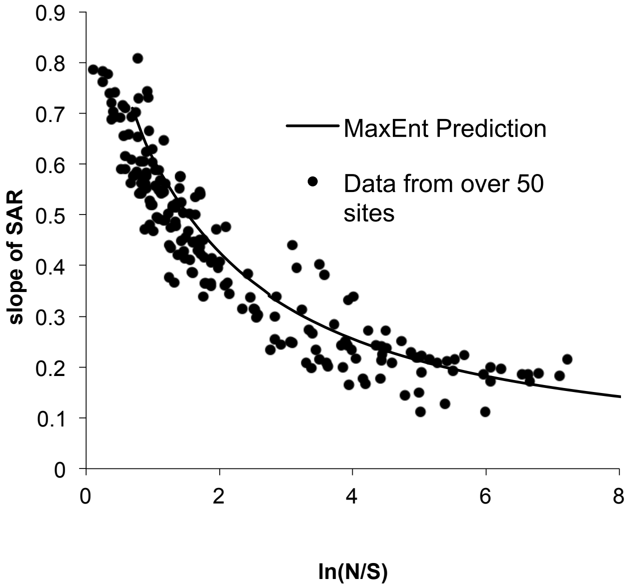
Maximum Entropy:
I. A Framework for Theory Construction
II. An application to Community Ecology

John Harte

Univ. of California
Berkeley

SFI Summer School

June 19,20, 2017



The Complex Systems Dilemma

The example of ecology:

- Vast list of mechanisms & traits:

  predation, mutualism, competition, dispersal, speciation, birth, death, pollination, cannibalism, migration, numerous biogeochemical processes, body size, speed, phenology, food preferences, rooting depth, mating strategies, coloration, mimicry, temperature/drought tolerance, nutrient acquisition strategies, SLA, decoying, masting, density dependence, sociality, Allee effect, disease/disease resistance, signalling … etc.

- Multiple abiotic factors that certainly exert influence
- Stochasticity
- Historical contingency
- Ambiguous system boundaries
- Difficulty conducting large-scale *controlled* experiments
- Difficulty isolating particular stresses acting on the system
- Increasing degradation/loss of objects of study

Basing complex systems models & theory on explicit mechanisms & traits often results in

Arbitrary choices of governing mechanisms/traits

Many adjustable parameters

Models/theory that are not readily falsifiable.

Ways to Predict the Form of Patterns in Complex Systems

1. Models based on explicit statistical assumptions:

   e.g., Random placement model (e.g., Coleman)
   Fractal models (e.g., Kunin; Richards; Green et al.)

2. Models incorporating specific combinations of governing mechanisms:

   e.g., Stochastic demographic Neutral Model (e.g., Hubbell)

   Community Lotka-Volterra-type models (Chesson; many authors)

   GCM's in Climate Science

3. Semi-statistical, Semi-mechanistic Models:

   Climate envelope models used to project species loss under climate change
   e.g., Margaret Davis, Thomas *et al*.; Loarie et al.

4. New application of something old: Information Theory/MaxEnt approach

   (Harte et al., Ecology 2008; Ecology Letters, 2009; 2015; Trends in Ecol. & Evol. 2014; Harte, 2011)

A General Rationale

for the use of the concept of information entropy in science:

1. In science we generally **begin with prior knowledge** and seek to expand that knowledge.

2. **Knowledge is not absolute, but rather probabilistic** in nature, and thus the expanded knowledge we seek can often be expressed mathematically in the form of presently unknown probability distributions.

3. Our **prior knowledge** can often be expressed **in the form of constraints** on those unknown distributions.

4. We seek **expanded knowledge that is "least biased",** in the sense that the expanded knowledge does not assume anything about the distributions other than the information contained in our prior knowledge.

Information Theory and Maximum Entropy

1940's – 1960's: Claude Shannon, Edwin Jaynes …formal & rigorous theory

The basic ideas go back to Laplace, late 1700's, and were used in the development of the theory of thermodynamics by J. Willard Gibbs ~ 1900.

Suppose you wish to know the form of a probability distribution P(n)?

All you know is some prior information about P(n) in the form of constraints.

For example: you might know the mean value, or the variance, or some combination of moments of the distribution.

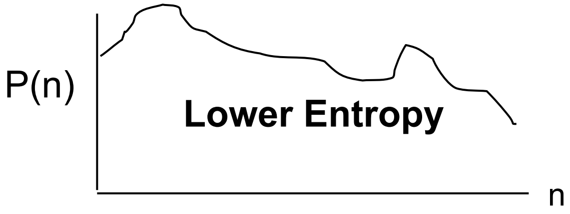
What can you infer is the least biased form of the probability distribution?

The Maximum Information Entropy Concept: Just what is being maximized?

Here “**entropy**” refers to **information entropy**, $I = - \Sigma_n P(n) \log(P(n))$,

Information entropy is a measure of the lack of structure or detail in the probability distribution describing your knowledge.

Maximizing information entropy ↔ finding the smoothest possible probability distribution that is compatible with the constraints that arise from prior knowledge.



If both of these distributions are consistent your prior knowledge, you should prefer the one with higher information entropy.
It makes fewer implicit unwarranted assumptions

Jaynes, E. T., 1982, On the Rationale of Maximum Entropy Methods, *Proc. IEEE.*, 70, 939;

A "derivation" of the main result

Constraints: $\sum_n f_k(n)p(n) = \langle f_k \rangle$ $\sum_n p(n) = 1$ (k = 1,…. K = # constraints)

Maximize: $I_I = -\sum_n p(n)\log(p(n))$ over the function *p*(*n*)

Approach: let $W = -\sum_n p(n)\log(p(n)) - \lambda_0(\sum_n p(n) - 1) - \sum_k \lambda_k(\sum_n f_k(n)p(n) - \langle f_k \rangle)$

(*λ*'s are Lagrange Multipliers)

$$dW/dp = -\log(p(n)) - 1 - \lambda_0 - \sum_k \lambda_k f_k(n) = 0$$

Solution:

$$p(n) = \frac{e^{-\sum_{k=1}^{K}\lambda_k f_k(n)}}{Z(\lambda_1, \lambda_2, \ldots, \lambda_K)}$$

Where:

$$Z = \sum_n e^{-\sum_{k=1}^{K}\lambda_k f_k(n)}$$

$$\frac{\partial \log(Z)}{\partial \lambda_k} = -\langle f_k \rangle$$

Some examples of outcomes of the MaxEnt procedure:

Let the constraints be of the form: $\boldsymbol{F_k = \langle f_k(n) \rangle = \Sigma_n P(n) f_k(n)}$

Constraint function $f(n)$	Form of $P(n)$
n	$e^{-\lambda n}$
mean = $(n_{max}+n_{min})/2$ (e.g., 6-sided fair die)	$1/(n_{max} - n_{min} + 1)$ (e.g., P(n) = 1/6)
n, n^2	Gaussian (normal) distribution
$\log(n)$, $\log^2(n)$	Lognormal distribution
$\log(n)$	$n^{-\lambda}$ (i.e., power law)
Discrete constraints, $P(n_i) = a_{i}$,, can be handled as well	

Some past applications of MaxEnt:

1. Improving image resolution in medicine, forensics (Skilling, …)

2. Inferring values for gaps in economic data such as in input-output tables (George Judge, Amos Golan)

3. Deriving the laws of stat. mech./thermodynamics (Jaynes)

4. Improving estimation of climate envelopes for species (Elith, Phillips) Predicting linkage distributions in networks

5. Predicting linkage distributions in networks

A brief historical detour

Origins of MaxEnt: The Laplace Problem

An urn with red and blue balls in it.

Initially in urn:

🔴 **# = ?**

🟣 **# = ?**

You reach in and pull out a red ball

What is the probability that if you reach in a second time you will get a red ball?

MILESTONES IN THE DEVELOPMENT OF INFORMATION THEORY

Laplace: Rule of Succession & Principle of indifference

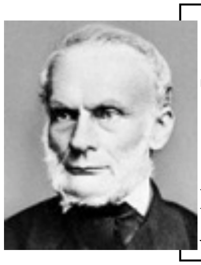
Shannon: Information entropy of a message

Szilard, Landauer: bit of information <=> kT of energy

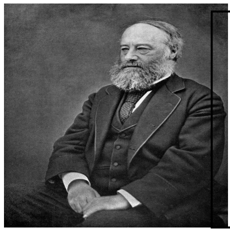
Jaynes: Objective Bayes, MaxEnt

Wheeler: "It from Bit": The suggestion that the laws of physics can be cast in terms of information….only information is truly fundamental.

Information and Thermodynamics

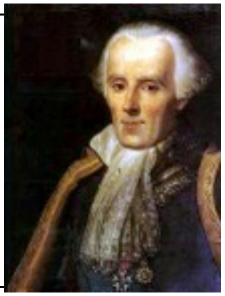


Clausius
The entropy concept; inequivalence of work and heat



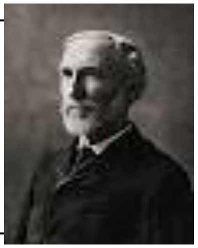
Joule:
Equivalence of work and heat

Laplace:
Probability as description of state of knowledge; principle of indifference, rule of succession



Boltzmann, Gibbs:

Thermodynamics and state likelihood



Shannon

Information entropy of an “alphabet”

***Thermodynamics re-derived from MaxEnt*!**

Jaynes:

Information entropy of a probability distribution; MaxEnt as logic of inference

"Information, in information theory,
is equivalent not to order but to disorder
in the sense that it takes more binary decisions—
more ones and zeros, more bits of information—
to describe disorderly situations of objects than orderly ones."

Schneider and Sagan (2005)

"To the physicist, entropy is a measure of uncertainty about the state of a physical system: one state among all the possible states it can be. These microstates may not be equally likely,
so the physicist writes the formula $S = -\Sigma p\log(p)$.

To the information theorist, entropy is a measure of uncertainty about a message: one message among all the possible messages that a communications source can produce. The possible messages may not be equally likely,
so Shannon writes the formula $H = -\Sigma p\log(p)$." Gleick (2011)

Maximum Entropy Theory Construction

Top Down

Macroscale: **State variables**:

↓ ***MaxEnt***

Microscale: **Probability distributions for:**

Abundances of species
Speakers of languages
Profits of companies
Sizes of organisms
Speeds of molecules
Linkages per node
ETC. ETC.

MaxEnt and the State Variable Concept

In Thermodynamics, these **macro-scale state variables** characterize the system:

P: pressure

V: volume

T: temperature

n: number of moles

From the constraints imposed by knowledge of the values of these state variables MaxEnt predicts:

Boltzmann distribution of energy levels, entropy law, equipartition, binomial distribution of molecules in space, ETC. (Jaynes 1957a, b)

In contrast, the Boltzmann derivation of thermodynamics

(maximize S = klog(W))

is a bottom-up approach.

A Speculation about Complexity:

Complex Systems vs. "Hyper-Complex Systems"

The former: Systems that can be decomposed into a macro- and a micro-level.

E.G.

Thermodynamics (bulk properties such as pressure vs. molecular kinetic energies)

Economics (firms, nations, total wealth vs. individual incomes)

Linguistics (languages vs. speakers)

Ecology (bulk properties such as species richness, # individuals, total metabolism vs. distributions of individuals among species or metabolism among individuals

These are the type systems for which MaxEnt has a natural application.

The latter: Complexity that cannot be decomposed in that way. Complexity across an effective continuum of scales.

E.G., Turbulence

We do not know how to apply MaxEnt to systems of the latter type,
which renders them hyper-complex!

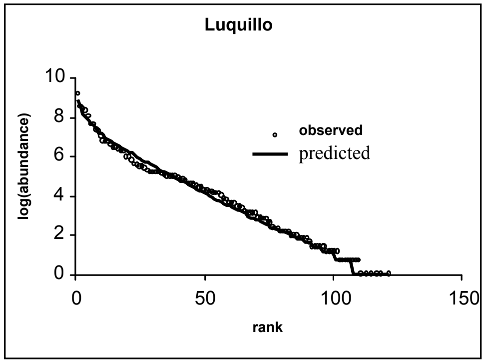
“It is interesting to contemplate
an entangled bank,
clothed with many plants of
many kinds,
with birds singing on the bushes,
with various insects flitting
about,
and with worms crawling
through the damp earth,
and to reflect that these
elaborately constructed forms,
so different in each other,
and dependent on each other in
so complex a manner,
have all been produced by
laws acting around us.”

