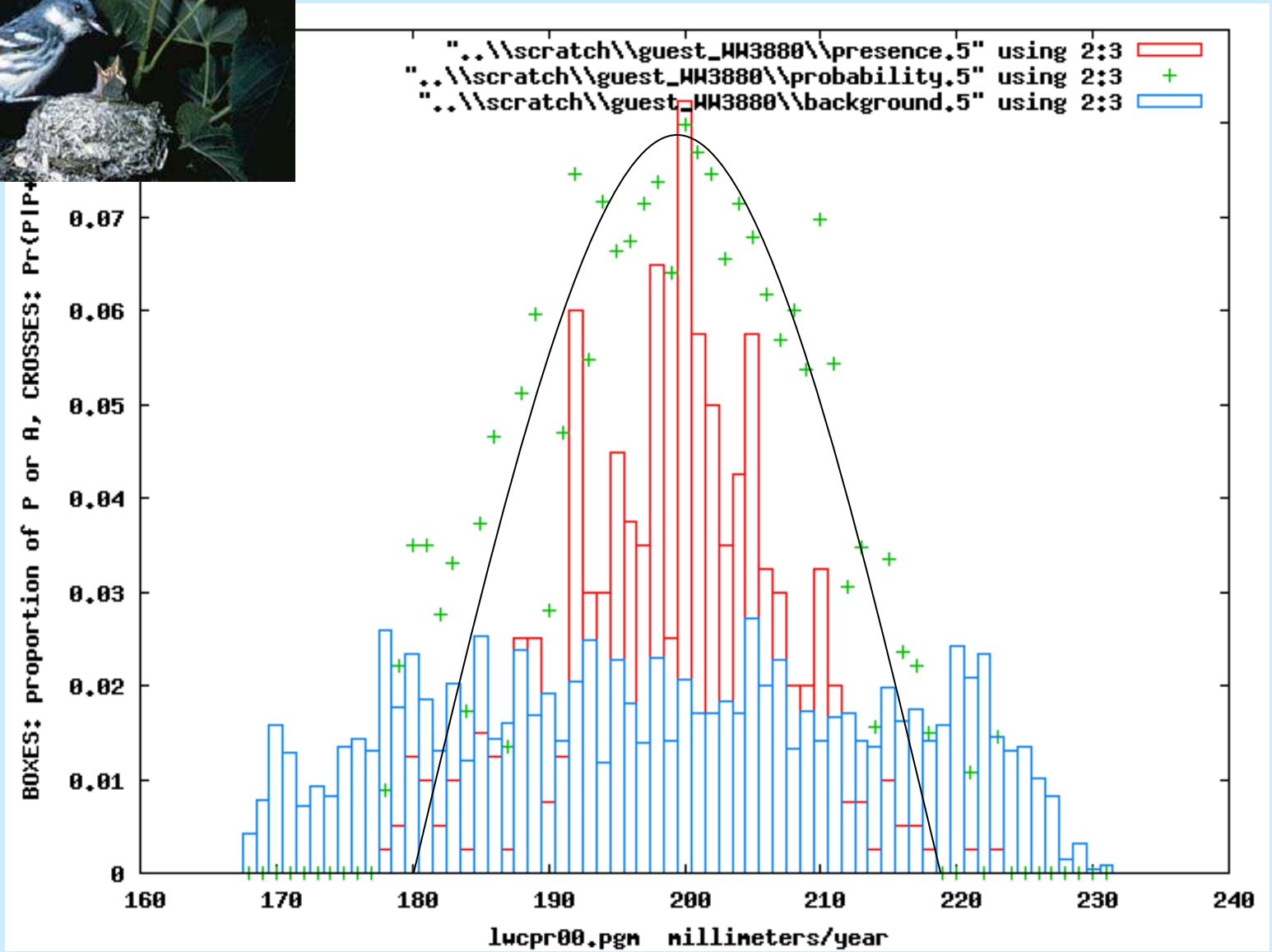
David R.B. Stockwell
San Diego Supercomputer Center
Email:davids@sdsc.edu
Web:http://biodi.sdsc.edu

Basic ENM process





Classical 1D parabolic response

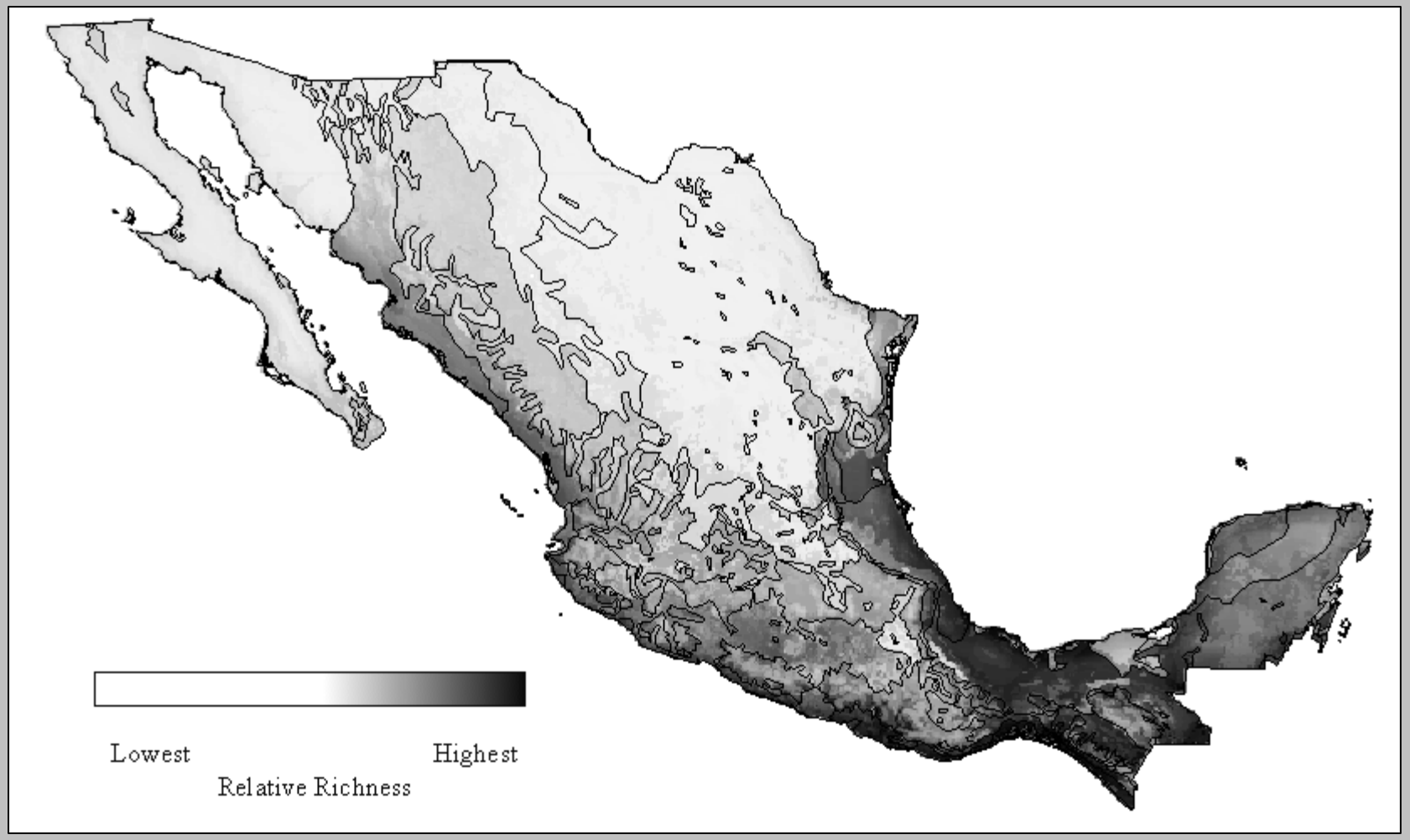


Why ENM?

