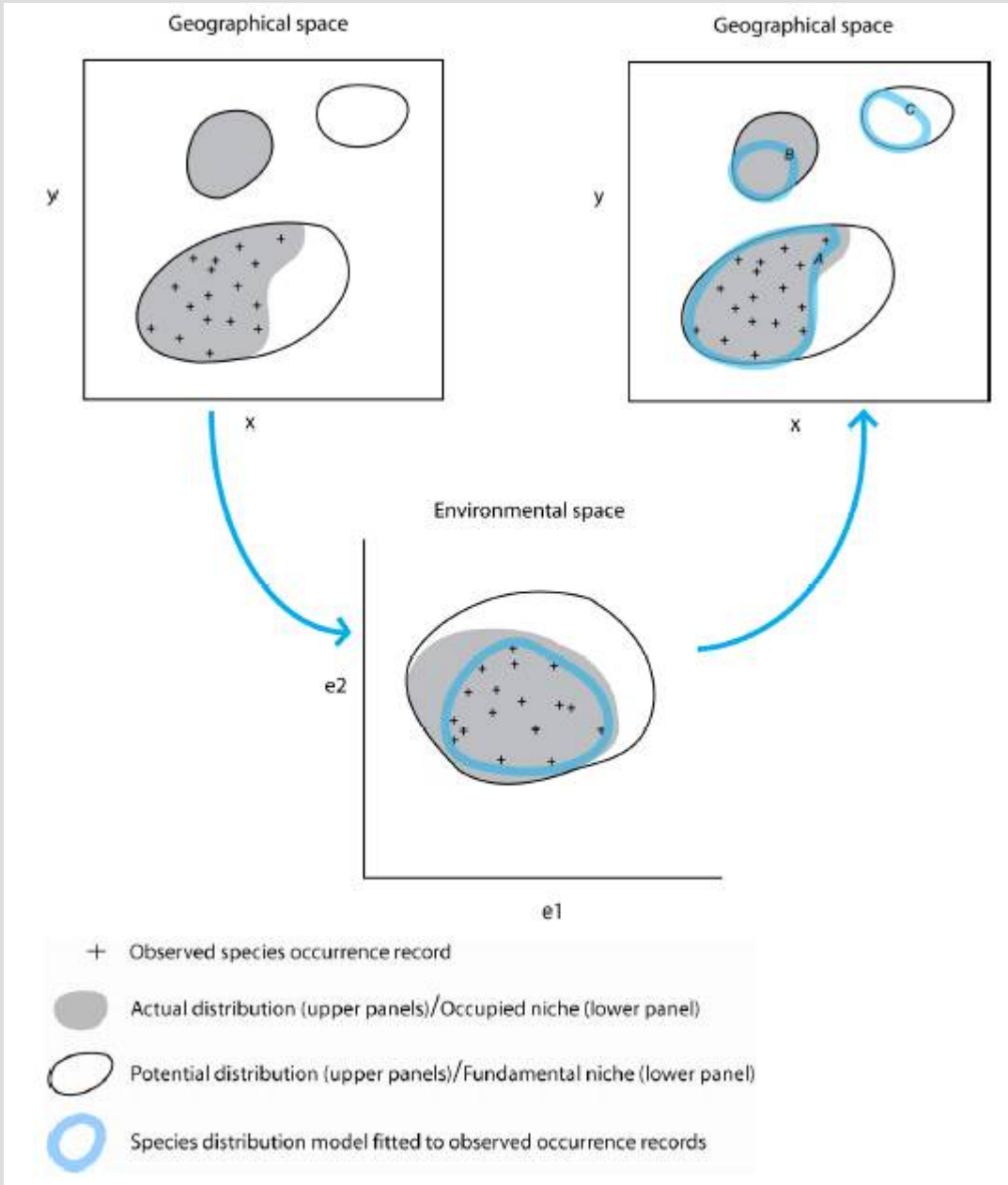
Species Distribution Modeling and GIS

Pat McNally

M.S. Candidate
Field of Soil and Crop Sciences
Environmental Information Science

GIS Day 2009

19 November 2009


Geographical space
Geographical space
y
y
x
x
B
C
A
Environmental space
e2
e1
Observed species occurrence record
Actual distribution (upper panels)/Occupied niche (lower panel)
Potential distribution (upper panels)/Fundamental niche (lower panel)
Species distribution model fitted to observed occurrence records