Charles Darwin
concluding paragraph of Origin
of the Species

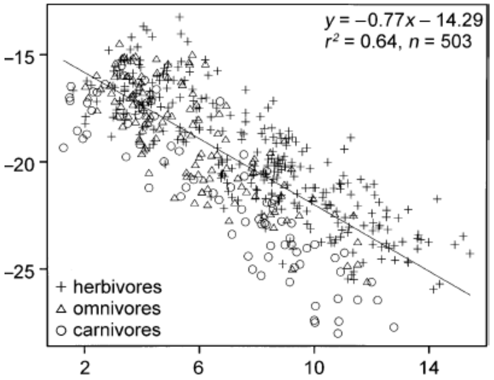
Macroecology: Patterns in the abundance, spatial distribution, & energetics of species…

Of special interest: patterns
that are widely observed
across habitats, taxa, & spatial scales

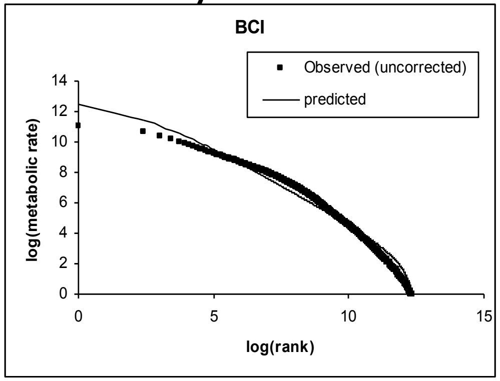
Species-Abundance Distribution



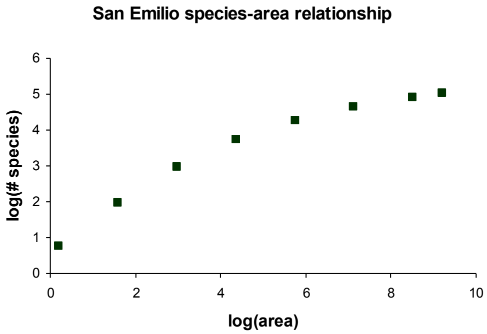
Body Size-Abundance Relationship

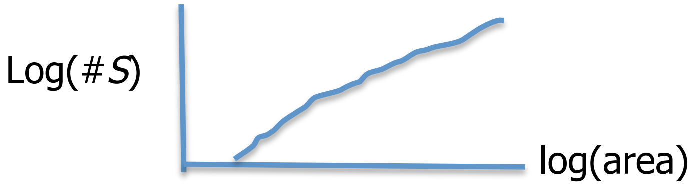
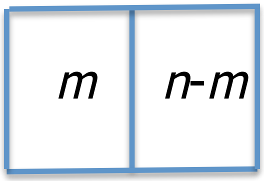
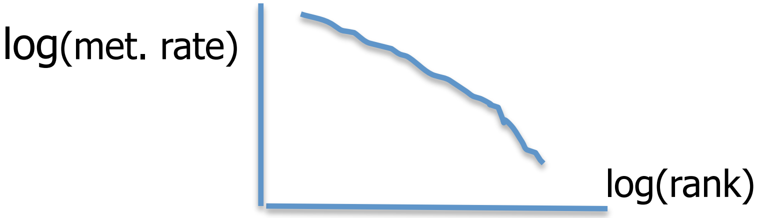
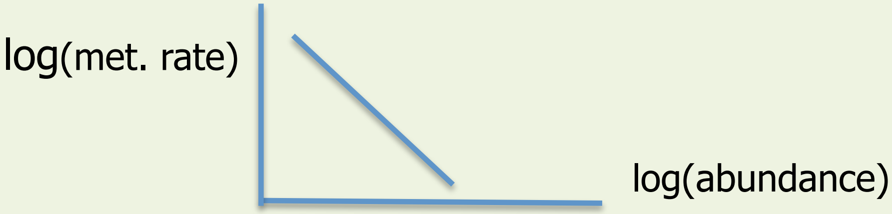


Distribution of Metabolic Rates
or body sizes over Individuals



Species-Area Relationship



Metric	**Prevalent Pattern in Ecology**
Species-Area Relationship	 S increases with area faster than log(A) and slower than A^{z}
Species-Abundance Distribution	 ~ Fisher log-series Distribution (hockey stick)
Spatial-Abundance Distribution: aggregation across a bisection	 All partitions (m) of n individuals in a species are equally likely (Laplace rule of succession)
Distribution of Metabolic Rates over Individuals	 ~ Power-law with exponential tail
Metabolism-Abundance Relationship	 Slope = -1

Species-Level Metrics

Symbol and Name of Metric		Description of Metric
$\Pi(n \mid A, n_0, A_0)$	intra-specific spatial-abundance distribution	probability that n individuals of a species are found in a cell of area A if it has n_0 individuals in A_0.
$B(A \mid n_0, A_0)$	box-counting range-area relationship	dependence on cell size of a box-counting measure of range for a species with n_0 individuals in A_0.
$C(A, D \mid n_0, A_0)$	intra-specific commonality	dependence on A and D of the fraction of pairs of cells of area A, separated by a distance D, that both contain a species with n_0 individuals in A_0.
$\Omega(D \mid n_0, A_0)$	O-ring measure of aggregation	average over each occurrence of an individual, of the density of individuals within a narrow ring at a radius D, divided by the density in the ring expected in a random distribution.
$\Theta(\varepsilon \mid n_{0,} S_0\ N_0\ E_0)$	intra-specific energy distribution	probability density function for an individual from a species with n_0 individuals to have a metabolic energy rate between ε and $\varepsilon + d\varepsilon$.
$\Delta(D \mid n_0, A_0)$	intra-specific dispersal distribution	probability density function for a species with n_0 individuals in A_0 to have a dispersal distance between D and $D + dD$

Community-Level Metrics

Symbol and Name of Metric		Description of Metric
$\Phi(n \vert S_0, N_0, A_0)$	species-abundance distribution (SAD)	probability that in a community with S_0 species and N_0 individuals, a species has abundance n
$S(A \vert S_0, N_0, A_0)$	species-area relationship (SAR)	average number of species in a cell of area A if S_0 species in A_0
$S(N \vert S_0, N_0)$	collector′s curve	average number of species found in a random sample of N individuals
$E(A \vert S_0, N_0, A_0)$	endemics-area relationship (EAR)	average number of species unique to cell of area A if S_0 species in A_0
$X(A, D \vert S_0, N_0, A_0)$	species turnover or “beta diversity”	average fraction of the species in cells of area A that are found in common to two cells of area A a distance D apart, if S_0 species in A_0.
$\Psi(\varepsilon \vert S_0, N_0, E_0)$	community energy distribution	probability for an individual in a community with S_0 species, N_0 individuals, and total metabolic rate, E_0, to have metabolic rate between ε and $\varepsilon + d\varepsilon$
$\Lambda(l \vert S_0, L_0)$	link distribution in a species network	probability that a species in a network with S_0 species and L_0 links is connected by l links to all other species

Why do we care about patterns and metrics in ecology?

1. **Extinction rates under habitat loss.**

   25% of Amazon rain forest has been cut.
   How many species lost?

2. **Scaling up biodiversity.**

   How many species of arboreal beetles
   in the Amazonian rain forest?

http://www.neystadt.org/john/album/

3. **Inferring process from pattern**

   (analogy: *Brahe -> Kepler -> Newton*)

Data → **Pattern** → **Force Law**

The Goal of the Maximum Entropy Theory of Ecology (METE)

To predict the shape of static and dynamic macro-scale patterns in ecology

- **Across taxa:** plants, arthropods, birds,…
- **Across spatial scale:** small patches to large biomes
- **Across habitats:** forests, meadows, deserts, tundra,…

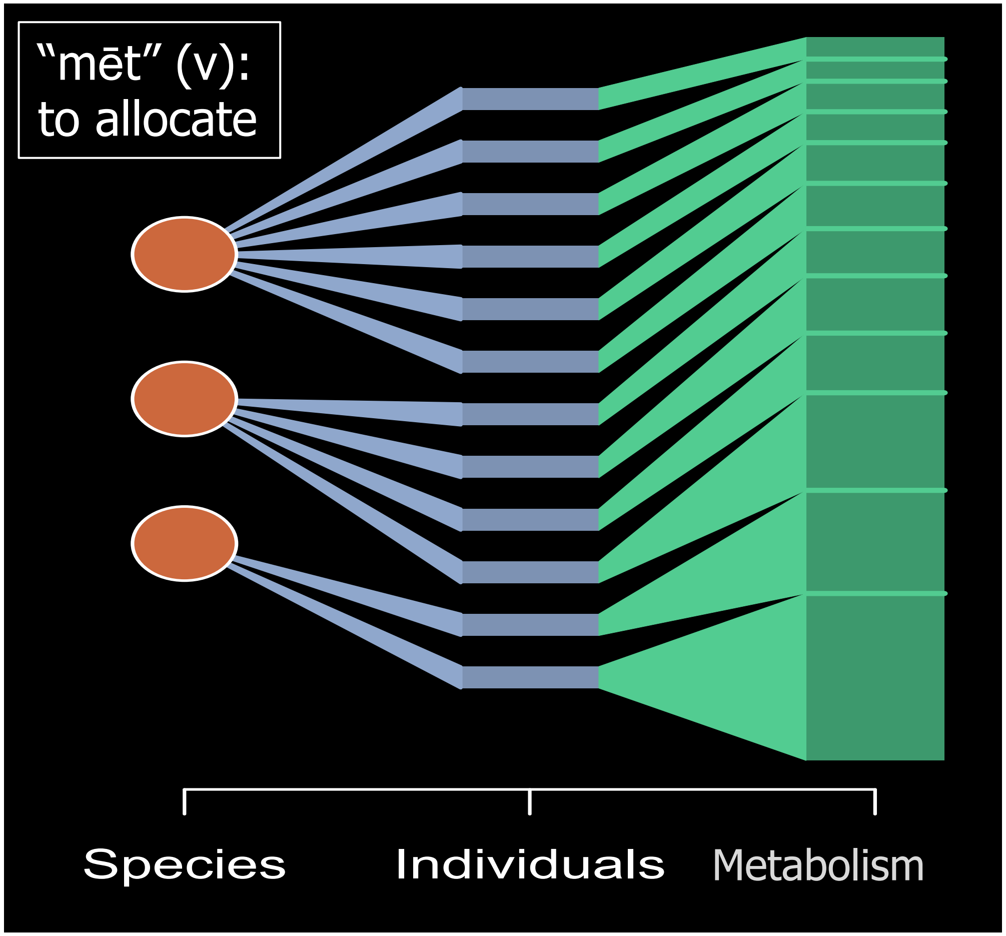
- **without adjustable parameters**
- **without pre-judging what specific mechanisms drive the system**

The ASNE Model of METE

The original ASNE realization of METE starts with the State Variables:
A = Area; S = # species; N = # individuals; E = total Metabolic throughput
And allocates energy to individuals and individuals to species in a least biased way

ASNE predicts the:

- **Species Abundance Distribution**
- **Species-Area Relationship**
- **Endemics-Area Relationship**
- **Spatial Distributions of Individuals**
- **Metabolic Rate Distributions across individuals**



The ASNE model
of the Maximum Entropy Theory of Ecology

Two probability distributions comprise this model:

1. $R(n,\ \varepsilon | S_0, N_0, E_0)$

abundance

metabolic rate

state variables

R is defined over the species and individuals in an area A_0.

a. Choose a species from the species urn.

b. Choose an individual from the individuals urn for that species.

c. *R* is the probability that the chosen species has *n* individuals and that the chosen individual has metabolic rate *ε*.

e.g., S_7

Species urn

e.g.,

S_1 S_7 S_{S_0}

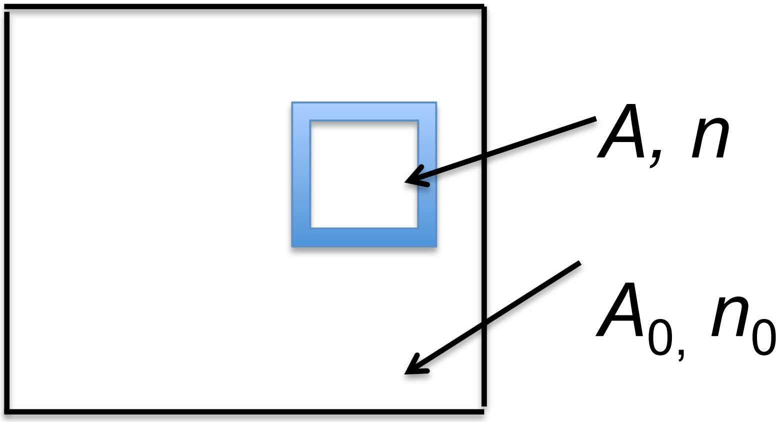
Individuals urns

2. … and a species-level spatial distribution,

$$\Pi(n \text{ in } A | n_0 \text{ in } A_0)$$

describing aggregation of individuals within species:

Π =
probability
that n individuals
in A if n_0 in A_0



From R and Π, most of the metrics of macroecology can be derived.

ASNE model of METE

$$R(n, \varepsilon \mid S_0, N_0, E_0) = \frac{e^{-\lambda_1 n} e^{-\lambda_2 n \varepsilon}}{Z(\lambda_1, \lambda_2)}$$

The Ecosystem Structure Function

$$\Pi(n \mid n_0, A, A_0) = \frac{e^{-\lambda_\Pi n}}{Z_\Pi}$$

Species-level Spatial distribution: probability *n* individuals in a cell of area *A* if $n_0$in A_0

No derivations here: you have to trust me on the math…or do it yourself.