Resolution

Use correlative variables of high resolution to improve resolution of scarce occurrence data.

Pattern of richness of birds - ENM vs. Veg map



Stockwell, D.R.B. and Peterson, A.T. (in press). Comparison of resolution of methods for mapping biodiversity patterns from point-occurrence data. *Ecological Indicators.*

GARP – Genetic Algorithm for Rule-set Production

1. Bioclimatic Envelope

```
IF      Dev=[1,2] AND StC=[0,1] AND SdC=[3,5] AND
   StQ=[1,3]   AND FlN=[2,3] AND Slp=[1,1] AND
   Ero=[3,3]
        THEN ExM= 3
```

2. Logistic Regression

```
IF       - Dev*0.10 - StC*0.10 + SdC*0.09`+ StQ*0.06
        - FlN*0.19 + Slp*0.40
        THEN ExM= 1
```

3. GARP rule

```
IF      Dev=[1,2] AND SdC=[0,0] AND StQ=[0,3]
        THEN ExM= 0
```

4. Atomic rule

```
IF      Dev= 0 AND StC= 1 AND SdC= 2`AND StQ= 1 AND
   FlN= 1
        AND Slp= 3 AND Ero= 1
        THEN ExM= 1
```

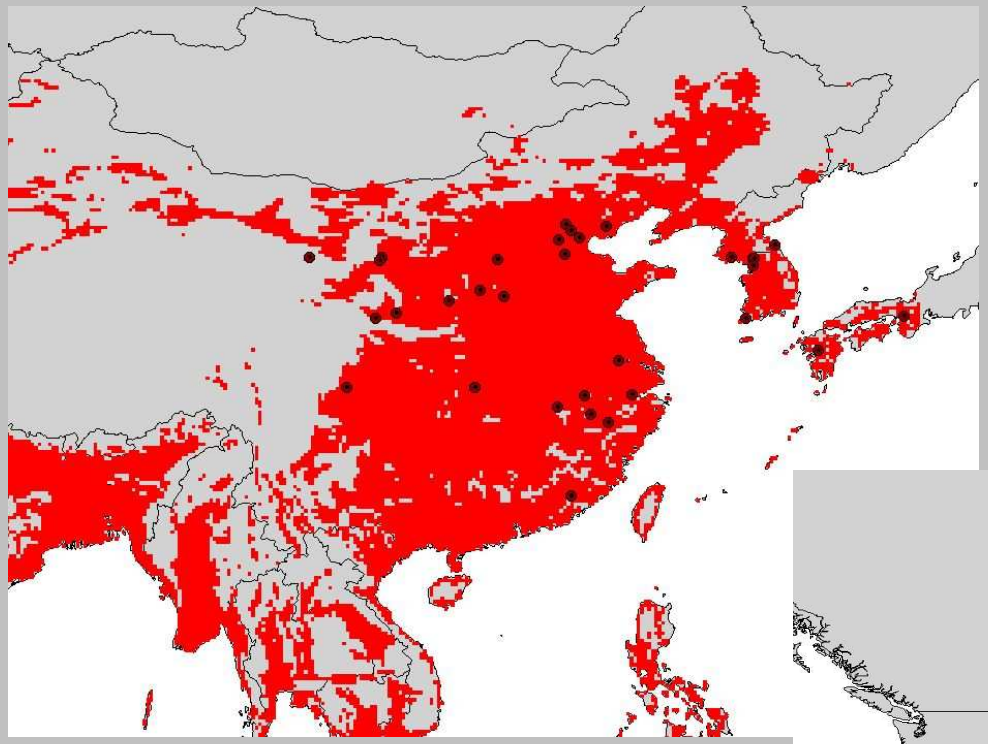
Stockwell D.R.B. and D. Peters 1999. The GARP Modeling System: problems and solutions to automated spatial prediction. *International Journal of Geographical Information Science* **13**:2 143-158.

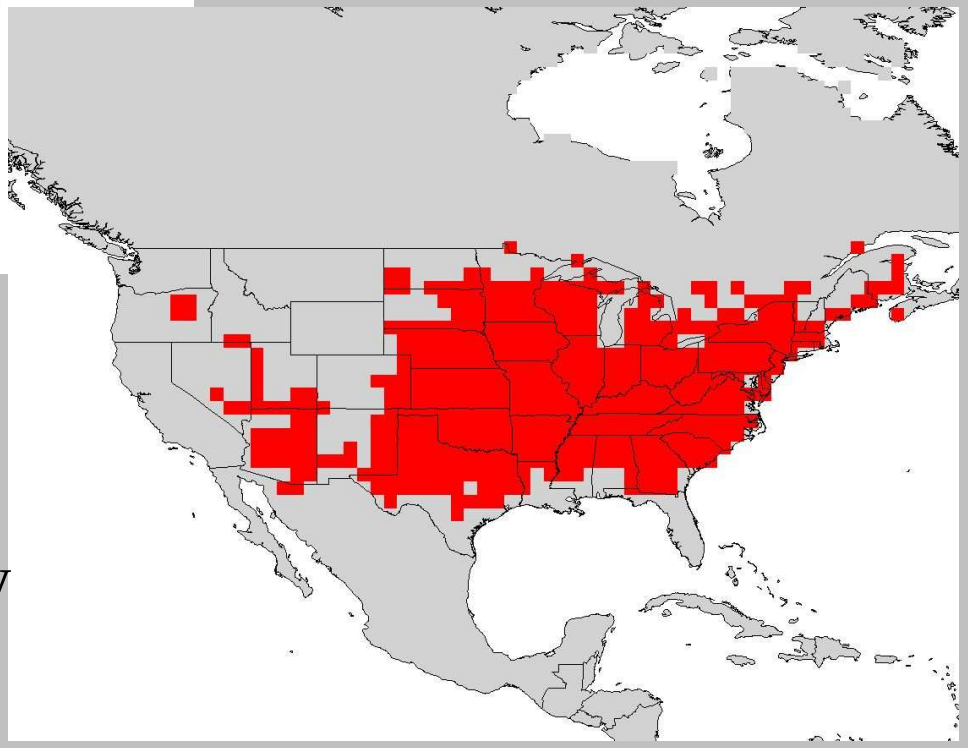
Why use multiple models?

Robustness

- Use of the consensus of multiple models compensates for problems in one model and provides adequate results on most occasions
- Does not necessarily provide the highest accuracy in a specific case