Species Distribution Models

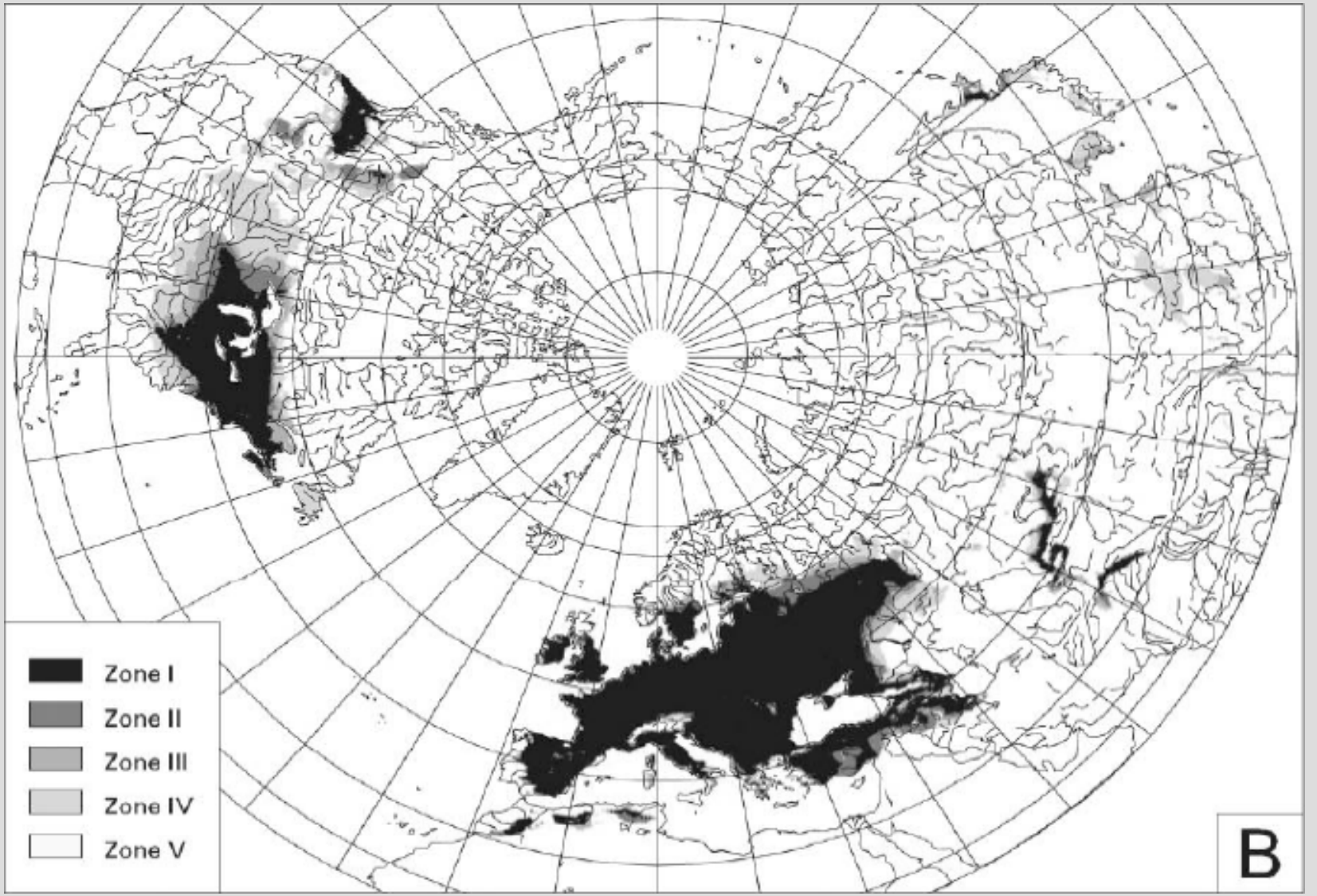
Models	Class of model	Data
• ***BIOCLIM***	envelope model	p
DOMAIN	multivariate distance	p
LIVES	multivariate distance	p
GLM	regression; generalized linear model	p/pa
GAM	regression; generalized additive model	p/pa
MARS	multivariate adaptive regression splines	p/pa
DT-GARP OM-GARP	rule sets derived with genetic algorithms	p/pa
• ***BRT***	boosted regression trees	p/pa
• ***MAXENT***	maximum entropy	p/b
MARS COMM	as for MARS; uses community data	p/c/b
GDM	generalized dissimilarity modeling; uses community data	p/c/b

Concerns with Species of Interest:
Garlic Mustard

- Ecological threat to an endangered species
- Wide range of potential distribution
- Invades healthy, intact forest ecosystems without a disturbance

