Quiz

True of False?

1. The fundamental niche is a subset of the realized niche
2. The realized niche is a subset of the fundamental niche
3. The fundamental niche is a subset of the realized environments
4. The realized niche is a subset of the realized environments

Niche Theory

John M. Drake

Odum School of Ecology
University of Georgia
Athens, Georgia USA 30602-2202

jdrake@uga.edu

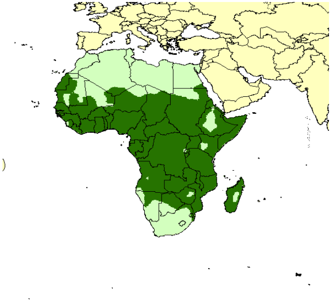
Table of Contents

Main question

If species X were to be introduced to place Y, would it persist?

Applications

- Biogeographic consequences of climate change
- Invasive species management and adaptation
- Survey design
- Species translocation for conservation
- Reserve design



Anopheles gambiae distribution in Africa

Related questions

- What is the *potential distribution* of a species?
- What *range of habitats* does a species use?
- Is this species a *specialist* or a *generalist*?
- What other species are *ecologically equivalent*

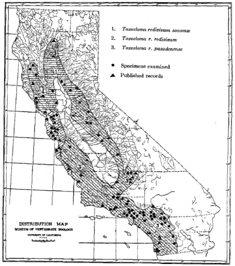
A species persists within its niche

The main question is a *subjunctive conditional*: If species X were to be introduced to place Y, would it persist?

Yes, the species would persist if the environment at Y belonged to the species *niche*, otherwise not.

What is a niche?

The *habitats* in which a species can survive and reproduce



Grinnell, J. (1917). "The niche-relationships of the California Thrasher". Auk 34: 427–433.

What is a niche?

The *set of environments* occupied by a species

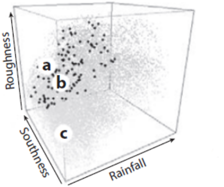


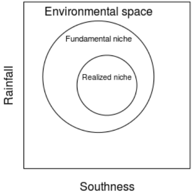
Image: Elith & Leathwick (2009)

- n-dimensional hypervolume

Hutchinson, G.E. (1957). Concluding remarks. Cold Spring Harbor Symposia on Quantitative Biology 22 (2): 415–427.

What is a niche?

The *set of environments* occupied by a species



- Fundamental niche - the *full set of environments* under which a species can persist (Grinnell's concept)
- Realized niche - the *actual set of environments* within which a species is found

Hutchinson, G.E. (1957). Concluding remarks. Cold Spring Harbor Symposia on Quantitative Biology 22 (2): 415–427.

Two kinds of space

Grinnell's concept is a *geographic space*, whereas Hutchinson's concept is an *ecological space*

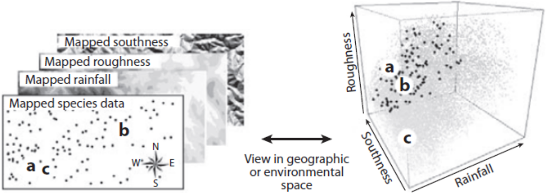


Image: Elith & Leathwick (2009)

Niche: A formal definition

Q. When does a species persist?
A. When growth rate exceeds replacement, i.e., $\lambda > 1$
Q. What does λ depend on?
A. Population size (z) and environment $x = [x_1, x_2, x_3, ...x_n] \in E$
Formal definition $N \in \mathbb{R}^n : x_n \exists z \Leftrightarrow \lambda(x, z) > 1$

What is a niche model?