$\beta = \lambda_1 + \lambda_2$

$\beta \ln(1/\beta) = S/N$

$\lambda_2 = S/E$

$$R(n, \varepsilon \mid S_0, N_0, E_0) = \frac{e^{-\lambda_1 n} e^{-\lambda_2 n \varepsilon}}{Z(\lambda_1, \lambda_2)}$$

MaxEnt gives

$$\Pi(n \mid n_0, A, A_0) = \frac{e^{-\lambda_\Pi n}}{Z_\Pi}$$

(Integrating over ε)

$$\boxed{\Phi(n) = \frac{1}{\lambda_2 Z n} e^{-\beta n}}$$

The Fisher log series Species Abundance distribution

There are no adjustable parameters: the state variables uniquely determine this and the other metrics

MaxEnt gives

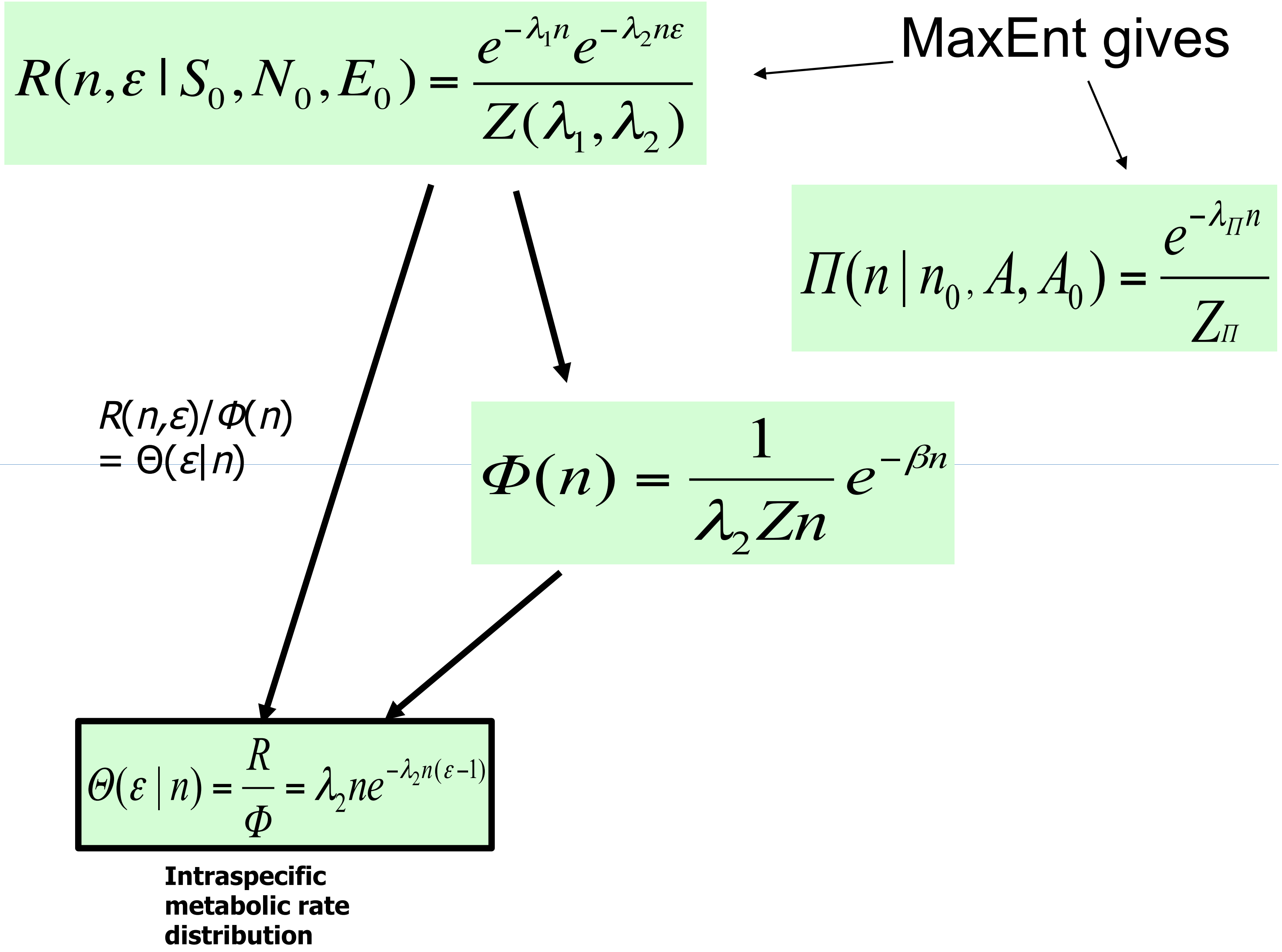
$$R(n, \varepsilon \mid S_0, N_0, E_0) = \frac{e^{-\lambda_1 n} e^{-\lambda_2 n \varepsilon}}{Z(\lambda_1, \lambda_2)}$$

$$\Pi(n \mid n_0, A, A_0) = \frac{e^{-\lambda_\Pi n}}{Z_\Pi}$$

(Multiplying by *nS/N* and summing over *n*)

$$\boxed{\Psi(\varepsilon \mid S_0, N_0, E_0) \approx \lambda_2 \cdot \beta \cdot \frac{e^{-\gamma}}{(1-e^{-\gamma})^2}}$$

Distribution of metabolic rates over individuals:
$\gamma(\varepsilon) = \lambda_1 + \lambda_2(\varepsilon)$



$$R(n,\varepsilon \mid S_0, N_0, E_0) = \frac{e^{-\lambda_1 n} e^{-\lambda_2 n\varepsilon}}{Z(\lambda_1,\lambda_2)}$$

MaxEnt gives

$$\Pi(n \mid n_0, A, A_0) = \frac{e^{-\lambda_\Pi n}}{Z_\Pi}$$

$R(n,\varepsilon)/\Phi(n)$
$= \Theta(\varepsilon|n)$

$$\Phi(n) = \frac{1}{\lambda_2 Z n} e^{-\beta n}$$

$$\boxed{\Theta(\varepsilon \mid n) = \frac{R}{\Phi} = \lambda_2 n e^{-\lambda_2 n(\varepsilon-1)}}$$

Intraspecific
metabolic rate
distribution

$$R(n,\varepsilon \mid S_0, N_0, E_0) = \frac{e^{-\lambda_1 n} e^{-\lambda_2 n \varepsilon}}{Z(\lambda_1, \lambda_2)}$$

MaxEnt gives

$$\Pi(n \mid n_0, A, A_0) = \frac{e^{-\lambda_\Pi n}}{Z_\Pi}$$

$$\Phi(n) = \frac{1}{\lambda_2 Z n} e^{-\beta n}$$

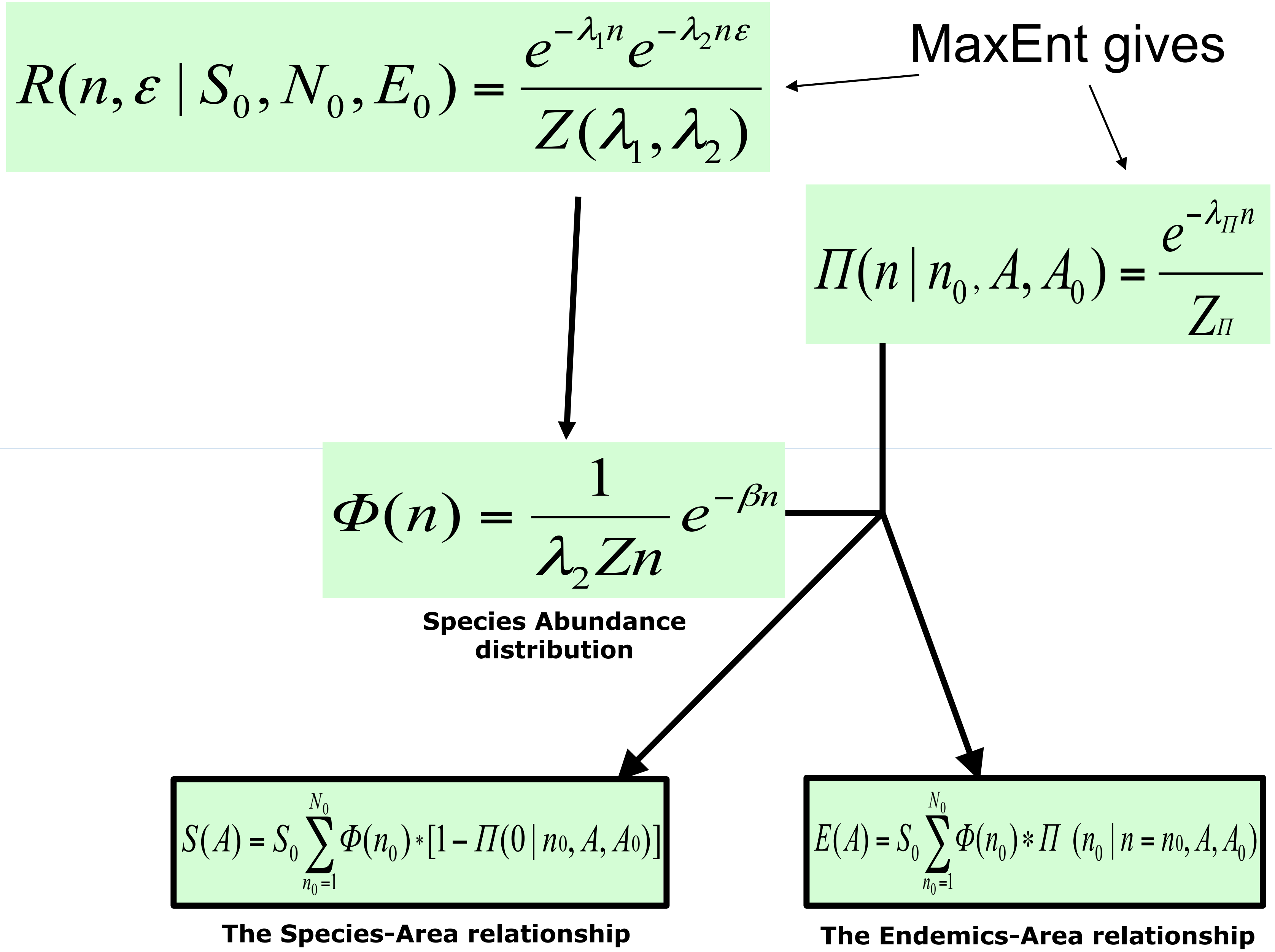
$$\Theta(\varepsilon \mid n) = \frac{R}{\Phi} = \lambda_2 n e^{-\lambda_2 n(\varepsilon - 1)}$$

(Taking the mean)

$$\boxed{< \varepsilon(n) >= 1 + \frac{1}{n\lambda_2}}$$

Abundance-metabolism relation for species

An approximate energy equivalence rule!



MaxEnt gives

$$R(n,\varepsilon \mid S_0, N_0, E_0) = \frac{e^{-\lambda_1 n} e^{-\lambda_2 n \varepsilon}}{Z(\lambda_1, \lambda_2)}$$