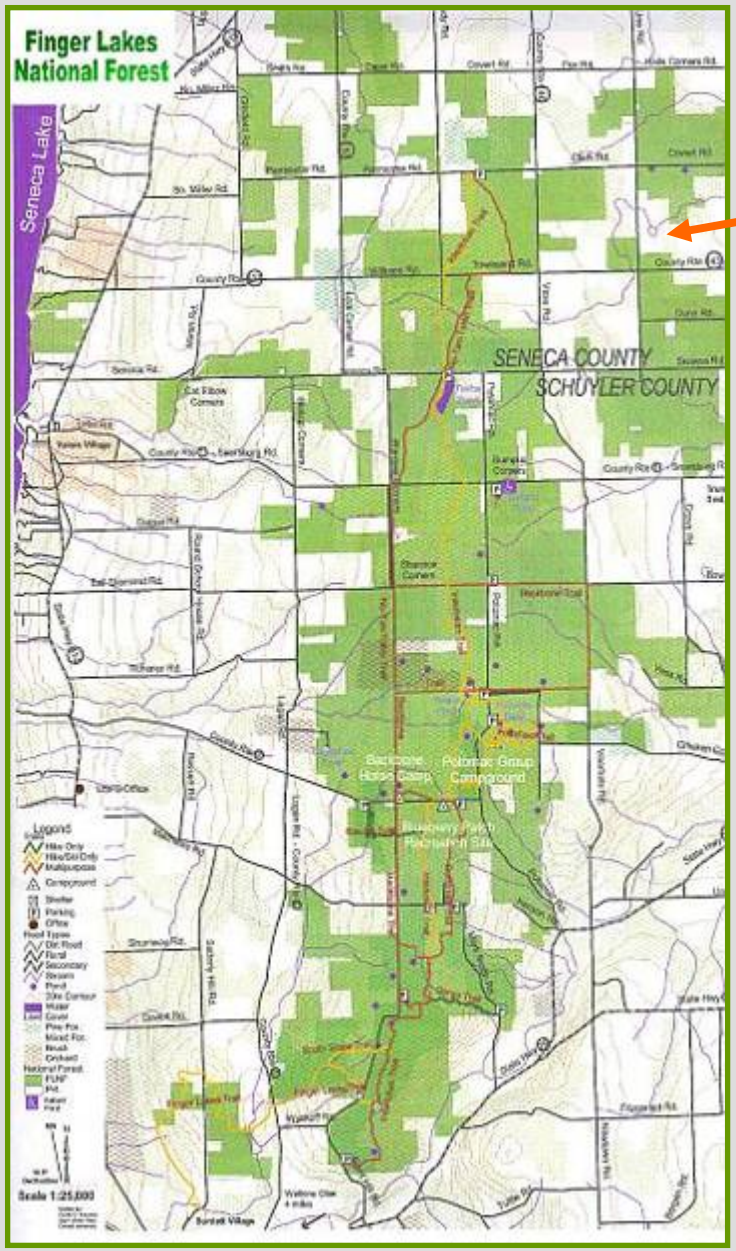
Current modeling approaches are global in nature

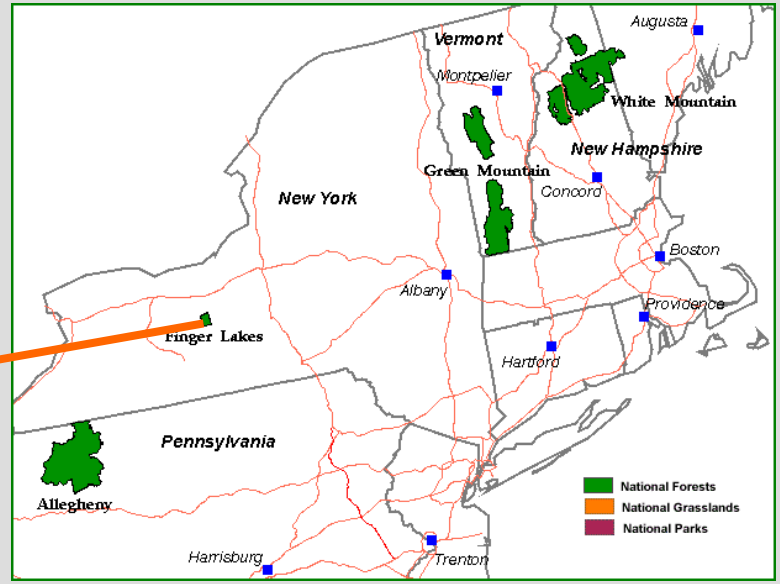
(need local/regional approaches for effective management)