Novel hypotheses - Invasive Pests

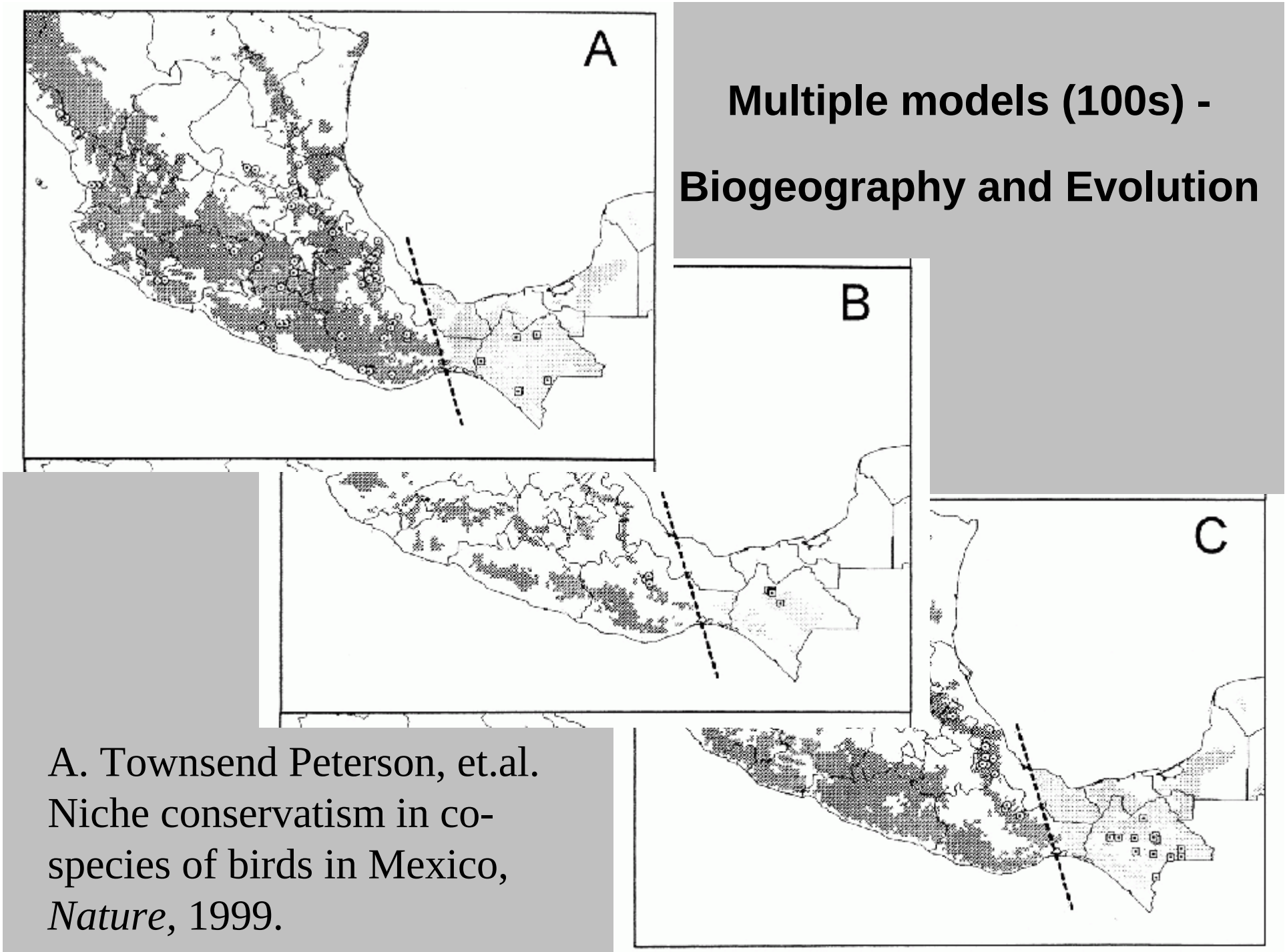




Predicted range of the Asian Longhorn Beetle in the USA (*Anoplophora glabripennis*) by A. Townsend Peterson (KU)