$$\Pi(n \mid n_0, A, A_0) = \frac{e^{-\lambda_\Pi n}}{Z_\Pi}$$

$$\Phi(n) = \frac{1}{\lambda_2 Z n} e^{-\beta n}$$

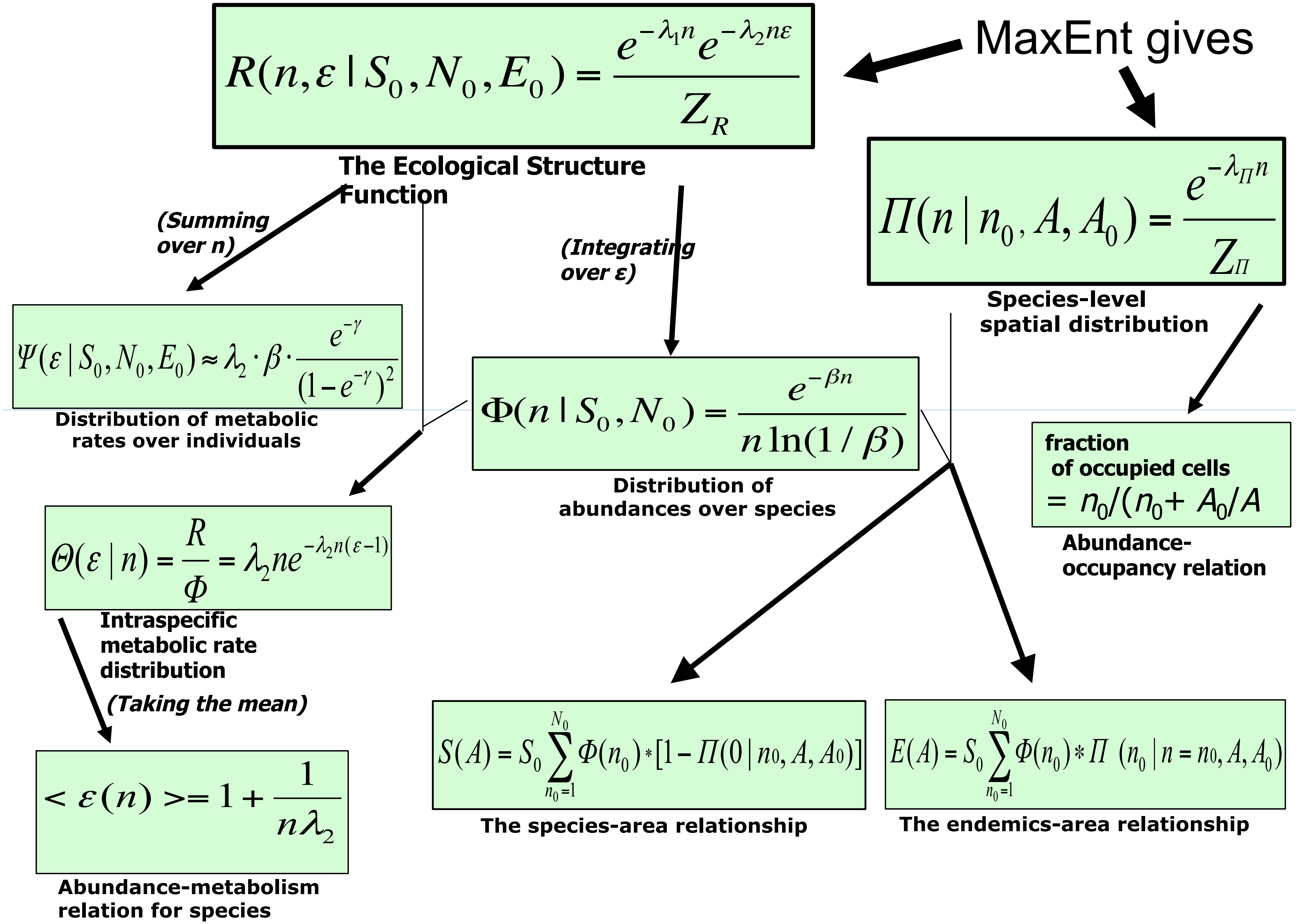
Species Abundance distribution

$$S(A) = S_0 \sum_{n_0=1}^{N_0} \Phi(n_0) * [1 - \Pi(0 \mid n_0, A, A_0)]$$

The Species-Area relationship

$$E(A) = S_0 \sum_{n_0=1}^{N_0} \Phi(n_0) * \Pi\ (n_0 \mid n = n_0, A, A_0)$$

The Endemics-Area relationship



MaxEnt gives

$$R(n,\varepsilon \mid S_0, N_0, E_0) = \frac{e^{-\lambda_1 n} e^{-\lambda_2 n\varepsilon}}{Z_R}$$

The Ecological Structure Function

$$\Pi(n \mid n_0, A, A_0) = \frac{e^{-\lambda_\Pi n}}{Z_\Pi}$$

Species-level spatial distribution

(Summing over n)

$$\Psi(\varepsilon \mid S_0, N_0, E_0) \approx \lambda_2 \cdot \beta \cdot \frac{e^{-\gamma}}{(1-e^{-\gamma})^2}$$

Distribution of metabolic rates over individuals

(Integrating over ε)

$$\Phi(n \mid S_0, N_0) = \frac{e^{-\beta n}}{n \ln(1/\beta)}$$

Distribution of abundances over species

fraction of occupied cells

= $n_0/(n_0 + A_0/A$

Abundance-occupancy relation

$$\Theta(\varepsilon \mid n) = \frac{R}{\Phi} = \lambda_2 n e^{-\lambda_2 n(\varepsilon - 1)}$$

Intraspecific metabolic rate distribution

(Taking the mean)

$$< \varepsilon(n) >= 1 + \frac{1}{n\lambda_2}$$

Abundance-metabolism relation for species

$$S(A) = S_0 \sum_{n_0=1}^{N_0} \Phi(n_0) * [1 - \Pi(0 \mid n_0, A, A_0)]$$

The species-area relationship

$$E(A) = S_0 \sum_{n_0=1}^{N_0} \Phi(n_0) * \Pi\ (n_0 \mid n = n_0, A, A_0)$$

The endemics-area relationship

Many Tests of Predictions

At ~ 25 distinct habitats: ~ 10^5 Species, ~10^9 individuals

36 serpentine meadow plots in CA

11 Smithsonian tropical forest plots

A 9.8 ha dry forest plot at San Emilio, Costa Rica

Plant census in Anza Borrego desert

Breeding bird censuses in southern Africa

Forest floor vegetation

Tree census data from the Western Ghats i. India

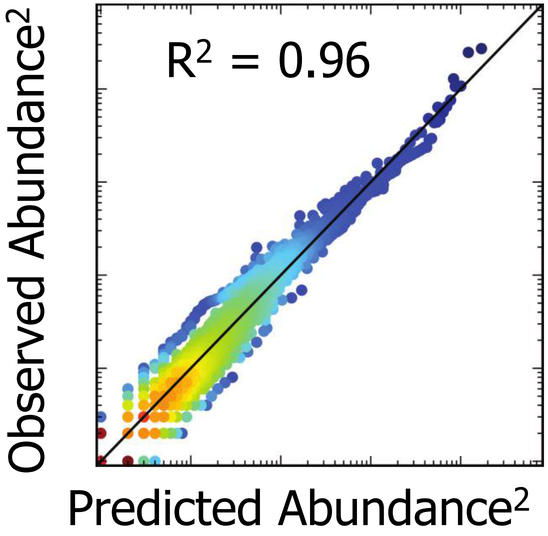
Hawaiian arthropods

Panamanian arthropods

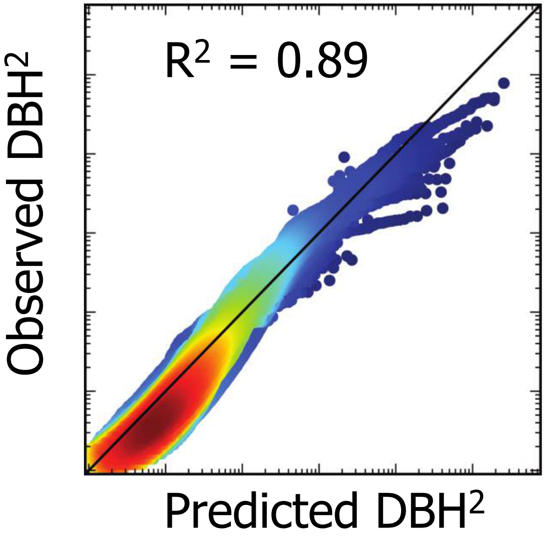
Human gut microbiome

Tests of abundance and metabolic rate distributions

Tests of predicted abundance distribution **for 15,848 plant, mammal, arthropod, and bird communities:** (White et al., 2012).

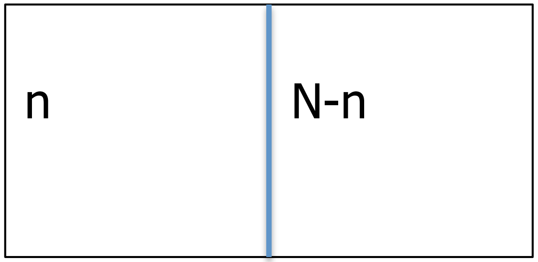


And metabolic rate distribution **for 60 forest communities:** (Xiao et al. 2015)



Bisection Graphs

(distribution of the fraction of individuals in the left side of a plot)



gas molecules: binomial distribution
p(n) sharply peaked at
$n = N/2$

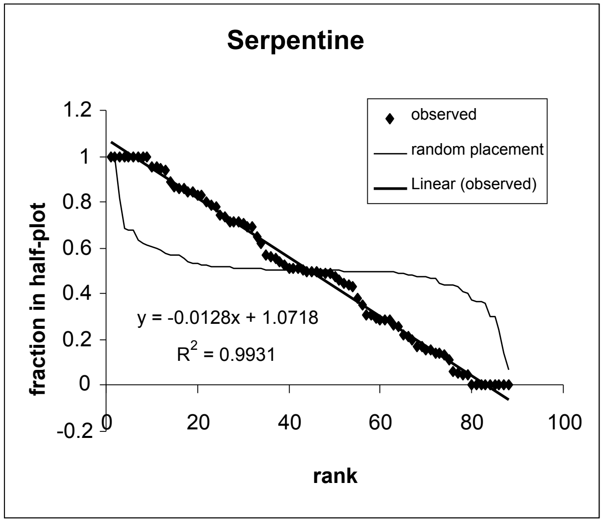
Laplace rule of succession for the assembly process:
If there are L on the left and R on the right, what is the probability the next one is on the left:

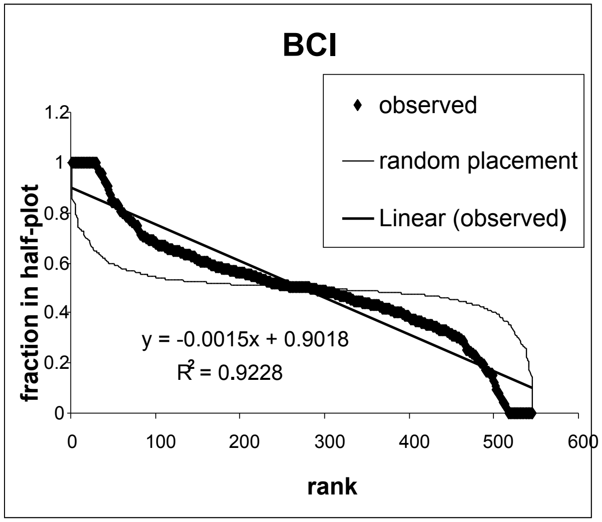
$$p(L) = (L+1)/(L + R + 2)$$

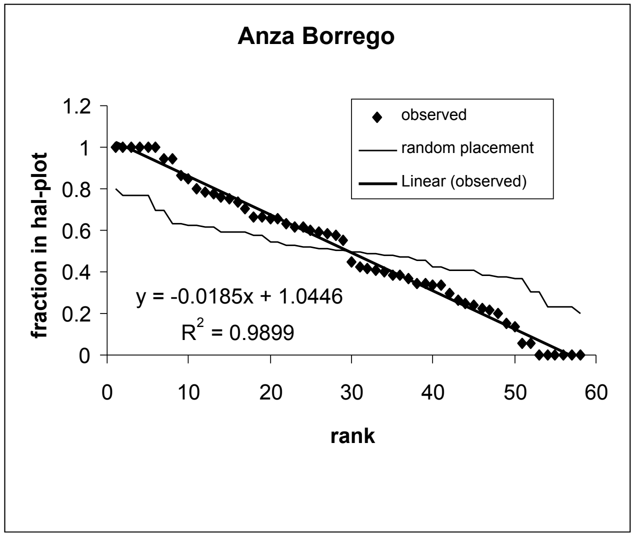
This results in a uniform distribution for p(L)

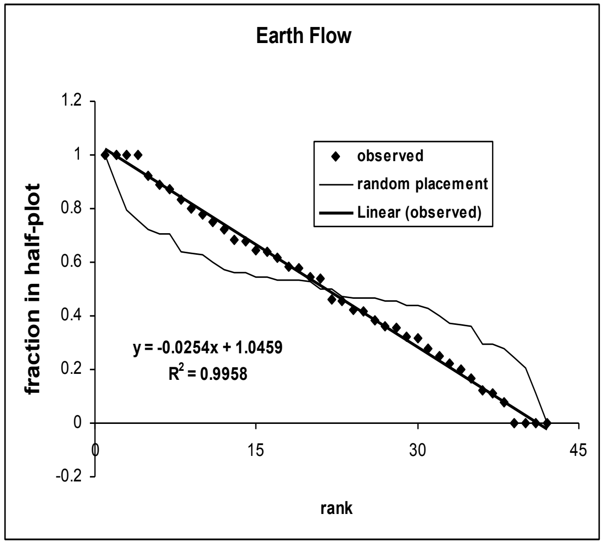
Bisection Graphs

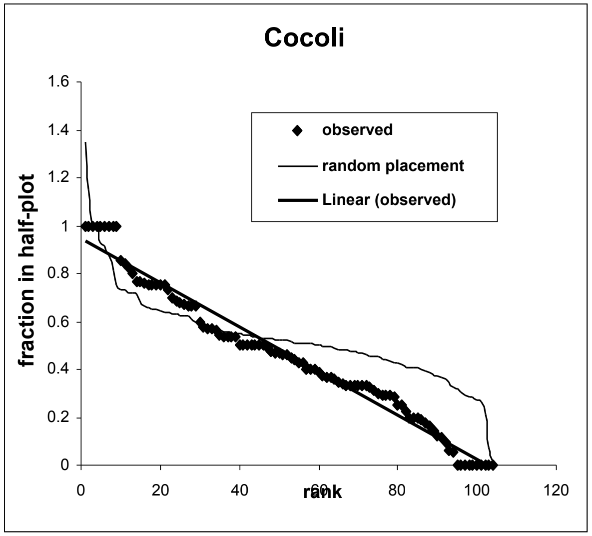
(distribution of the fraction of individuals in the left side of a plot)