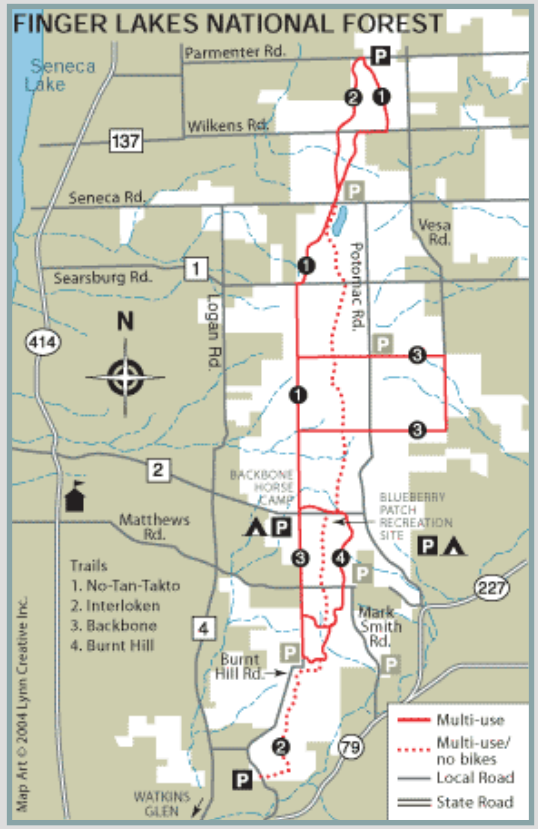


Climatically modelled range and long-term probability zones for invasion in North America.
(Welk et. al. 2002)

Study Area


Finger Lakes
National Forest
Seneca Lake
SENECA COUNTY
SCHUYLER COUNTY
Backbone Horse Camp
Potomac Group Campground
Blueberry Patch Recreation Site
Legend
Scale 1:25,000


Vermont
Montpelier
Augusta
White Mountain
New Hampshire
Green Mountain
Concord
New York
Boston
Albany
Providence
Finger Lakes
Hartford
Pennsylvania
Allegheny
Harrisburg
Trenton
National Forests
National Grasslands
National Parks