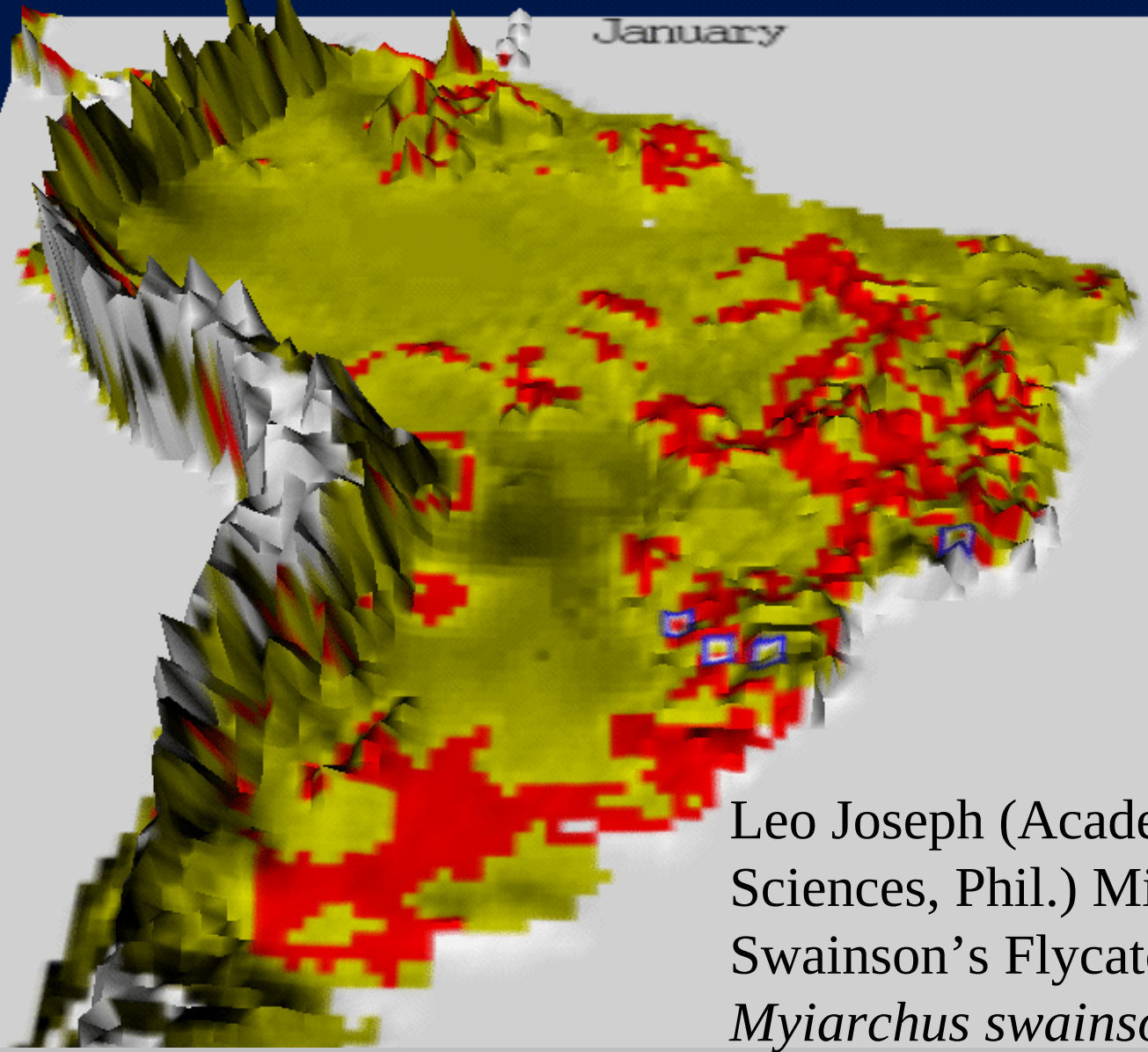
Multiple models (100s) -

Biogeography and Evolution



A. Townsend Peterson, et.al. Niche conservatism in co-species of birds in Mexico, *Nature*, 1999.

Novel hypotheses - Dynamic Migration Models

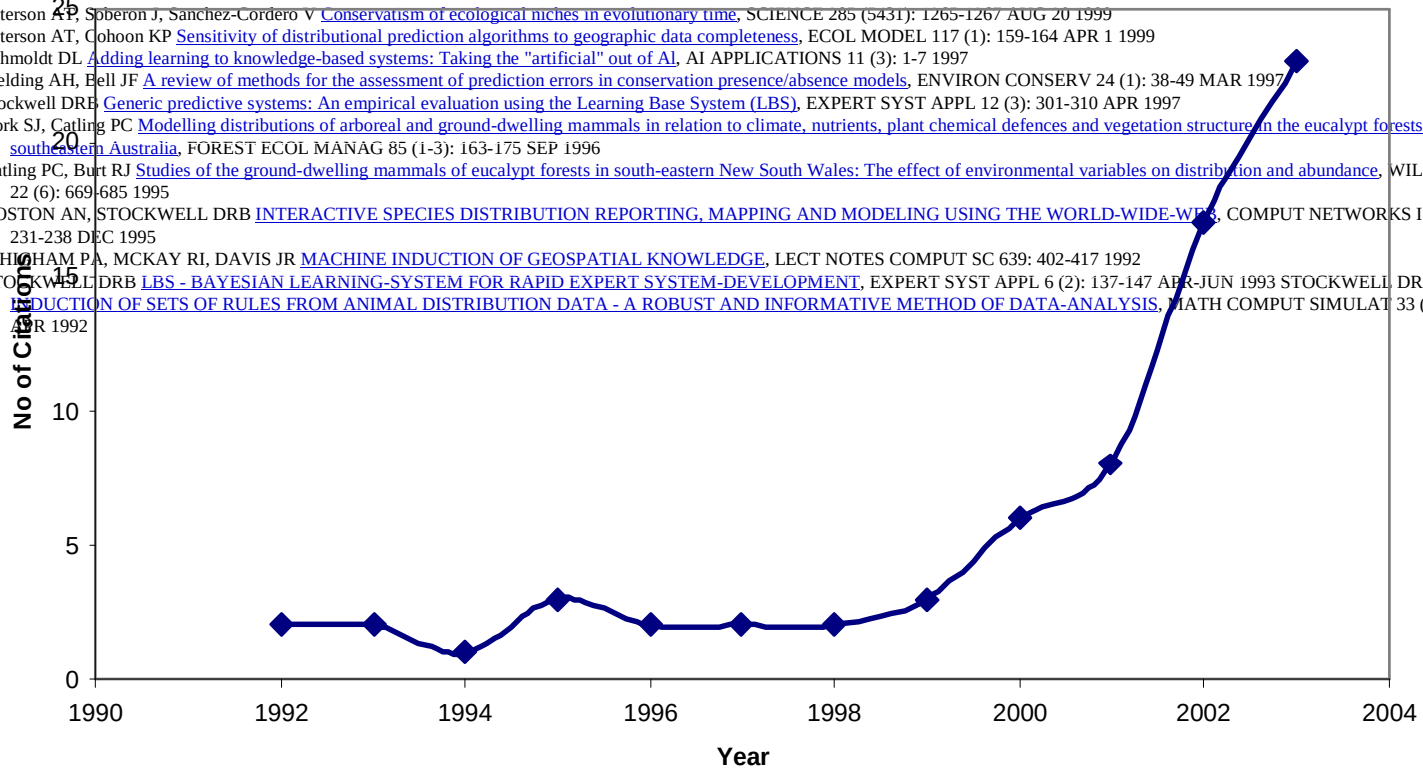


Leo Joseph (Academy of Sciences, Phil.) Migration of Swainson's Flycatcher *Myiarchus swainsoni* across South America

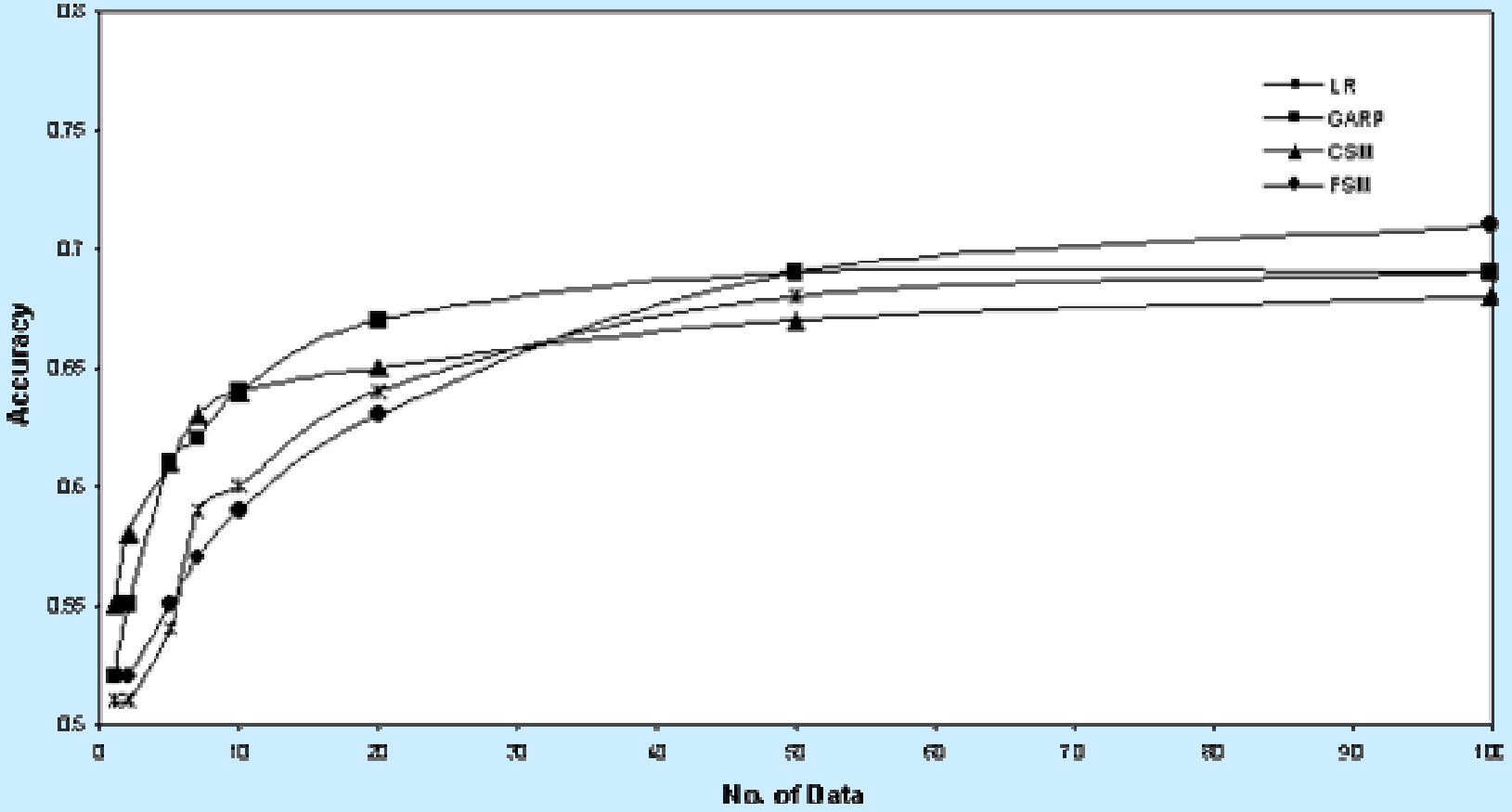
Growth in GARP Citations

Williams PH, Margules CR, Hilbert DW Data requirements and data sources for biodiversity priority area selection, J BIOSCIENCE 27 (4): 327-338 Suppl. 2 JUL 2002