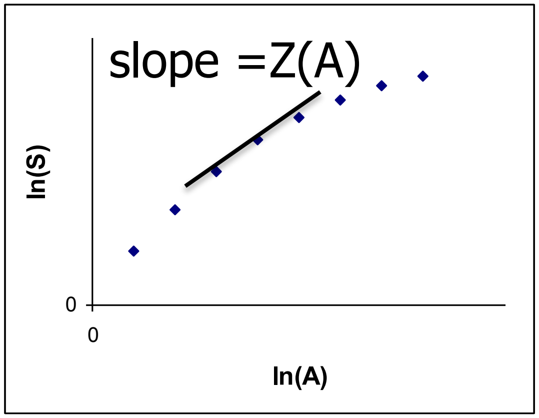






The most unexpected prediction: The Species Area Relationship

The traditional way
SARs are plotted:



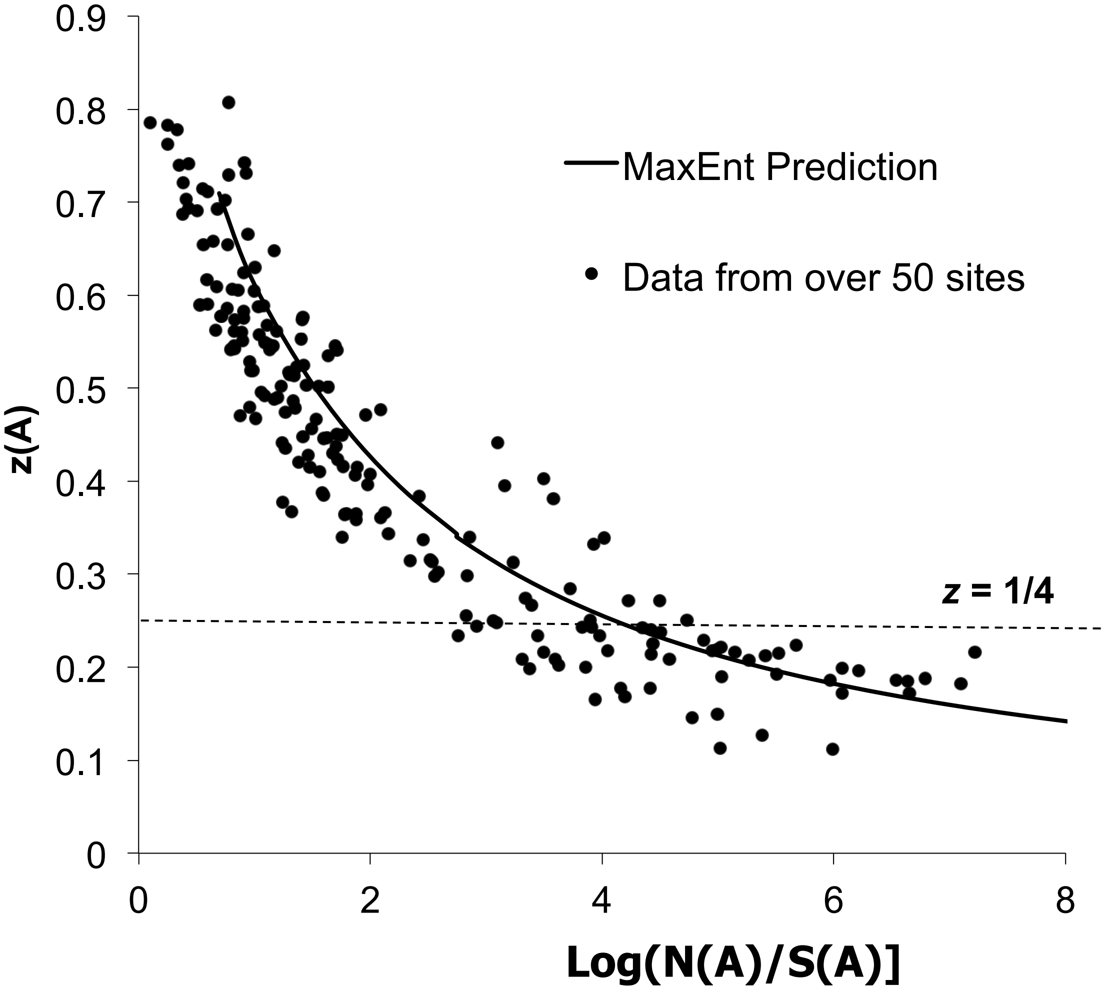
METE predicts:

all species-area curves collapse onto a universal curve

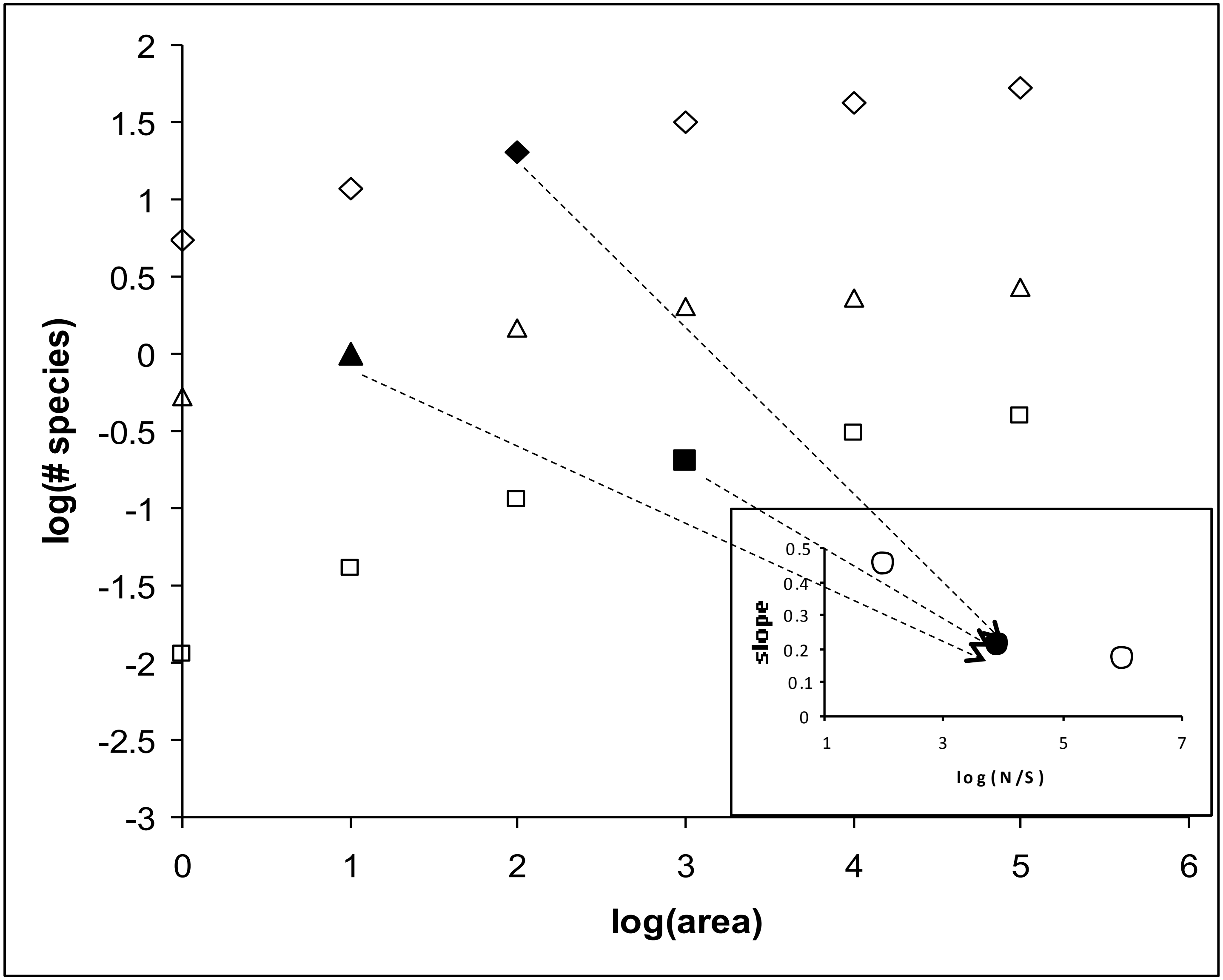
if

**The local slope, Z(A),
Is plotted against the local value of S(A)/N(A)**

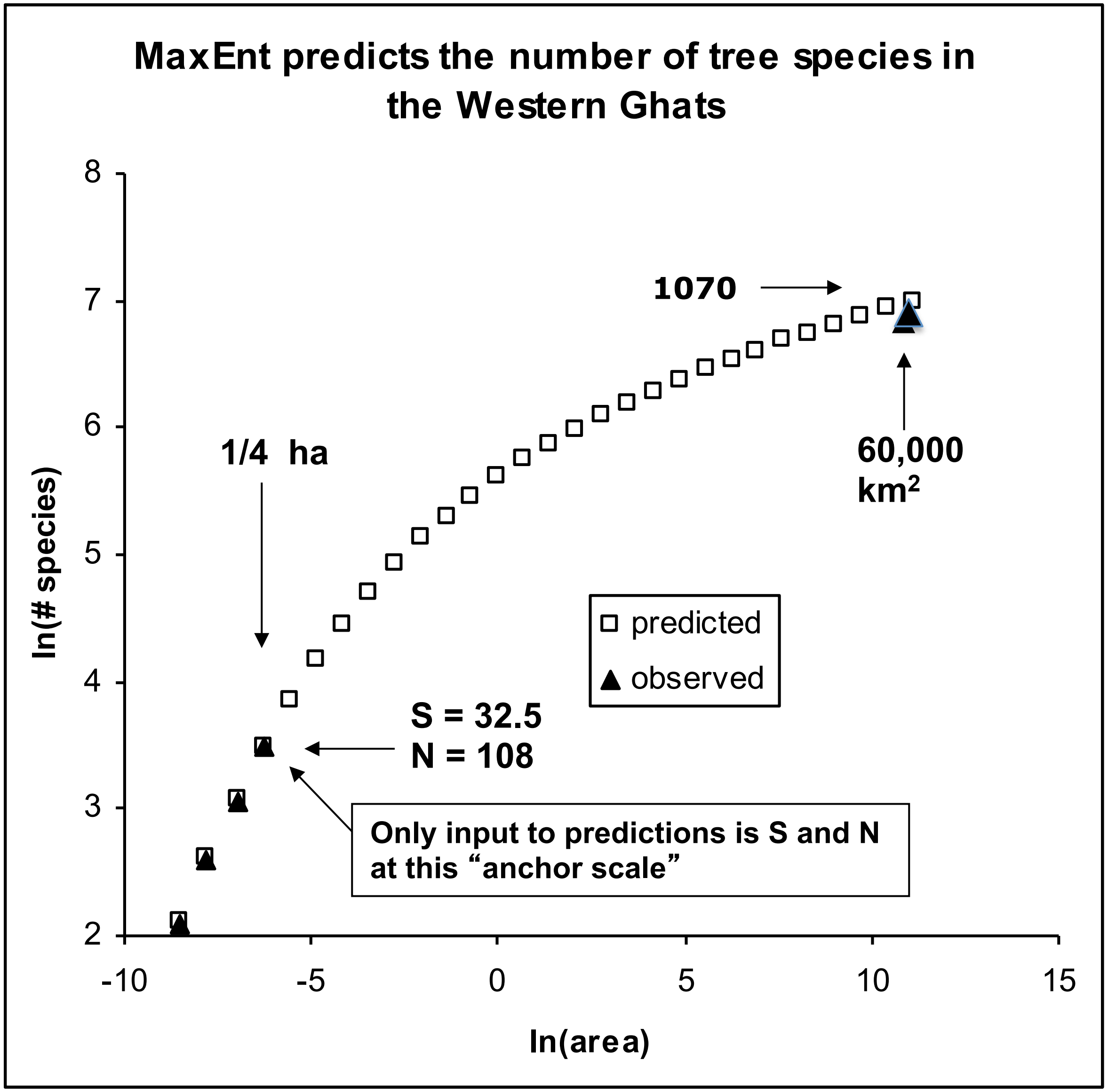
Harte et al., *Ecology Letters*, 2009;



The Concept of Non-Power-Law Scale Collapse



Now upscaling from small plot census data is possible



MaxEnt predicts the number of tree species in the Western Ghats

1070

1/4 ha

60,000 km²

S = 32.5
N = 108

Only input to predictions is S and N at this "anchor scale"

Downscale from "anchor value"

Upscale from "anchor value"

Presently, ~ 1040 tree species known to be in the preserve, but new ones being found at a rate of ~ 2-3/year.