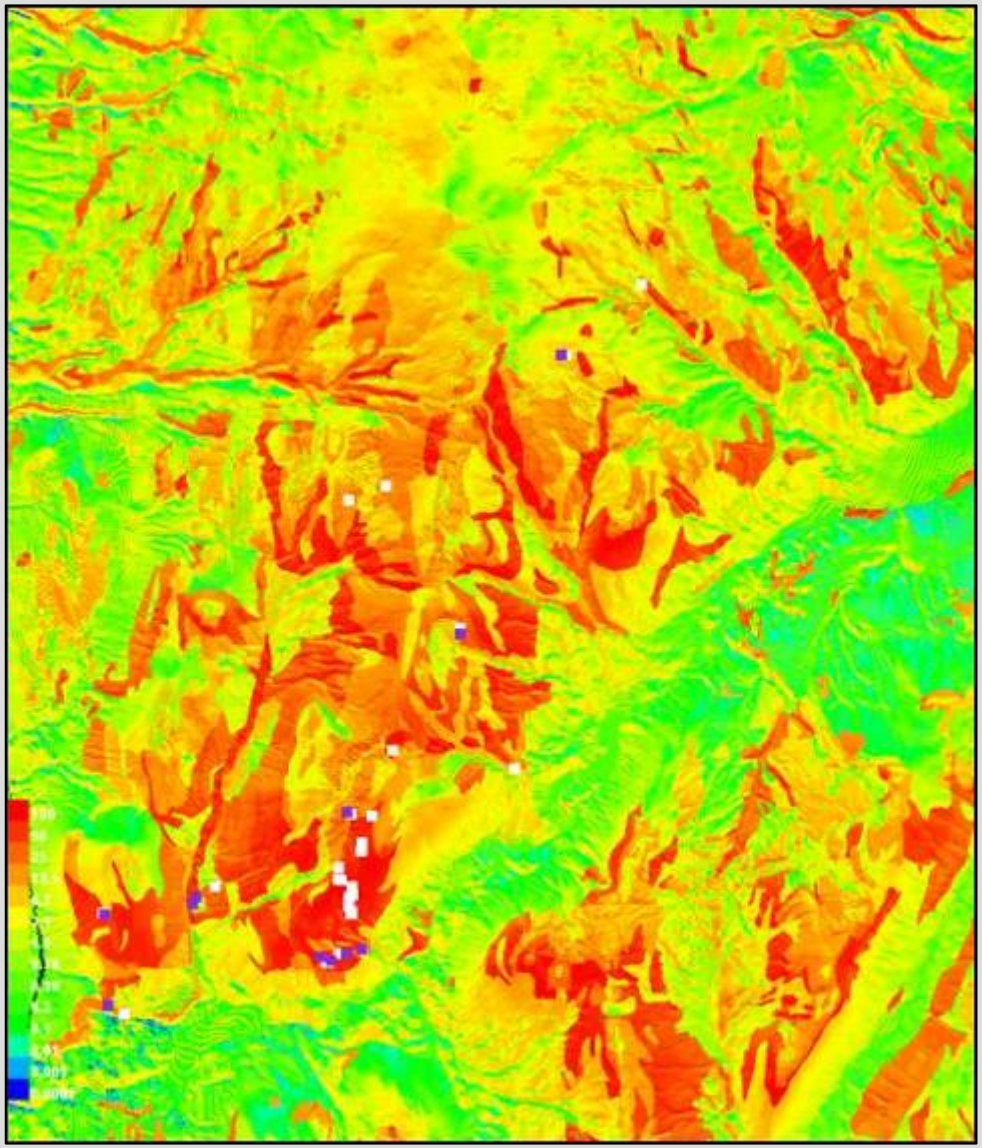

FINGER LAKES NATIONAL FOREST
Seneca Lake
Parmenter Rd.
Wilkens Rd.
137
Seneca Rd.
Vesa Rd.
Searsburg Rd.
Potomac Rd.
Logan Rd.
414
N
2
Matthews Rd.
BACKBONE HORSE CAMP
BLUEBERRY PATCH RECREATION SITE
227
Mark Smith Rd.
Burnt Hill Rd.
4
79
WATKINS GLEN
Trails
1. No-Tan-Takto
2. Interloken
3. Backbone
4. Burnt Hill
Map Art © 2004 Lynn Creative Inc.
Multi-use
Multi-use/ no bikes
Local Road
State Road

Predictor Variables (MaxEnt)

Infestation Susceptibility:

f (soil type, elevation, slope aspect, slope gradient)

MaxEnt Model Results



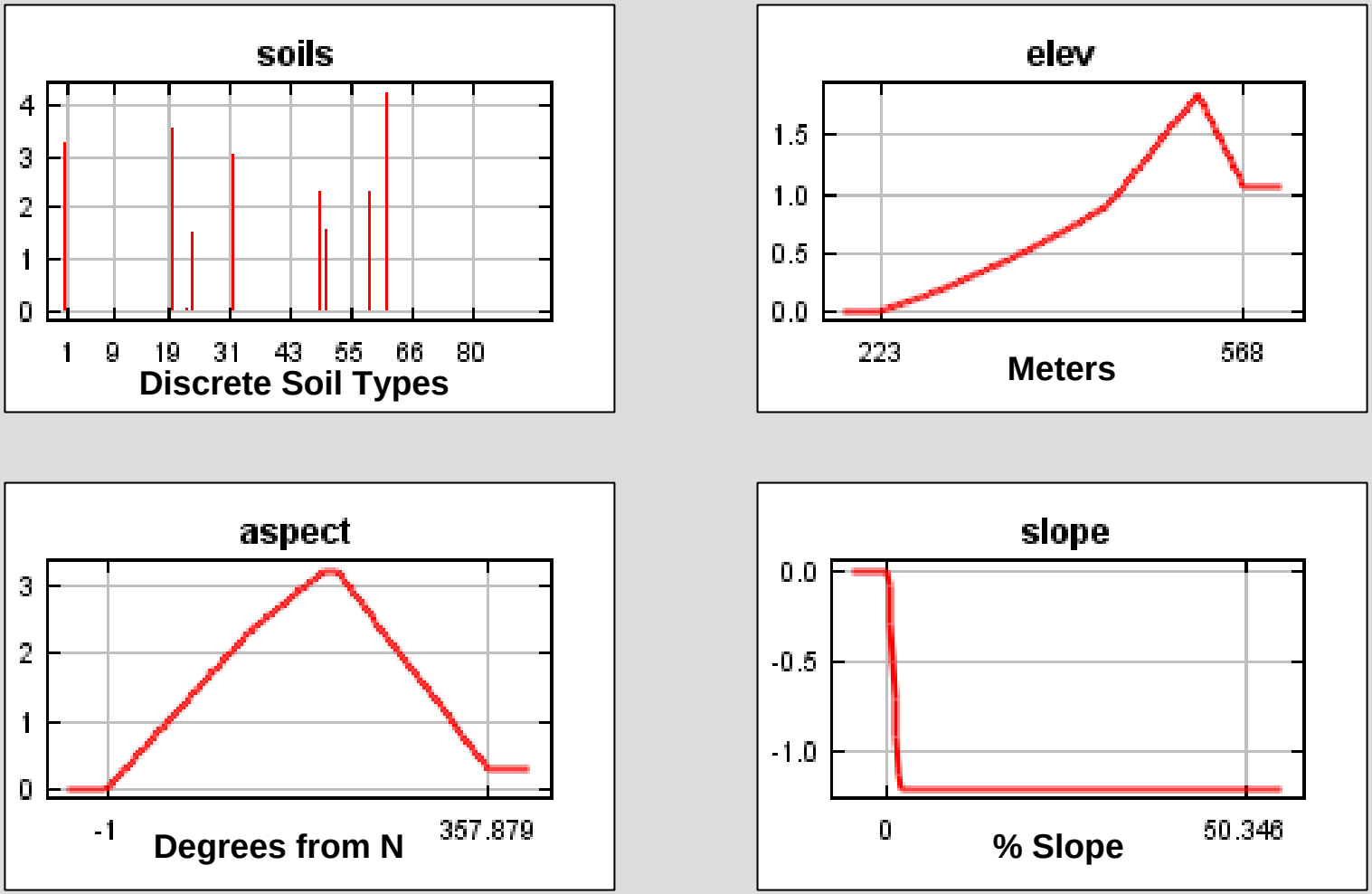
Probability surface for *Alliaria petiolata* infestation