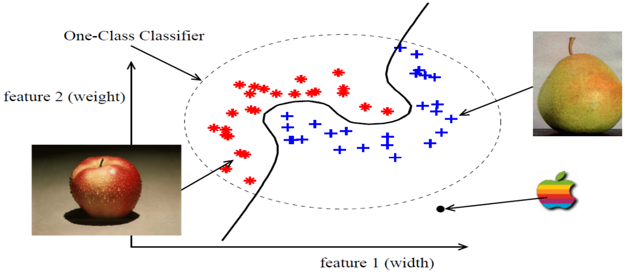
A *niche model* is a representation that takes an environment (a vector of measurements) as input and returns a binary quantityaccording to whether that environment belongs to the niche.
A niche model could be:

- An equation
- A computer program

This suggests that a strategy for modeling ecological niches is to estimate a *decision boundary* between niche and non-niche regions of environmental space.

Two approaches to niche modeling as a decision boundary

- Niche modeling as a *binary classification* problem
- Niche modeling as *support estimation* or *one-class classification*

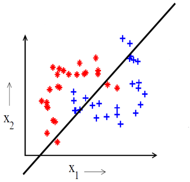


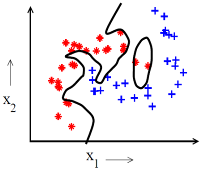
Tax, D.M.J. (2001). One-class classification. PhD thesis.

Niche modeling as binary classification

Benefit: Allows use of classical and modern classification algorithms

- Logistic regression or discriminant analysis
- Neural networks, support vector machines, boosted regression trees





Tax, D.M.J. (2001). One-class classification. PhD thesis.

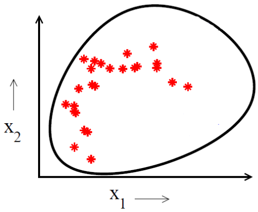
Obstacle: Requires two-class data (ideally in equal proportion)

- Solution 1: "Absence sampling" (but when can one say for certain that a species is not present?)
- Solution 2: "Pseudo-absence sampling" (but this both requires that the translation of geographic-environmental space is fixed and is sensitive to choice of geographic space)
- Solution 3: "Background sampling" (same problems as Solution 2)

Niche modeling as one-class classification

Benefit: Treats the niche problem in the way conceptualized by Hutchinson

- Does not require artificial data
- Infinite number of ways to be "non-niche"



Tax, D.M.J. (2001). One-class classification. PhD thesis.

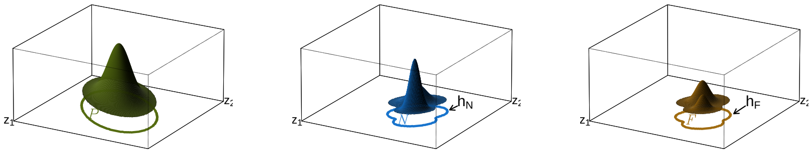
Obstacle: Statistical problem is less familiar and not well defined

- Distribution estimation
- Support vector machines
- Nearest neighbor and other "instance learning" algorithms

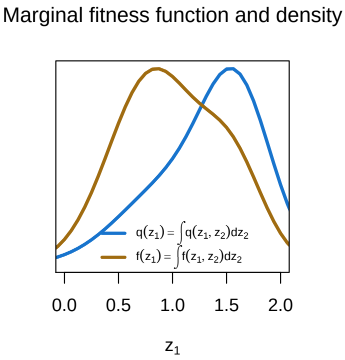
Objective: learning ecological niches from data

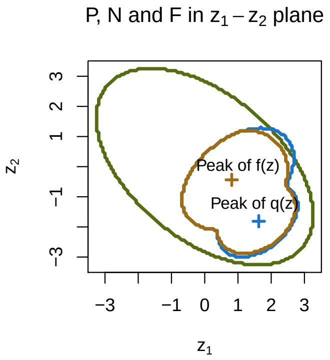
- Modeling from presence-background and presence-only data
- Object of model is ecological niche boundaries (rather than density, mean/variance, etc.)
- Modeling with non-stationary, non-iid data, irrelevant variables, etc...