We are extrapolating here over a scale range of

60,000 km² / ¼ ha

= 2^{24}

Harte et al., Ecology Letters, 2009

MaxEnt Predicts Food Web Structure (linkage distributions)

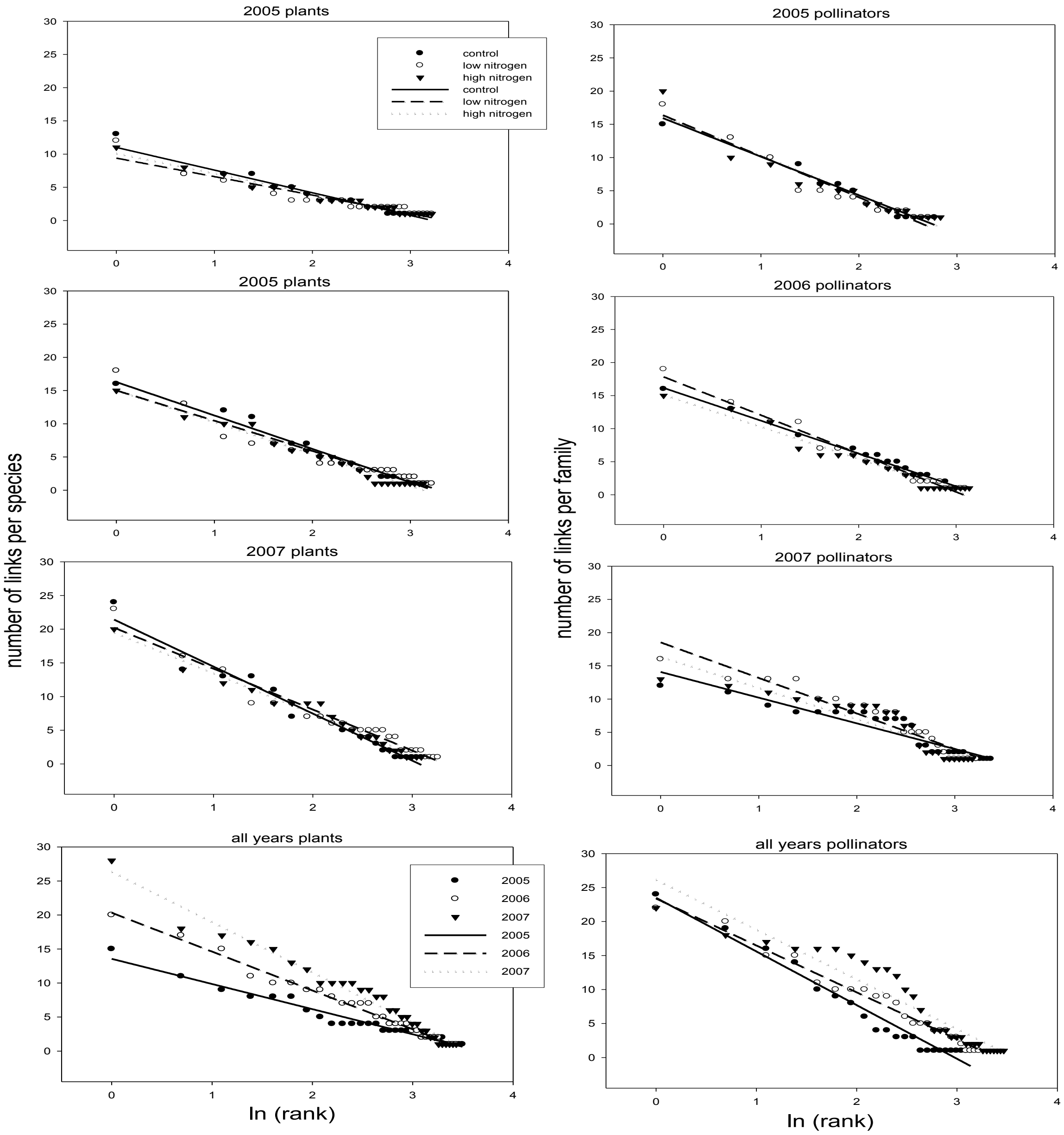
With the addition of the state variable *L* (total number of linkages in web), METE predicts:

linkages per species = a - b·log(rank)

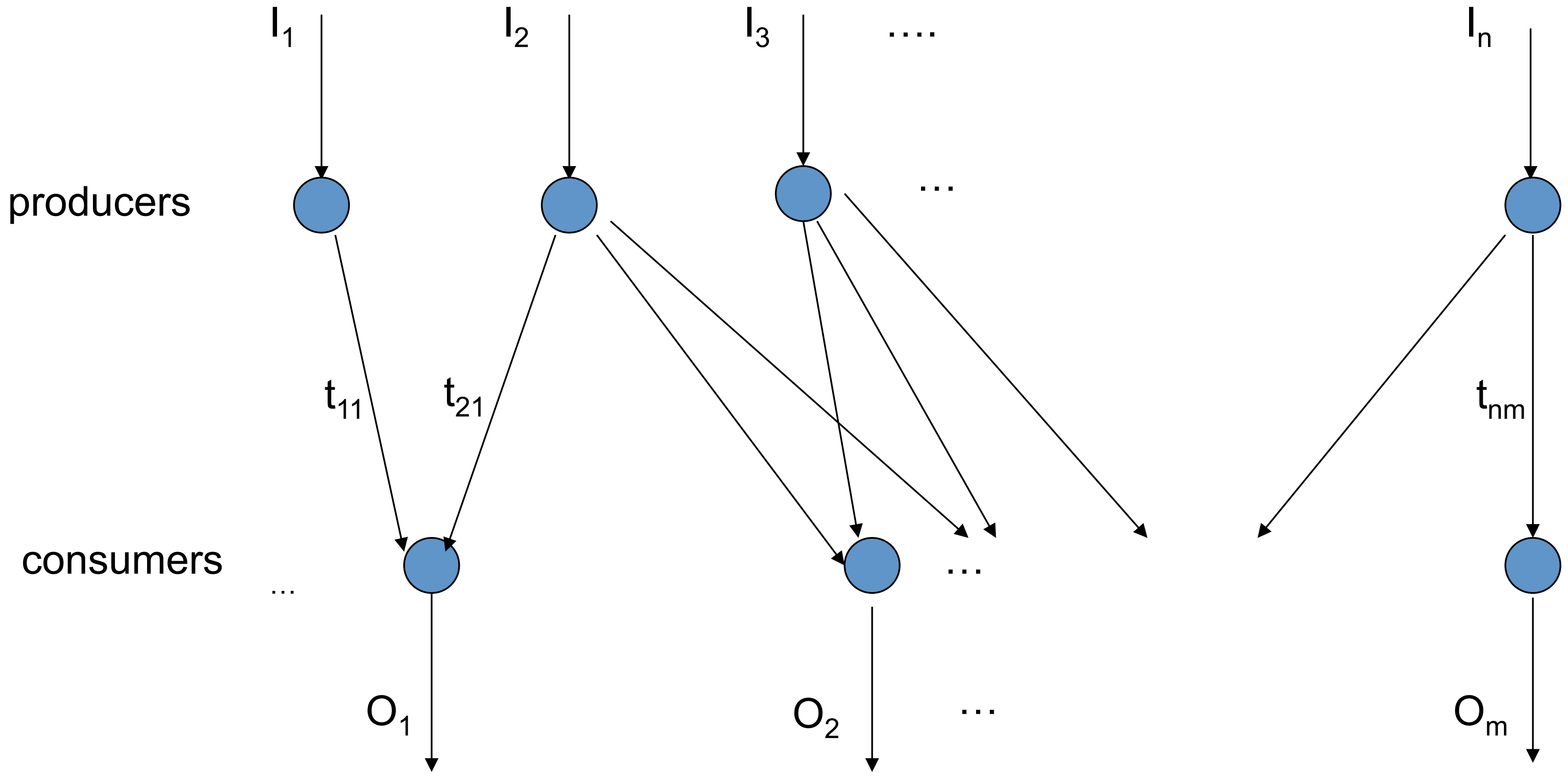
(See also new work on MaxEnt & food web link distributions by Rich Williams, 2009, 2010)

Plant-pollinator data from Laura Burkle generally follow the MaxEnt prediction.

Deviations exist however, and need to be better understood.



FLOWS IN NETWORKS: The MaxEnt Solution



$T = \Sigma_{ij} t_{ij}$;

$$\boxed{t_{ij} = I_i * O_j / T}$$

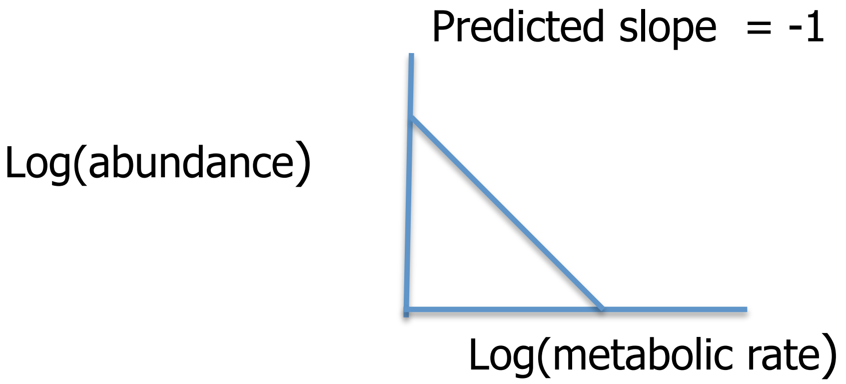
E. G.:

The I's are NPP's of plants

The O's are metabolic rates of herbivores

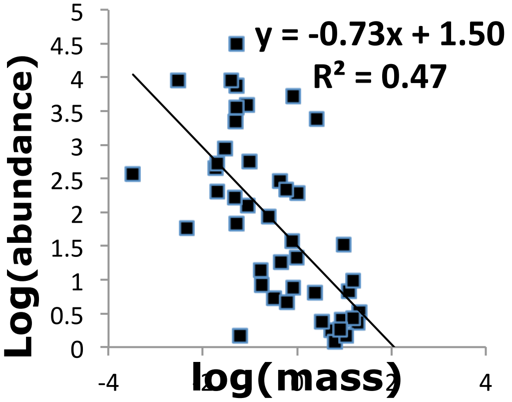
But: METE 1.0 has a problem!

It predicts exact Energy Equivalence: species divide up the energy pie equally.



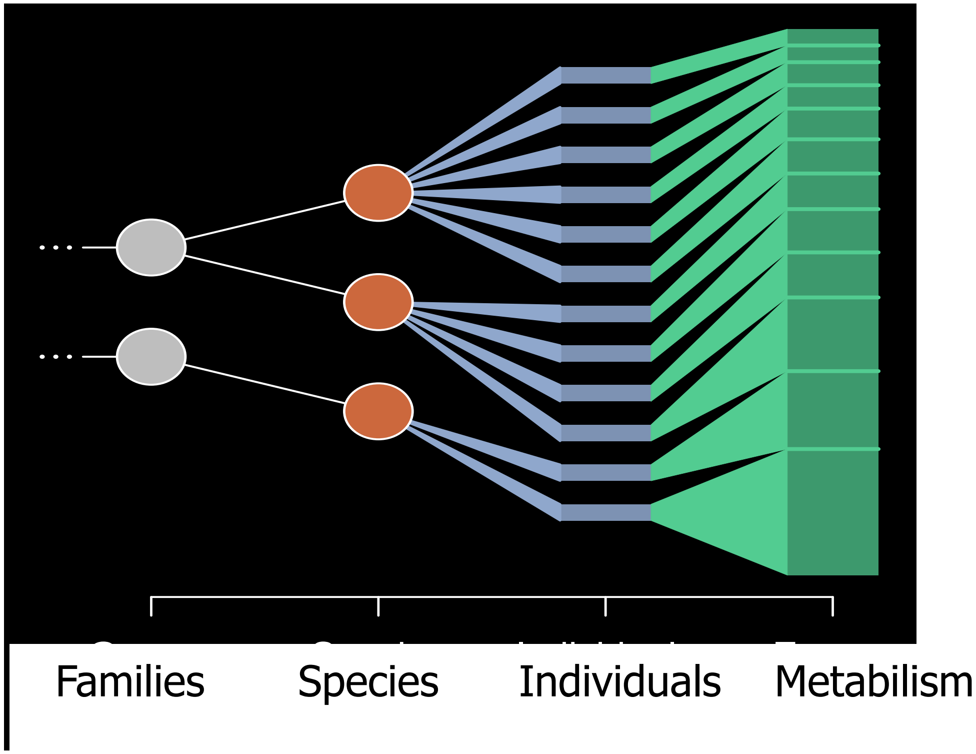
Rocky intertidal inverterbrate data
from Marquet et al., Science 1990

Using metabolism
~ mass$^{3/4}$



There is way more scatter
than predicted!