Warmer colors show areas with higher probability of infestation.

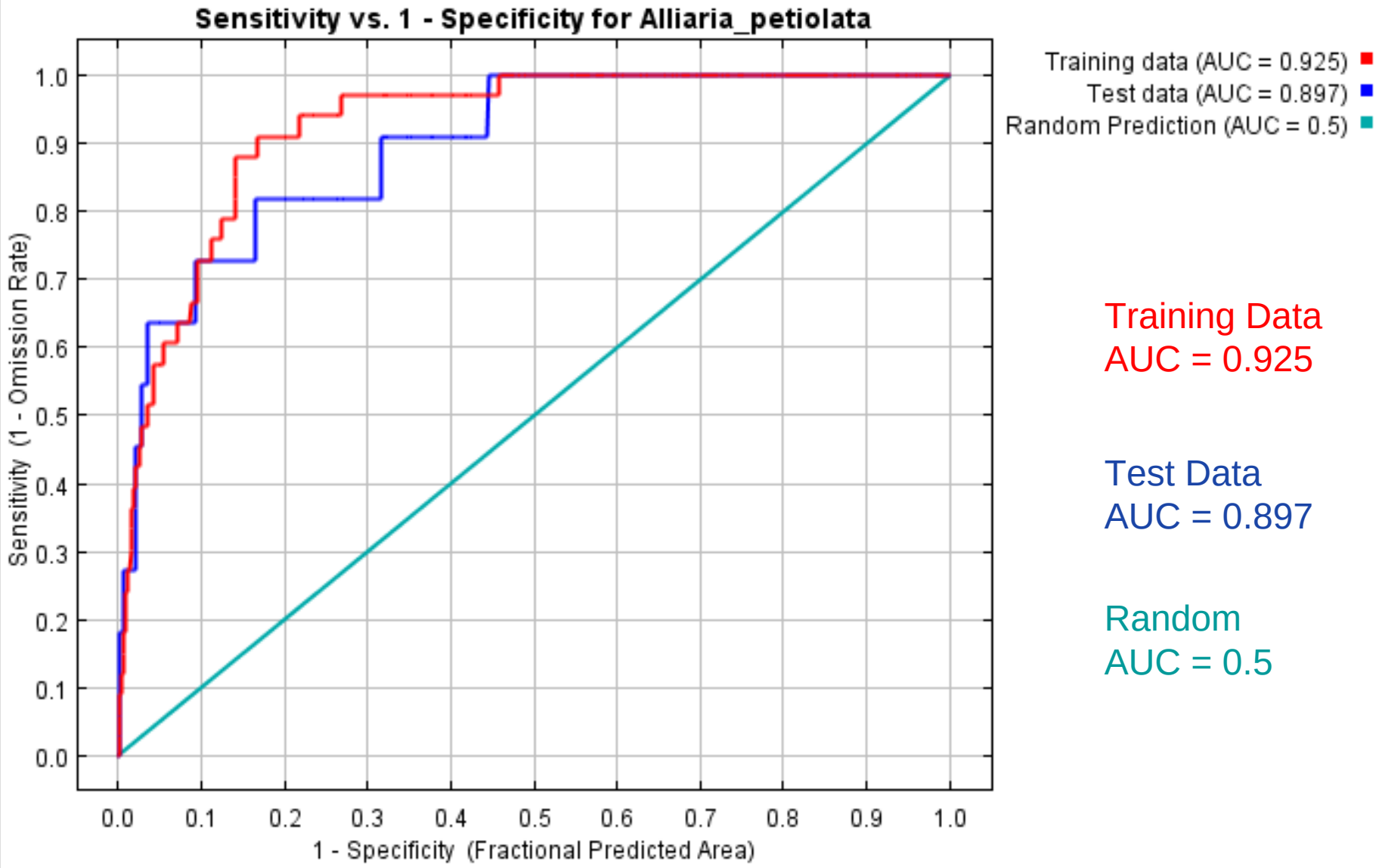
White dots show the presence locations used for model development

Violet dots show model test data point locations.