34. Turner BJ, Chikumbo O, Davey SM Optimisation modelling of sustainable forest management at the regional level: an Australian example, ECOL MODEL 153 (1-2): 157-179 JUL 15 2002
35. Lim BK, Peterson AT, Engstrom MD Robustness of ecological niche modeling algorithms for mammals in Guyana, BIODIVERS CONSERV 11 (7): 1237-1246 JUL 2002
36. Peterson AT, Sanchez-Cordero V, Ben Beard C, et al. Ecologic niche modeling and potential reservoirs for Chagas disease, Mexico, EMERG INFECT DIS 8 (7): 662-667 JUL 2002
37. Peterson AT, Ball LG, Cohoon KP Predicting distributions of Mexican birds using ecological niche modelling methods, IBIS 144 (1): E27-E32 JAN 2002
38. Anderson RP, Gomez-Laverde M, Peterson AT Geographical distributions of spiny pocket mice in South America: insights from predictive models, GLOBAL ECOL BIOGEOGR 11 (2): 131-141 MAR 2002
39. Stockwell DRB, Peterson AT Effects of sample size on accuracy of species distribution models, ECOL MODEL 148 (1): 1-13 FEB 1 2002
40. Soberon J, Golubov J, Sarukhan J The importance of Opuntia in Mexico and routes of invasion and impact of Cactoblastis cactorum (Lepidoptera : Pyralidae), FLA ENTOMOL 84 (4): 486-492 DEC 2001
41. McKay RI Variants of genetic programming for species distribution modelling - fitness sharing, partial functions, population evaluation, ECOL MODEL 146 (1-3): 231-241 Sp. Iss. SI DEC 1 2001
42. Bobbin J, Recknagel F Knowledge discovery for prediction and explanation of blue-green algal dynamics in lakes by evolutionary algorithms, ECOL MODEL 146 (1-3): 253-262 Sp. Iss. SI DEC 1 2001
43. Mackey BG, Lindenmayer DB Towards a hierarchical framework for modelling the spatial distribution of animals, J BIOGEOGR 28 (9): 1147-1166 SEP 2001
44. Peterson AT, Sanchez-Cordero V, Soberon J, et al. Effects of global climate change on geographic distributions of Mexican Cracidae, ECOL MODEL 144 (1): 21-30 OCT 1 2001
45. Peterson AT Predicting species' geographic distributions based on ecological niche modeling, CONDOR 103 (3): 599-605 AUG 2001
46. Peterson AT, Vieglais DA Predicting species invasions using ecological niche modeling: New approaches from bioinformatics attack a pressing problem, BIOSCIENCE 51 (5): 363-371 MAY 2001
47. Pearce J, Ferrier S The practical value of modelling relative abundance of species for regional conservation planning: a case study, BIOL CONSERV 98 (1): 33-43 MAR 2001
48. Soberon JM, Llorente JB, Onate L The use of specimen-label databases for conservation purposes: an example using Mexican Papilionid and Pierid butterflies, BIODIVERS CONSERV 9 (10): 1441-1466 OCT 2000
49. Godown ME, Peterson AT Preliminary distributional analysis of US endangered bird species, BIODIVERS CONSERV 9 (9): 1313-1322 SEP 2000
50. Whigham PA Induction of a marsupial density model using genetic programming and spatial relationships, ECOL MODEL 131 (2-3): 299-317 JUL 1 2000
51. Sanchez-Cordero V, Martinez-Meyer E Museum specimen data predict crop damage by tropical rodents, P NATL ACAD SCI USA 97 (13): 7074-7077 JUN 20 2000
52. Cork SJ, Hume LD, Foley WJ Improving habitat models and their utility in koala conservation, CONSERV BIOL 14 (3): 660-668 JUN 2000
53. Joseph L, Stockwell D Temperature-based models of the migration of Swainson's Flycatcher (Myiarchus swainsoni) across South America: A new use for museum specimens of migratory birds, P ACAD NAT SCI PHILA 150: 293-300 APR 14 2000
54. Peterson AT, Soberon J, Sanchez-Cordero V Conservatism of ecological niches in evolutionary time, SCIENCE 285 (5431): 1265-1267 AUG 20 1999
55. Peterson AT, Cohoon KP Sensitivity of distributional prediction algorithms to geographic data completeness, ECOL MODEL 117 (1): 159-164 APR 1 1999
56. Schmoldt DL Adding learning to knowledge-based systems: Taking the "artificial" out of AI, AI APPLICATIONS 11 (3): 1-7 1997
57. Fielding AH, Bell JF A review of methods for the assessment of prediction errors in conservation presence/absence models, ENVIRON CONSERV 24 (1): 38-49 MAR 1997
58. Stockwell DRB Generic predictive systems: An empirical evaluation using the Learning Base System (LBS), EXPERT SYST APPL 12 (3): 301-310 APR 1997
59. Cork SJ, Catling PC Modelling distributions of arboreal and ground-dwelling mammals in relation to climate, nutrients, plant chemical defences and vegetation structure in the eucalypt forests of southeastern Australia, FOREST ECOL MANAG 85 (1-3): 163-175 SEP 1996
60. Catling PC, Burt RJ Studies of the ground-dwelling mammals of eucalypt forests in south-eastern New South Wales: The effect of environmental variables on distribution and abundance, WILDLIFE RES 22 (6): 669-685 1995
61. BOSTON AN, STOCKWELL DRB INTERACTIVE SPECIES DISTRIBUTION REPORTING, MAPPING AND MODELING USING THE WORLD-WIDE-WEB, COMPUT NETWORKS ISDN 28 (1-2): 231-238 DEC 1995
62. WHIGHAM PA, MCKAY RI, DAVIS JR MACHINE INDUCTION OF GEOSPATIAL KNOWLEDGE, LECT NOTES COMPUT SC 639: 402-417 1992
63. STOCKWELL DRB LBS - BAYESIAN LEARNING-SYSTEM FOR RAPID EXPERT SYSTEM-DEVELOPMENT, EXPERT SYST APPL 6 (2): 137-147 APR-JUN 1993 STOCKWELL DRB, NOBLE IR INDUCTION OF SETS OF RULES FROM ANIMAL DISTRIBUTION DATA - A ROBUST AND INFORMATIVE METHOD OF DATA-ANALYSIS, MATH COMPUT SIMULAT 33 (5-6): 385-390 APR 1992



Accuracy of Models by Number of Occurrence Points



Stockwell, D.R.B. and Townsend Peterson, A.T. 2002. Effects of sample size on accuracy of species distribution models. *Ecological Modelling* **148**:1-13

Why continue collecting?

Accuracy

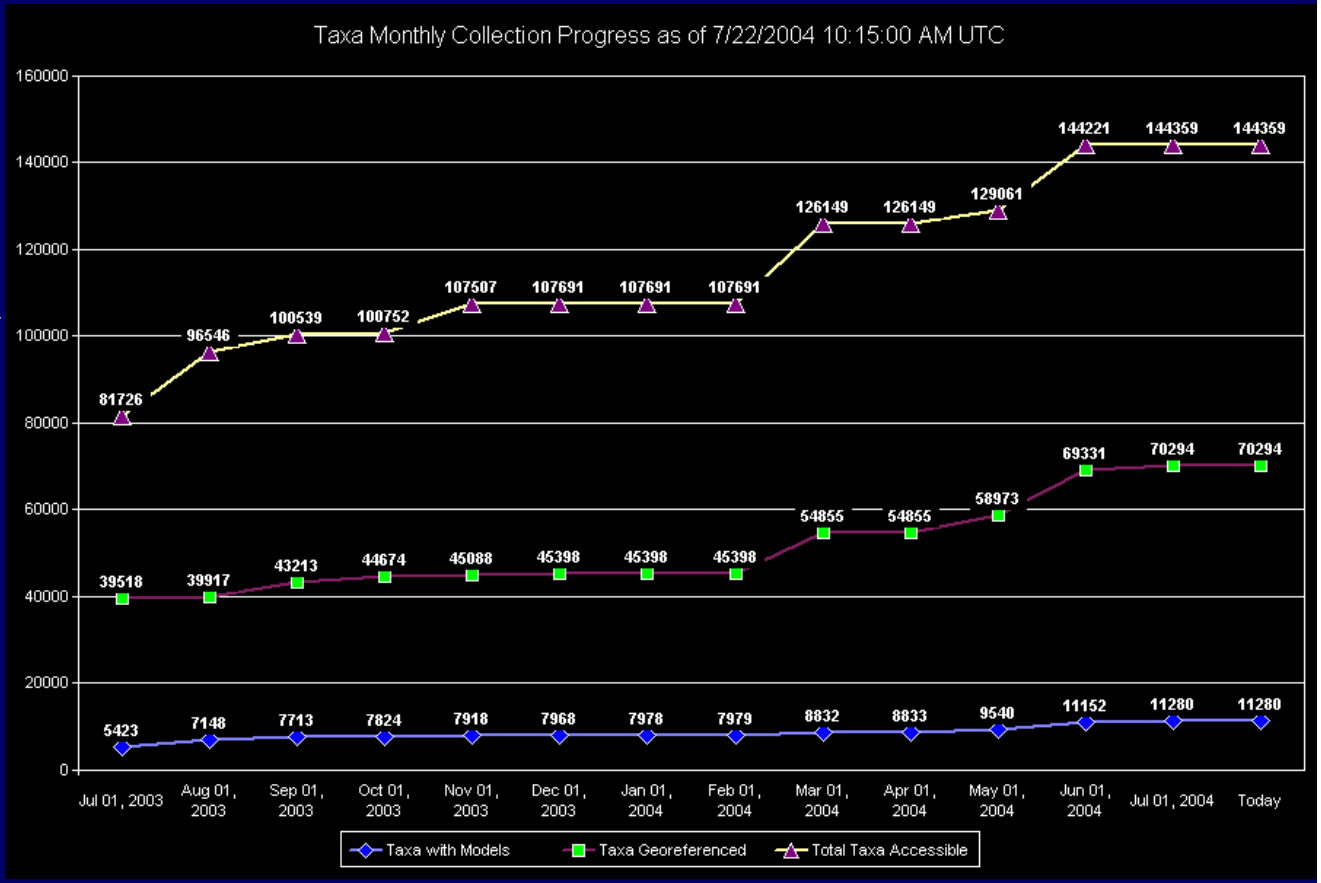
- Most methods perform well with enough data
- Adequate occurrence points is a major limitation to accuracy
- 10 points produces 90% of maximum accuracy
- Of all museum specimens, 49% have any georeferences, 8% have > 10 georeferences