Probability density of environment p(z) | Habitat selection function q(z)>1 | Density of occupied environments f(z)=p(z)



Modeling niche boundary may be easier than we think



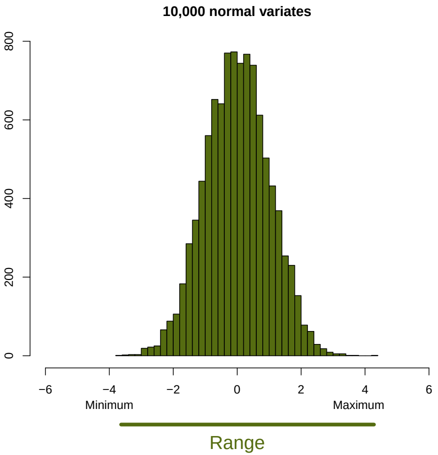


Can we learn the niche boundary from data?

What is a range (one dimension)?

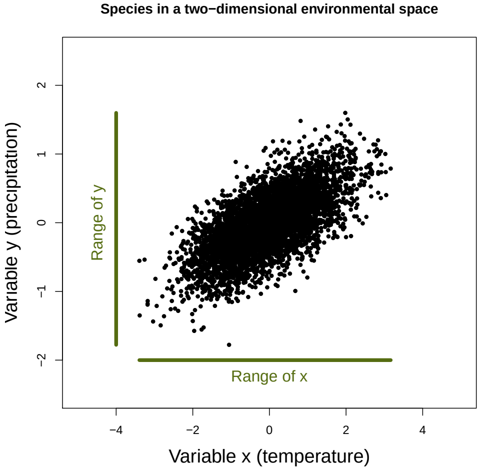
Objectives:

- A presence-only method that is ecologically "interpretable"
- A method that maximally concentrates on estimating species niche boundaries
- A method that is insensitive to contingent species absences
- A method that can be parameterized or partly parameterized with experiments on species tolerances



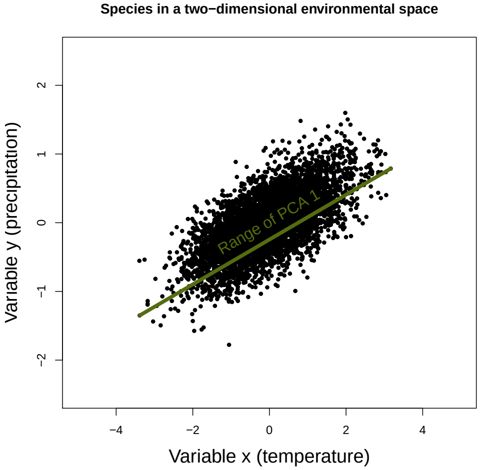
What is a range (two dimensions)?

- In two dimensions there are two ranges...
- In n dimensions there are n ranges



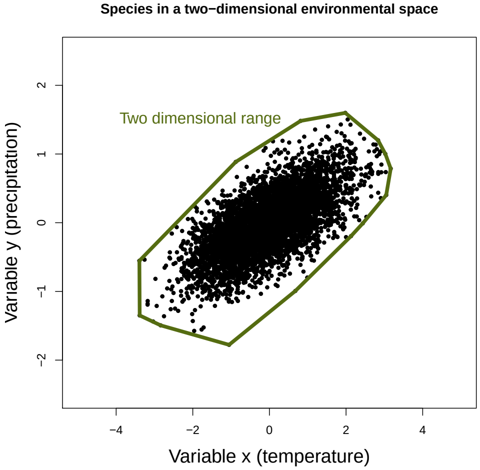
But why this coordinate system?