MaxEnt Response Curves



ROC Curve

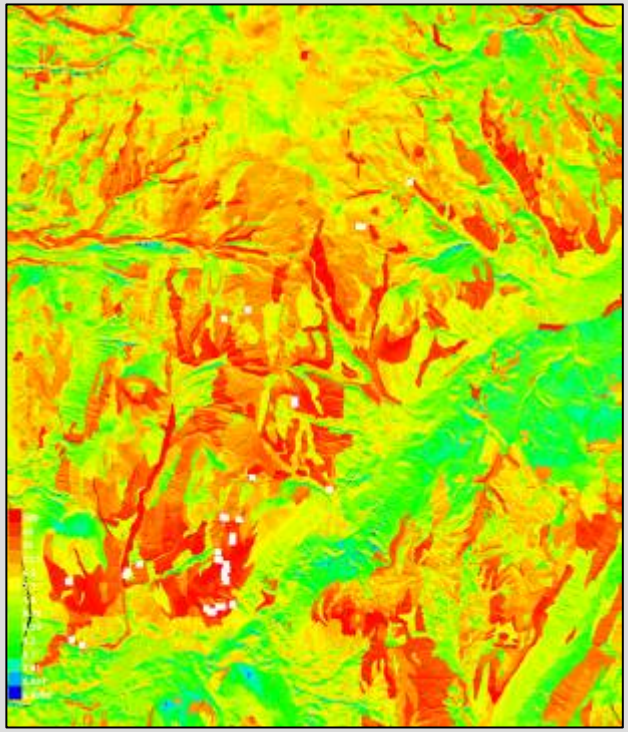


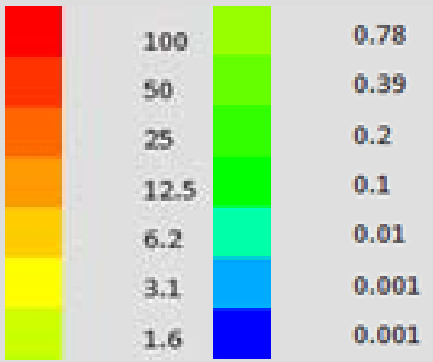
Training Data
AUC = 0.925

Test Data
AUC = 0.897

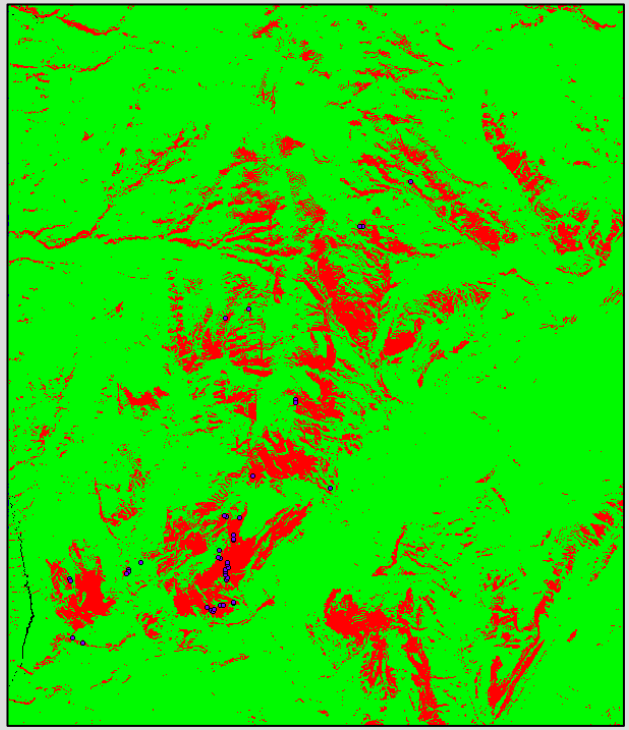
Random
AUC = 0.5

Maxent

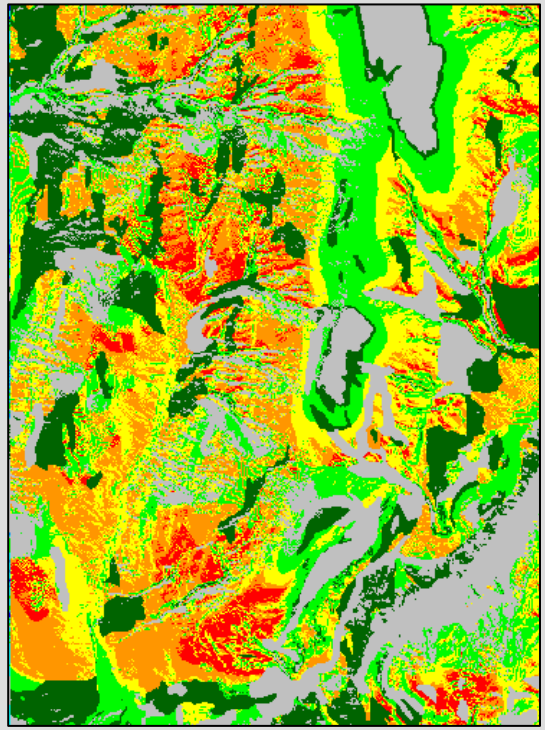


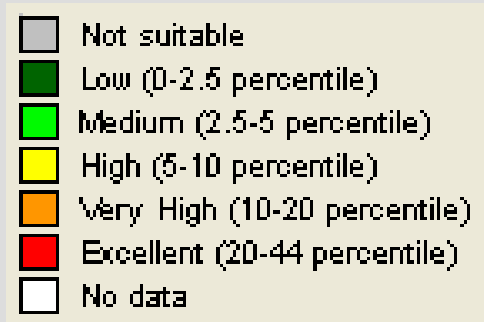


BRT in R



Bioclim Diva GIS





Software

- MaxEnt –
  - http://www.cs.princeton.edu/~schapire/maxent
  - http://groups.google.com/group/Maxent
- Boosted Regression Trees in R http://www.r-project.org
- Diva-GIS – Incorporates distribution modeling tools – http://www.diva-gis.org
- OpenEV – Open Source Image Processing http://openev.sourceforge.net
- openModeller – http://openmodeller.sourceforge.net
- SAGA-GIS – System for automated geoscientific analyses – http://www.saga-gis.org

WorldClim - Global Climate Data

- **http://www.worldclim.org/** – Interpolated global climate surfaces
- **Time Periods**
  - **Current** conditions (interpolations of observed data, representative of 1950-2000)
  - **Future** conditions (downscaled from global climate model (GCM) output, IPPC 3rd assessment; 4th assessment coming soon)
  - **Past** conditions (downscaled global climate model output)
- **Grids - Generic or ESRI**
- **Resolution**
  - **30 arc-seconds (~1 km)**
  - **2.5 arc-minutes (~20 sq km @ equator)**
  - **5 arc-minutes**
  - **10 arc-minutes**

http://biodiversityinformatics.amnh.org/content.php?content_id=140

Biodiversity Informatics Facility

Species Distribution Modeling

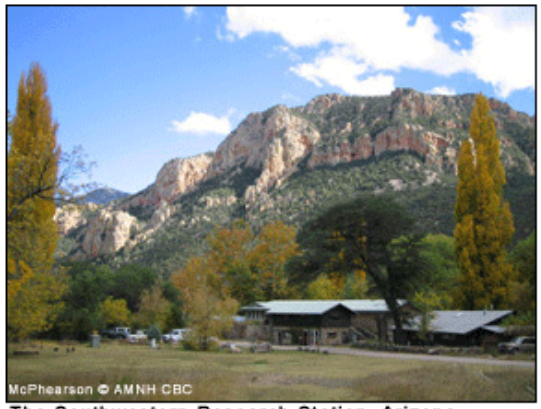
Guide ○ Online Resources ○ Training ○ Search

Training

Species Distribution Modeling > Training

AMNH Training Workshop: Species Distribution Modeling

COURSE DATES: March 29 - April 2, 2010



The Southwestern Research Station, Arizona.

The American Museum of Natural History's Center for Biodiversity and Conservation runs a week-long workshop on **Species Distribution Modeling.**

Applications are now being accepted for the training workshop Species Distribution Modeling, to be held at the American Museum of Natural History's Southwestern Research Station, Arizona.

Models that predict species' potential distributions by combining observed occurrence records with digital data layers of environmental variables have great potential for application across a range of biogeographical analyses. Applications include guiding field surveys to detect unknown populations, projecting potential impacts of climate change, predicting species' invasions, supporting reserve planning, and investigating niche evolution. The workshop focuses on the theoretical and practical aspects of this approach (sometimes termed 'ecological niche' or 'bioclimate envelope' modeling) and is designed for students, researchers and practitioners of conservation biology. Using a mixture of lectures, hands-on computer lab applications, discussions, and case studies, participants will learn to:

- Obtain and process data necessary for species distribution modeling;
- Run distribution models using a variety of approaches;
- Validate and interpret model results;
- Apply these techniques to a range of applications.