Data Sources: Museum Data http://speciesanalyst.net

- **Institution Server Database Status**
- KUNHM habanero.nhm.ukans.edu:210 KUBirds OK
- KUNHM habanero.nhm.ukans.edu:210 KUMammals OK
- KUNHM habanero.nhm.ukans.edu:210 KBSPlants OK
- UNAM fcbiologia.fciencias.unam.mx:210 Mamife OK
- UNAM fcbiologia.fciencias.unam.mx:210 MexBirds OK

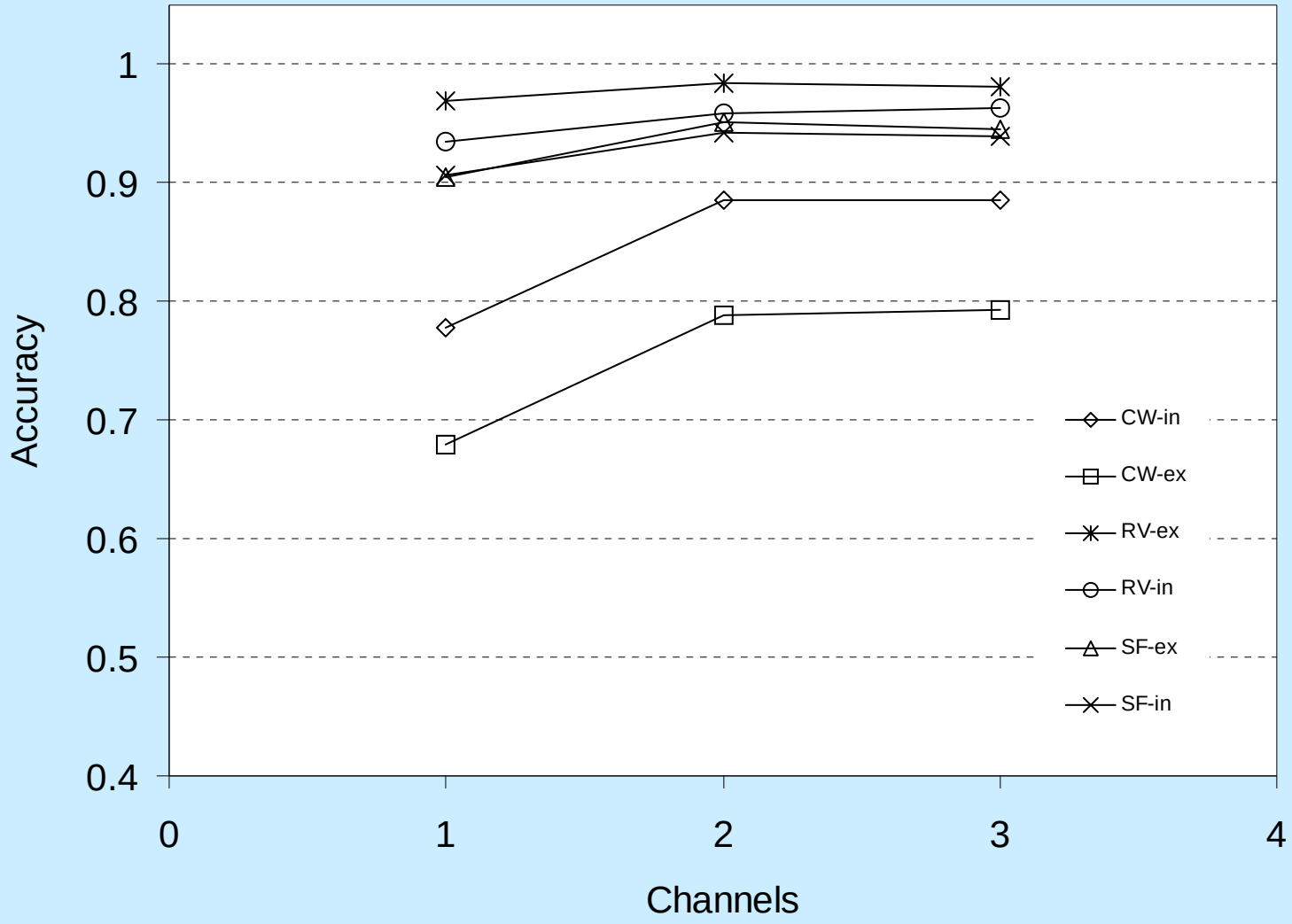
Lifemapper - www.lifemapper.org

- a screensaver grid computing project
- develop a fauna and flora using the worlds museum data
- installed base of over 30,000 screensavers on personal computers.



Stockwell, D.R.B., Beach J.H., Stewart A., Vorontsov G., Vieglais D., and Scachetti Pereira R. (in press) The use of the GARP genetic algorithm and Internet grid computing in the Lifemapper world atlas of species biodiversity

Accuracy by number correlates



Why fewer correlates?