- The *n* ranges depend on choosing a coordinate system



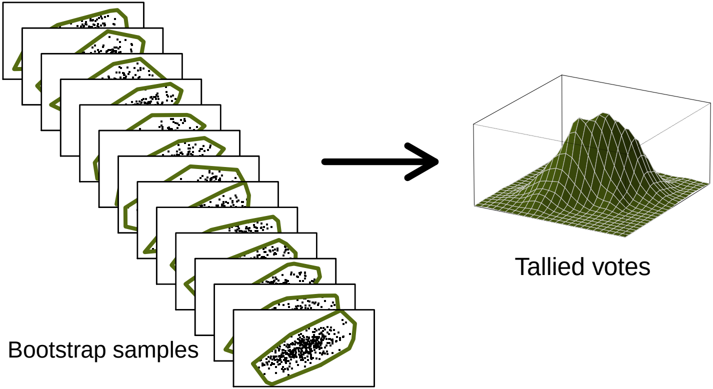
Convex hull

- Convex hull is the boundary of the smallest convex set containing a specified set of points (e.g. environmental conditions at which a species is found)
- Equivalently, the convex hull is the set of endpoints of the ranges in all coordinate systems (rotations)



Range bagging: the proposal

Choose marginal dimension (δ)
Choose boostrap coverage ($w = 1$)
Choose number of bootstrap samples ("votes") (v)



Range bagging: the proposal

Algorithm

Input: $x \times n$ table of environmental records
Tuning variables: δ the dimension of the marginal niche, w boostrap sample size, v iterations (votes)

for $i = 1$ to v, **do**

1. **Sample step.** For a model of order δ, randomly select (without replacement) $\delta < n$ environmental variables. From the resulting $x \times \delta$ table of records, randomly select (without replacement) $w \leq x$ records to be included in the bootstrap sample, X_i.
2. **Marginal niche estimation step.** As a *base learner*, obtain the marginal niche C_i of the points in X_i. If $\delta = 1$, the marginal niche is simply the interval between the minimum and maximum values in the bootstrap sample. For $\delta > 1$, the marginal niche is the convex hull of the bootstrap sample.

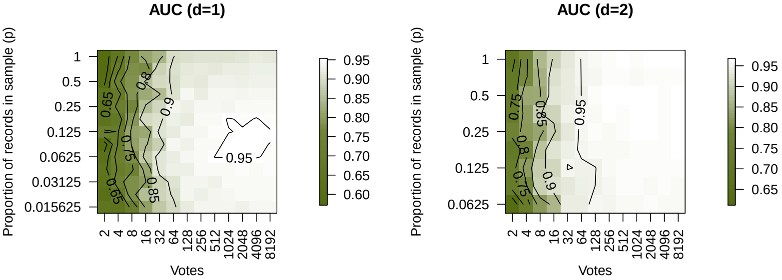
end for

Output: $\mathbb{N} = \{C_1, C_2, C_3, \ldots C_v\}$

To evaluate, calculate the proportion of $\mathbb{N}$ containing a candidate environmental condition.

Range bagging: performance

Performance on two-spined blackfish *Gadopsis bispinosus*



Conclusions:

- Performance comparable to MaxEnt using presence-only data and very little tuning
- Performance does not degrade with the number of votes

Conclusions

- New methods for presence-background and presence-only modeling perform almost as well as existing methods
- Many opportunities for extension to other "data problems" (especially non-stationary data and regional transferability)
- Potential for further use of ensemble methods for variance reduction

References

Drake, J. M., and J. M. Bossenbroek. 2009. Profiling ecosystem vulnerability to invasion by zebra mussels with support vector machines. Theoretical Ecology **2**:189–198.

Drake, J. M., C. Randin, and A. Guisan. 2006. Modelling ecological niches with support vector machines. Journal of Applied Ecology **43**:424–432.

Elith, J., S. J. Phillips, T. Hastie, M. Dudík, Y. E. Chee, and C. J. Yates. 2011. A statistical explanation of MaxEnt for ecologists. Diversity and Distributions **17**:43–57.

Valentini, G., and T. G. Dietterich, 2003. Low bias bagged support vector machines. Pages 752–759 *in* ICML-2003.