Explanation

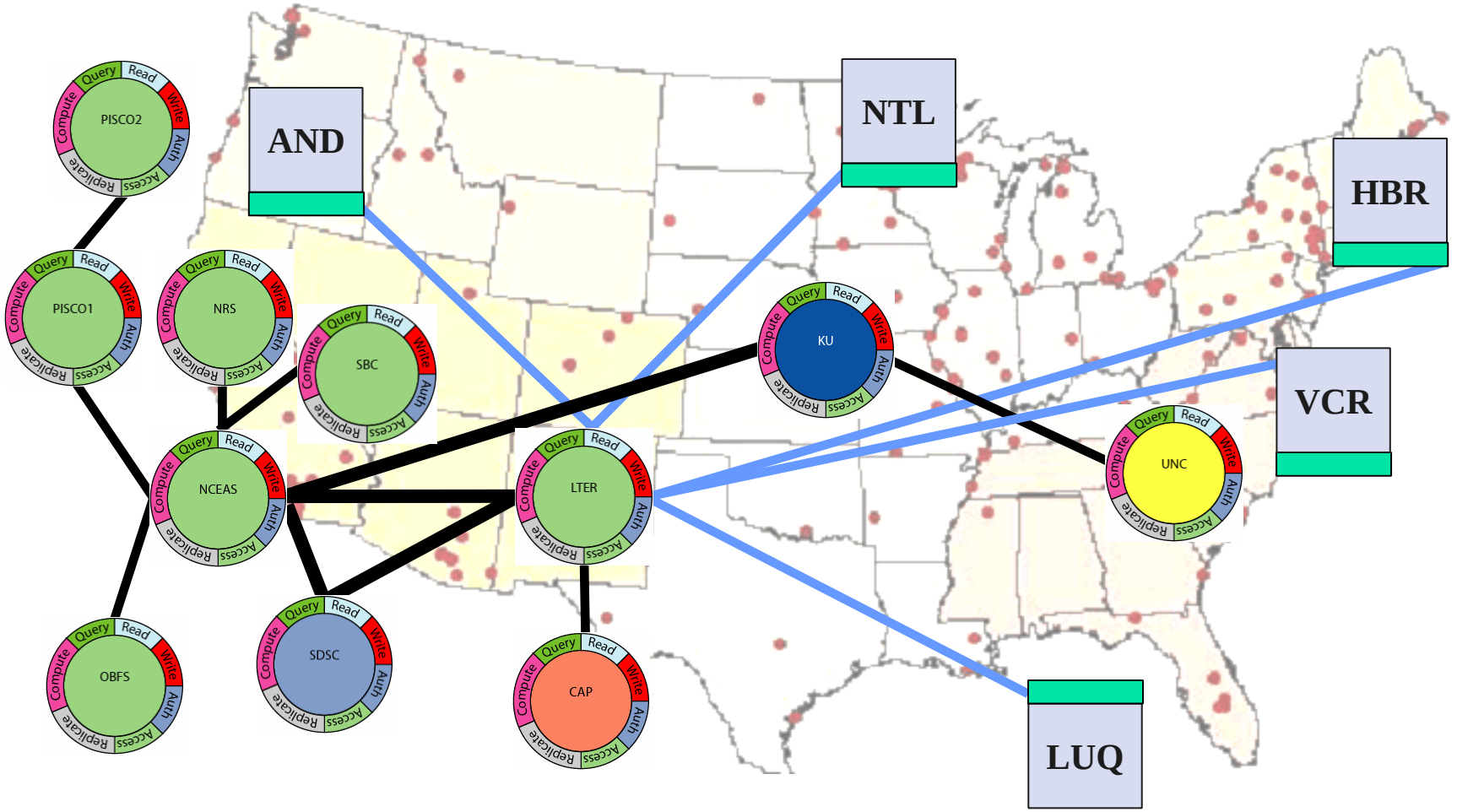
- Identify those factors that maximize accuracy
- Parsimonious model - with the right variables, and non-linear response can be as few as one or two

Some sources of correlative data

Terrestrial >500 Marine >500

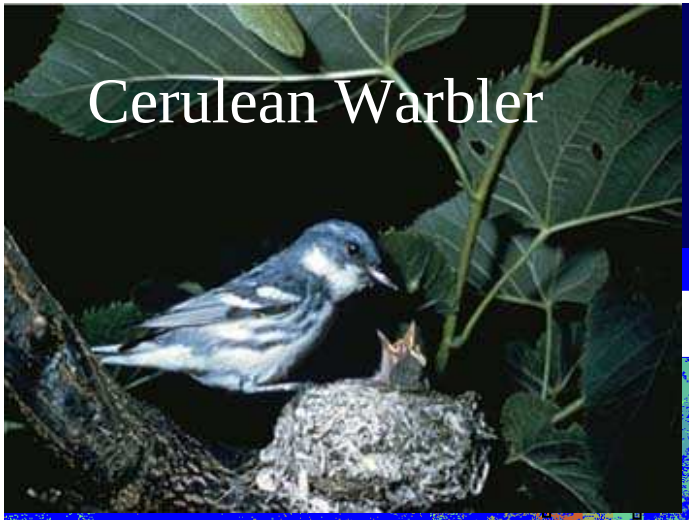
- Global Ecosystems database (1deg - 1km) Topographic (DEM), Atmospheric, Climatic and Meteorologic, Hydrologic, Oceanographic, Ecosystems and Biogeochemical Dynamics, Geological and Geophysical Data - 10GB
- Landsat 1km - %cover of treecover, evergreen, deciduous, broadleaf - 4GB
- Marine - productivity, annual temperatures and deviations, salinity, at verious depths
- Satellite - MODIS 12 level 3-4 land and 15 level 3-4 ocean products 1km to 250m tiled, 1TB to 4TB per annum
- Digital Elevation - min, max, median elevation, slope, aspect, rugosity, hydrological variables - 30m - 1TB
- BIOCLIM variables

SEEK EcoGrid
http://seek.ecoinformatics.org



Metacat node
VegBank node
Xanthoria node
SRB node
DiGIR node
Legacy system

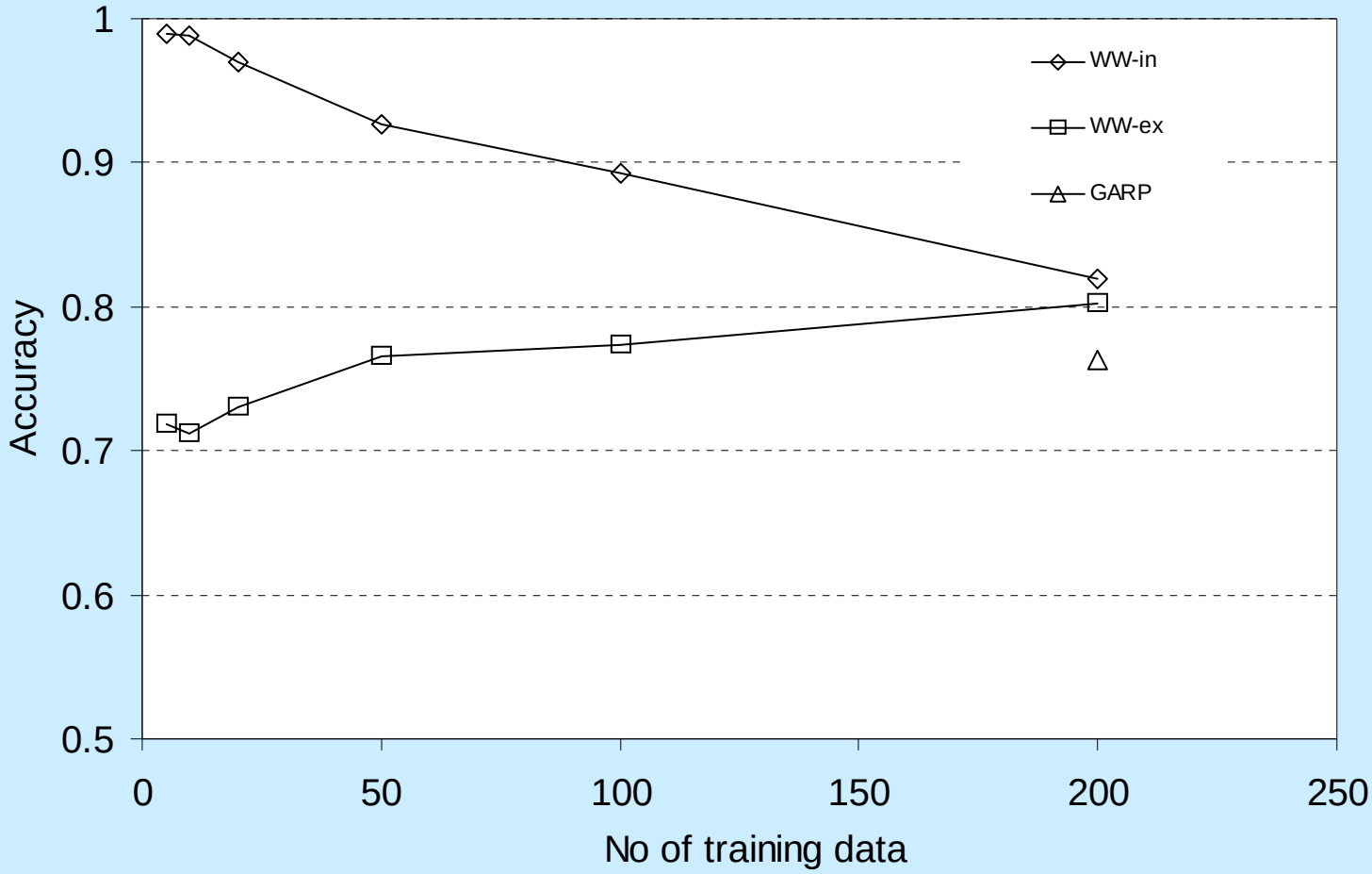
LTER Network (24)
Natural History Collections (>> 100)
Organization of Biological Field Stations (180)
UC Natural Reserve System (36)
Partnership for Interdisciplinary Studies of Coastal Oceans (4)
Multi-agency Rocky Intertidal Network (60)



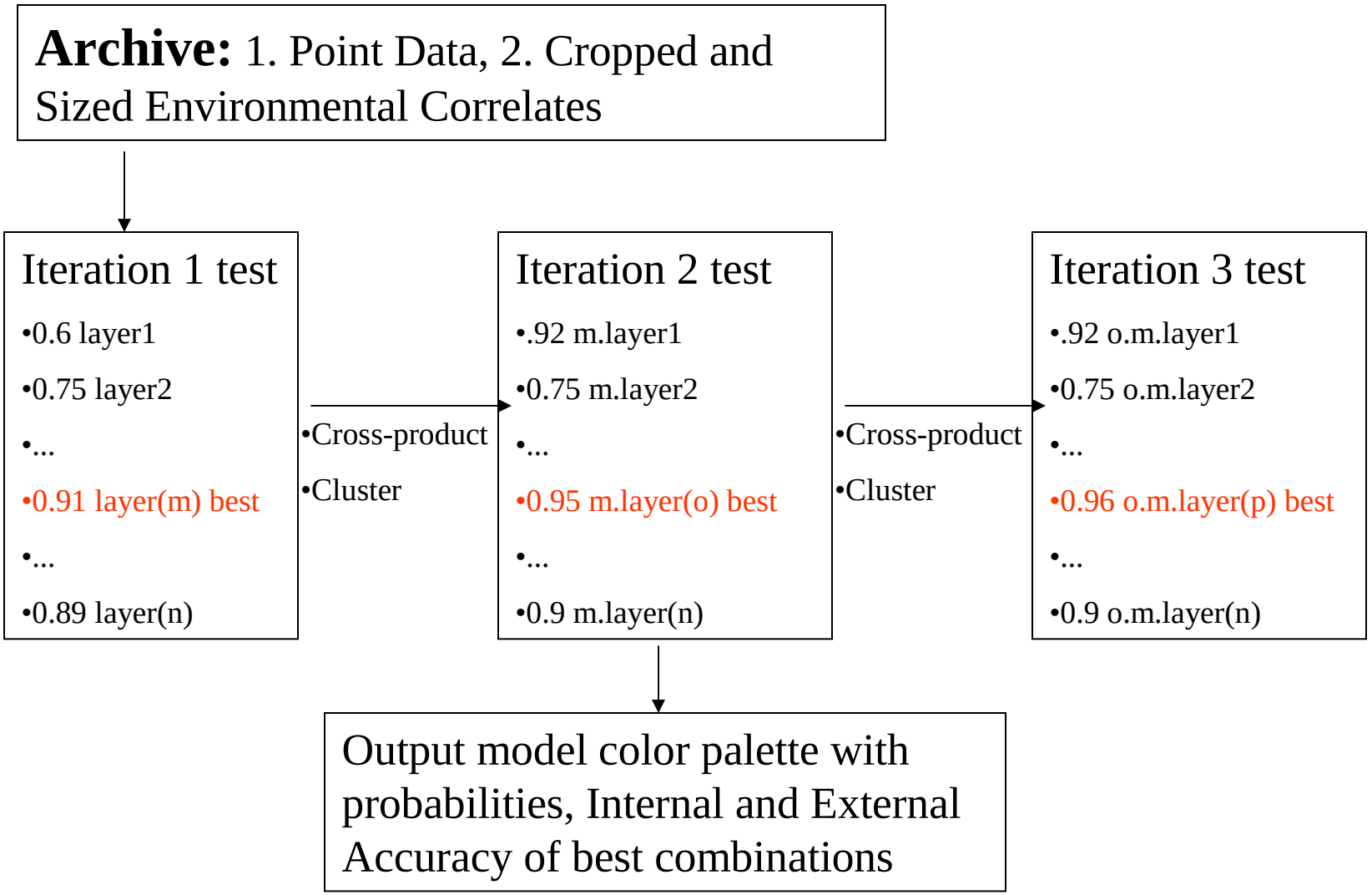
WhyWhere?

Answers the question
"Where is it and why?"
on a global scale.

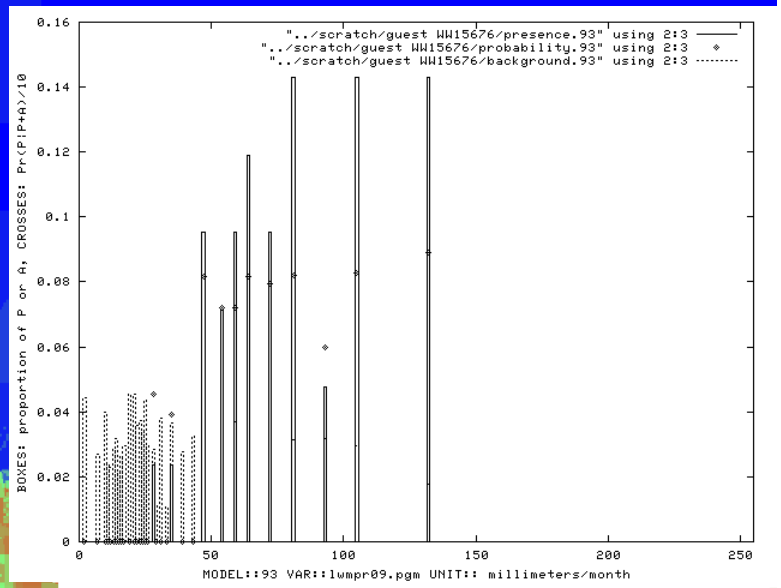
WhyWhere? vs. GARP ↑14%

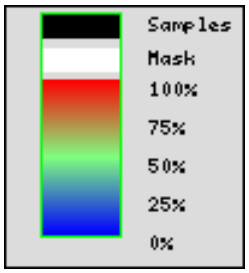


WhyWhere? Parallel Prediction Algorithm



WW Distribution of Jaguar - *Panthera onca*





Víctor Sánchez-Cordero, Sahotra Sarkar, David Stockwell and Howei Liu

Competition limits the southern distribution of bobcats *Lynx rufus*

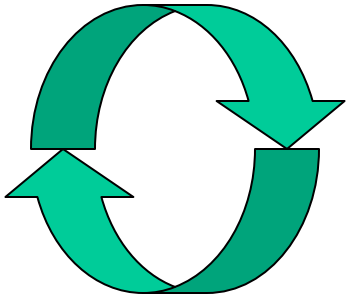
Socio-Scientific Summary

- Increased Resolution ⇐ Correlative Models
- Increased Robustness ⇐ Consensus Models
- Increased Accuracy ⇐ Increased Occurrence Data ⇐ Natural History Museums
- Increased Explanation ⇐ Increased Correlative Data ⇐ NASA

Advance in ENM

1. Empirical and Theoretical Statistical Studies

4. New Science Applications



2. Algorithmic Software Developments

3. Information Infrastructure Development

http://biodi.sdsc.edu/ww_home.html