Solution: Add an additional state variable for # Families:

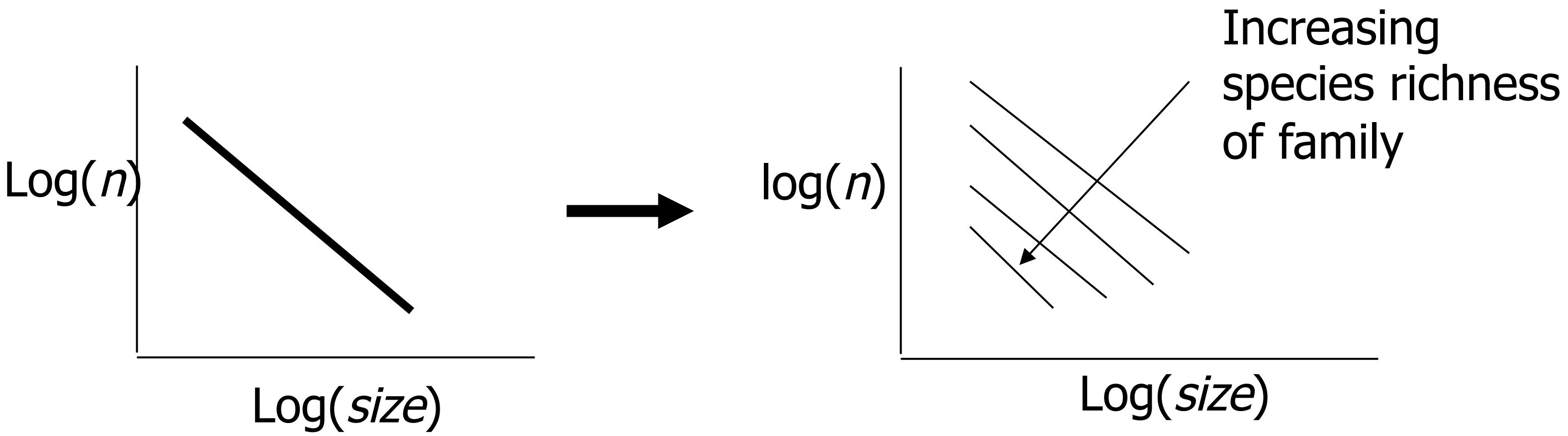


Suppose you add a new state variable: F = # families

If (A,S,N,E) → (A,F,S,N,E)

(F = family or other higher order category)

MaxEnt predicts the size-abundance relationship is modified. It splits apart!



And this greatly improves the accuracy of the predicted size-abundance relationship for:

- **intertidal invertebrates**
- **tropical forest trees**
- **North American mammals!**

Harte et al., *Ecology Letters* 2015.

The taxonomically extended theory predicts observed patterns in macroecology that depend on species richness of higher taxonomic levels:

1. The most abundant species should belong to families or genera that contain relatively few species.

(*Consistent with Amazon tree data: ter Steege et al., 2013*)

2. Rare species should be over-represented in species-rich families or genera.

(*Consistent with vascular plant data: Schwartz & Simberloff, 2001; Lozano & Schwartz, 2005*)

3. Species with the largest body sizes, and therefore largest metabolic rates of individuals, belong to families or genera with the fewest species. Moreover, the variance of body size across species should be greatest in families or genera with the fewest species

(*Both predictions consistent with mammal data: Smith et al., 2004*)

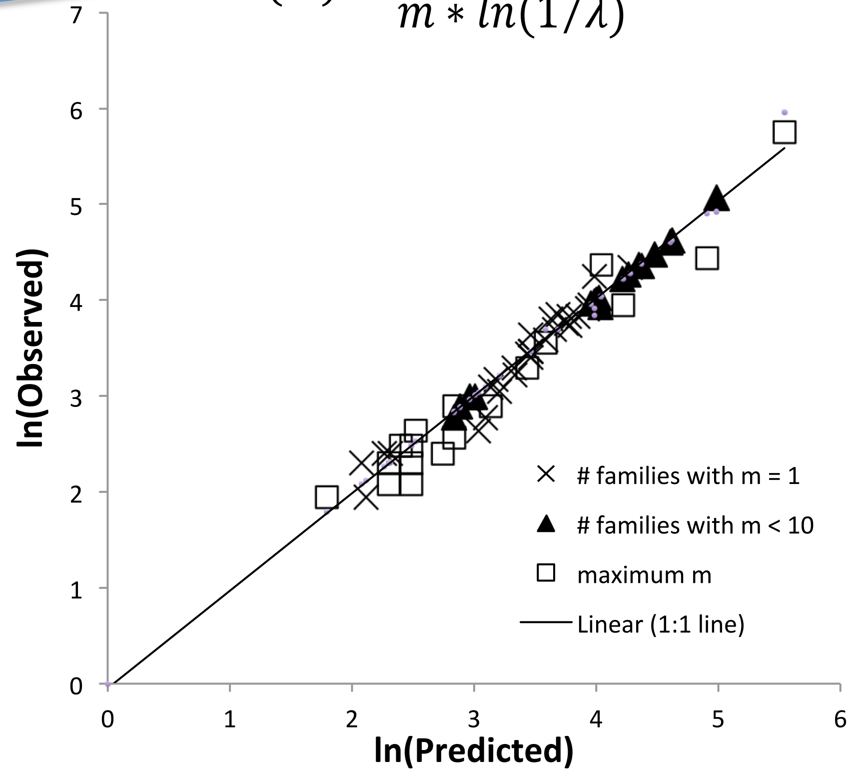
AFSNE also predicts the distribution of # species across families for arthropods, plants, birds, and microorganisms

m = # species per family

$\lambda = \lambda(F_0, S_0)$

The prediction requires knowledge only of the # of species and # families in a data set.

$$\Gamma(m) \approx \frac{e^{-\lambda * m}}{m * \ln(1/\lambda)}$$



Harte et al,. *Ecology Letters* 2015.

- Arthropod data from Basset et al. (2011), and Gruner (2007)
- Bird data consist of ten transects chosen randomly from the Breeding Bird Census (Sauer et al. 2014)
- Plant data from: census plots at Cape Point Preserve (Slingsby, pers. comm.); the Smithsonian Tropical Forest Research Institute plots at BCI (Condit 1998, Condit et al. 2004; Hubbell et al. 2005), Luquillo (Thompson et al. 2002); Sherman and Cocoli (Pyke et al. 2001; Condit et al. 2004); Yasuni (Valencia et al. 2003; 2004);
- Microbiome data (Wu et al. 2013; Larry Smarr pers. comm.).

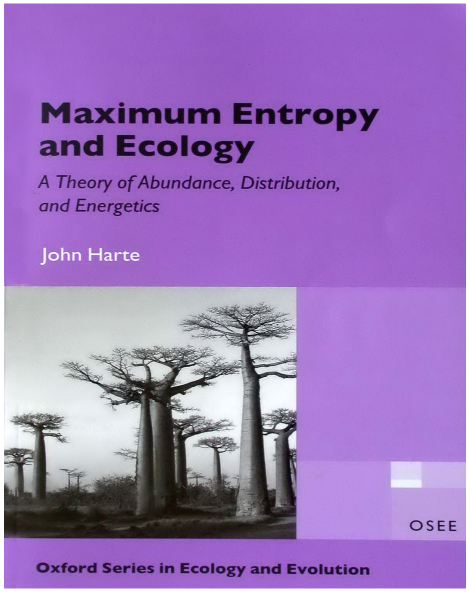
Summary of Static METE

(And a shameless self-promotion)

Ecological theory based on MaxEnt
allows us to predict the shapes of patterns in macroecology,

and thereby improve predictions of extinction rates under habitat loss, as well as estimates of species richness at large spatial scales.

I have spared you the math, but if you want
to learn more about MaxEnt, and how to use it, see:



Maximum Entropy and Ecology

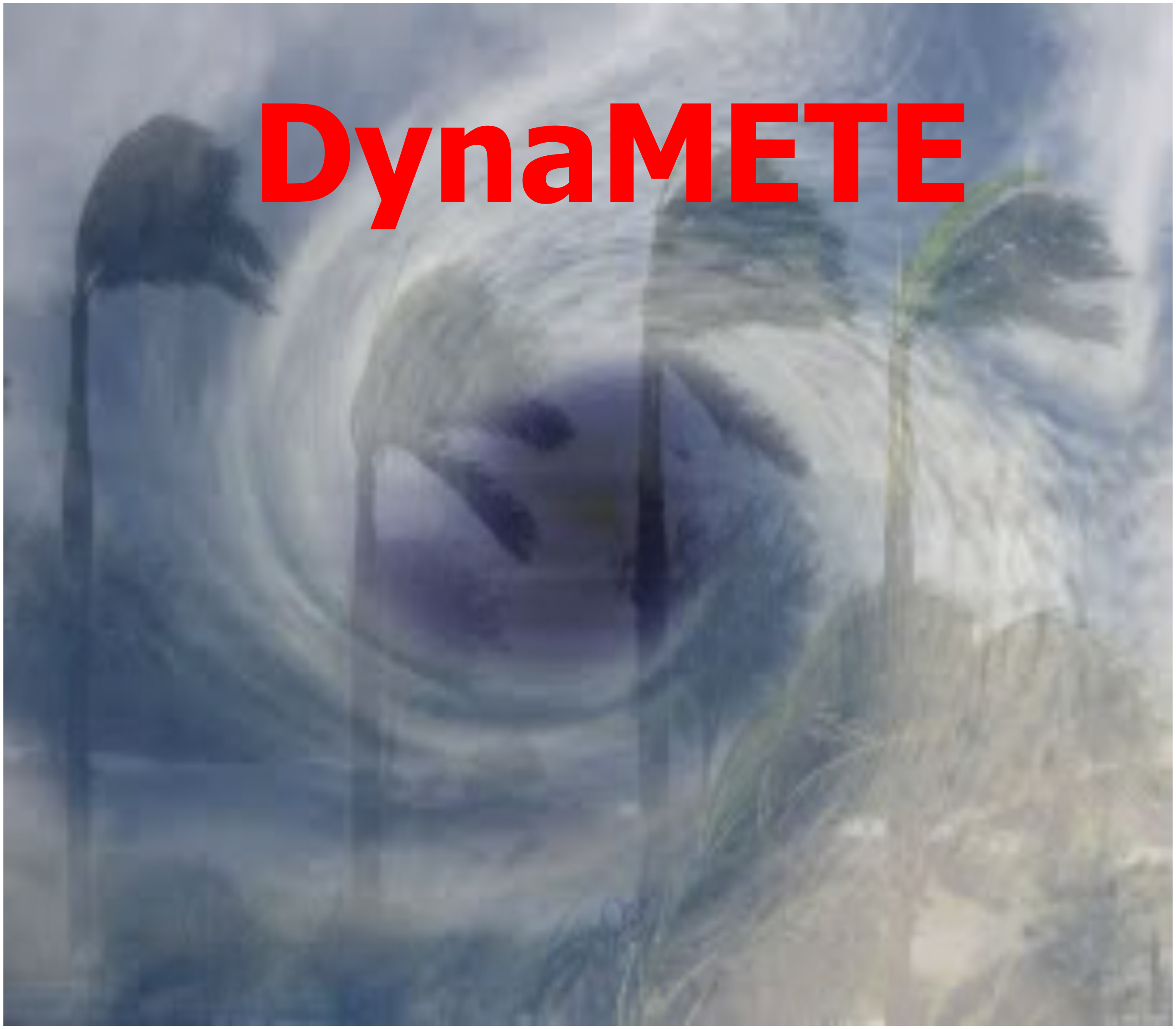
A Theory of Abundance, Distribution, and Energetics

John Harte

OSEE

Oxford Series in Ecology and Evolution

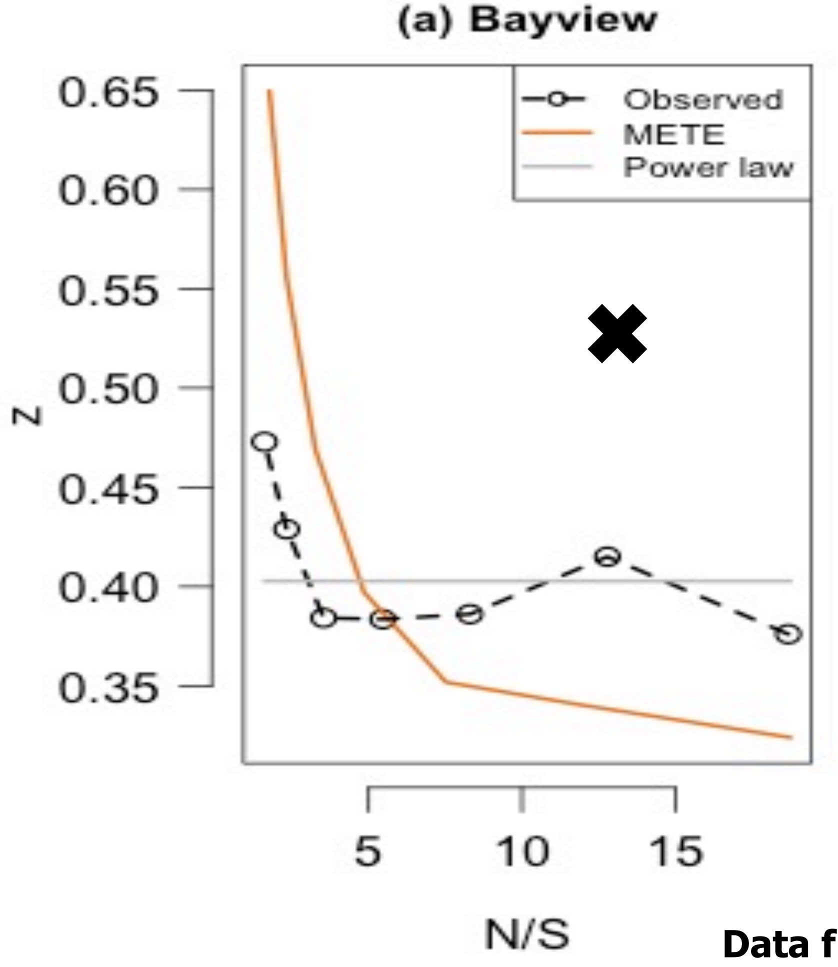
A big problem remains …we need:



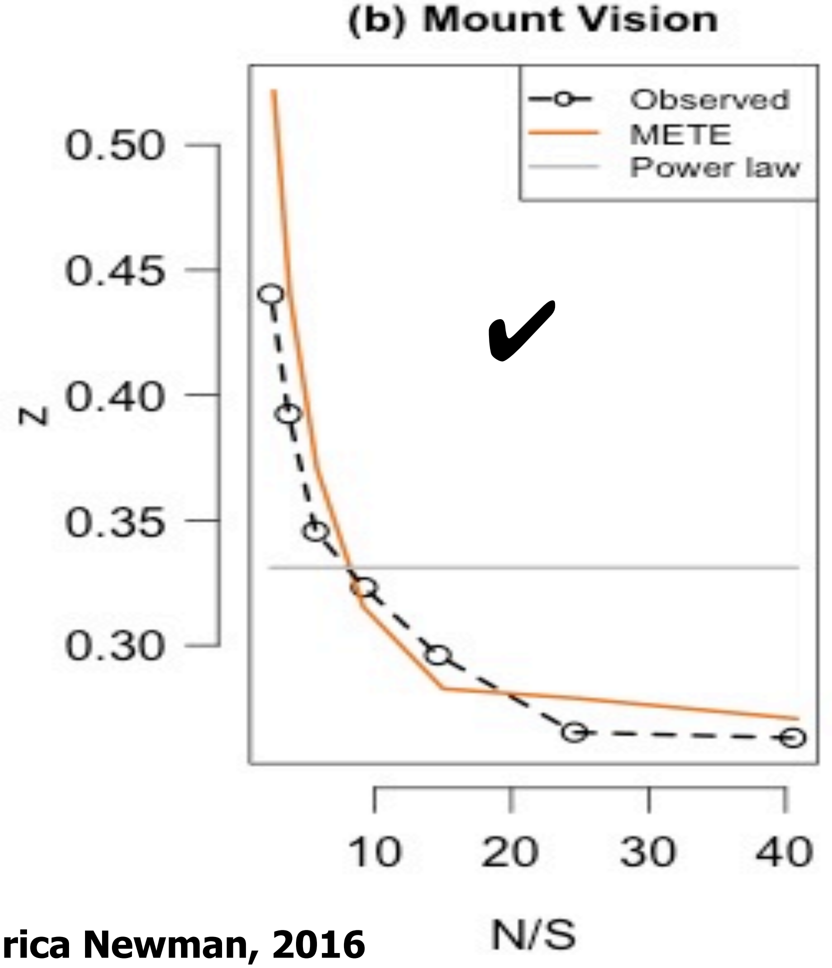
The METE predictions fail for ecosystems in which the State Variables are rapidly changing!

We have assembled many examples of this. Here is one: aftermath of fire in a fire-adapted Bishop Pine Forest

recently burned, rapidly changing



mature, quasi-steady-state



Data from Erica Newman, 2016

Further evidence that the theory fails to predict patterns in ecosystems undergoing relatively rapid change

1.

Abundance distribution of Rothampsted Moths

Relatively undisturbed fields:
Fisher log series distribution
(predicted by METE)

Fields recently fallowed
and in transition:
Lognormal distribution

Kempton and Taylor (1974)

2.

Abundance distribution of trees in Smithsonian tropical forest plots

The most disturbed plot (Barro Colorado Island in Panama) shows the most deviation from METE

3.

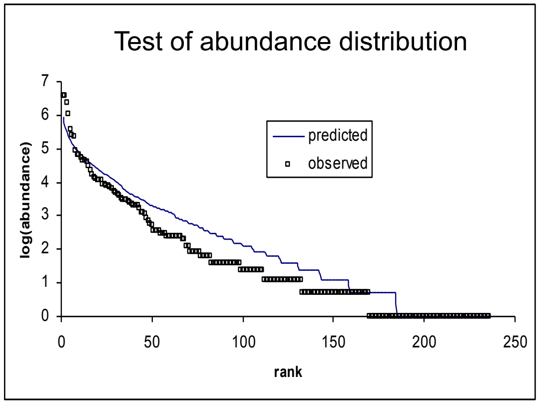
Abundance distributions of Hawaiian Arthropods

sites of different ages and stages of diversification

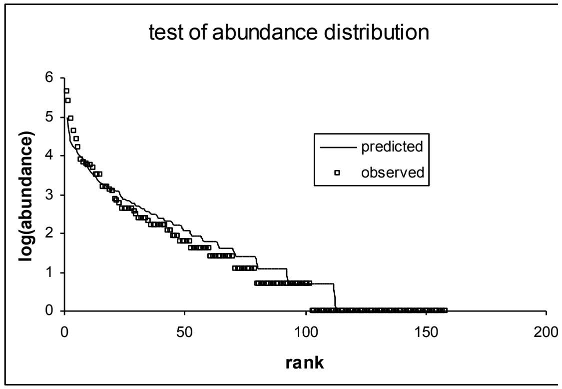
Data from Dan Gruner

Similar pattern of success and failure for body size distributions!

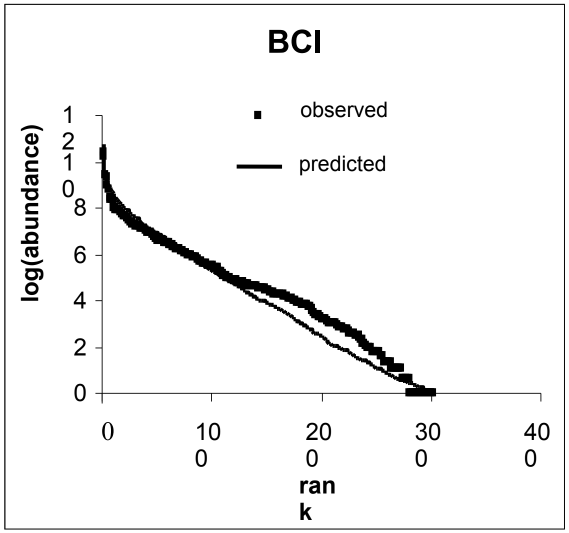
150 y



4 My



Another Example
The BCI 50 ha plot: A dynamic system



The creation of Gatun Lake isolated the plot from its immigrant source pool. It is losing species (Condit et al.; Egbert Leigh, pers. comm.)

The predicted log-series abundance distribution fails at BCI.

Other examples:

Arthropods in newly-fallowed fields at Rothampsted (Kempton and Taylor*, J. Animal Ecol*. 1974)
Arthropods on Hawaiian islands of differing ages (Rominger et al., *Glob. Ecol. Biogeogr.* 2016)
Rodents in exclosure experiments at Portal (Supp et al., *Ecology* 2012)

Stochastic Dynamic MaxEnt theory: DynaMETE

1. The probability distribution, $P(S,N,E,t)$, of the state variables is governed by a Master Equation that is driven by ontogenic growth, birth, death, migration, speciation, and extinction transition probabilities.

2. The dynamic structure function is the superstatistical distribution of the MaxEnt solution over $P(S,N,E,t)$:

$$D(n,\varepsilon,t) = \int_1^{S_0} dS \int_S^{N_0} dN \int_N^{E_0} dE \left\{ R(n,\varepsilon \mid S,N,E) * P(S,N,E,t) \right\}$$

3. Derive the time-dependent metrics of macroecology from R', just as in the static theory. E.G., the dynamic species-abundance distribution $\Phi(n,\varepsilon,t) = \Sigma_\varepsilon D(n,\varepsilon,t)$

The Basic Idea of the Master Equation for a Probability Distribution

dP(n,t)/dt = flow in – flow out   of probability =

+ rate(n - 1 → n) • P(n - 1, t)

+ rate(n + 1 → n) • P(n+1, t)

- [rate(n → n - 1) +rate(n → n + 1)] • P(n, t)

e.g. If n = abundance,

Rate(n - 1 → n) = birth rate + immigration rate

Six Transition processes drive the master equations

Transition Process	Effect of Transition	Approach
Ontogenic Growth	$E \rightarrow E+1$	Logistic-type growth of individuals (can be modified by metabolic scaling theory)
Birth	$N \rightarrow N+1$	Fixed per capita birth rate and density-dependent death rate
Death	$N \rightarrow N-1$	
Speciation	$S \rightarrow S+1$	Proportional to S or N? S.
Immigration	$N \rightarrow N+1$, $S \rightarrow S$ or $N \rightarrow N+1$, $S \rightarrow S+1$	Constant influx rate of individuals, drawn from metacommunity species pool.
Local extinction	$S \rightarrow S-1$, $N \rightarrow N-1$	"Last flower standing". For the t → t+1 transition, the abundance distribution at time t tells us how many singleton species (n=1) there were, and the death rate then tells us the probability of an extinction.

Simplification:

$$\boldsymbol{P(S, N, E, t) = F(S, t; f(N,E)) \cdot G(N, t; g(S,E)) \cdot H(E, t; h(S,N))}$$

which turns a ghastly coupled master Eq. into three manageable weakly-coupled master eqs.

In the eq. for $F(S,t+1)$, the function $f(N,E)$ is an average of a derived function over $G(N,t)$ and $H(E,t)$ at time t.

And similarly for the functions $g(S,E)$ and $h(S,N)$

Ontogenic Growth

For a single individual: $d\epsilon/dt|_{\text{ontogenic}} = w(\epsilon) = w_0\epsilon^{2/3} - w_1\epsilon$

Sum over individuals of $w_1\epsilon = w_1E$

But what about: Sum of $w_0\epsilon^{2/3}$???

The sum of $\epsilon^{2/3}$ *does not equal (sum of* $\epsilon)^{2/3} = E^{2/3}$

We write the sum over individuals of $w_0\epsilon^{2/3}$ as:

$$\sum_{S,N} F(S)G(N)N \sum_{n,\varepsilon} w_0 \varepsilon^{2/3} n \frac{S}{N} R(n,\varepsilon|S,N,E)$$

$$= w_0 \sum_{S,N} F(S)G(N)S \frac{\lambda_2}{\ln\left(\frac{1}{1-e^{-\beta}}\right)} \sum_{n} n e^{-\lambda_1 n} \sum_{\varepsilon} \varepsilon^{\frac{2}{3}} e^{-\lambda_2 n \varepsilon}$$

$$= w_0 \sum_{S,N} F(S)G(N)S \frac{1}{{\lambda_2}^{2/3} \ln\left(\frac{1}{1-e^{-\beta}}\right)} \sum_{n} \frac{e^{-\beta n}}{n^{2/3}}$$

$$= w_0 \sum_{S,N} F(S)G(N) \frac{S e^{-\beta}}{\beta^{\frac{1}{3}} {\lambda_2}^{2/3} \ln\left(\frac{1}{1-e^{-\beta}}\right)}$$

$$= w_0 E^{2/3} < \frac{N^{1/3} e^{-\beta}}{\ln^{2/3}\left(\frac{1}{1-e^{-\beta}}\right)} >_{F(S),G(N)}$$

$\beta \ln(1/\beta) = S/N$

$\lambda_2 = S/E$

Birth and death:

$$\frac{dn}{dt} = r(n,\varepsilon) = \frac{b_0 n}{\varepsilon^{1/3}} - \frac{d_0 n}{\varepsilon^{1/3}}\left(1 + \frac{n}{K}\right) = \frac{b_0 n}{\varepsilon^{1/3}} - \frac{d_0 n}{\varepsilon^{1/3}} - \frac{d_1 n^2}{\varepsilon^{1/3}}$$

This results in

$$r_{\mathrm{dem}}(N) = r_b(N) - r_d(N) = b_0 N^{4/3} < \frac{\ln^{1/3}\left(\frac{1}{1-e^{-\beta}}\right) e^{-\beta}}{E^{1/3}} >_{F(S),H(E)}$$

$$-d_0 N^{4/3} < \frac{\ln^{1/3}\left(\frac{1}{1-e^{-\beta}}\right) e^{-\beta}}{E^{1/3}} >_{F(S),H(E)}$$

$$-d_1 N^{7/3} < \frac{\ln^{4/3}\left(\frac{1}{1-e^{-\beta}}\right) e^{-\beta}}{SE^{1/3}} >_{F(S),\mathrm{H(E)}}$$

Extinction:

$$\sigma = d(1)S\phi(1)$$

Using the same technique as above:

$$\sigma(S) = (d_0 + d_0)S^{4/3} < \frac{e^{-\beta}}{\ln\left(\frac{1}{1-e^{-\beta}}\right)E^{1/3}} >_{G(N),H(E)}$$

Speciation = $\lambda_0 S$

Immigration = m_0

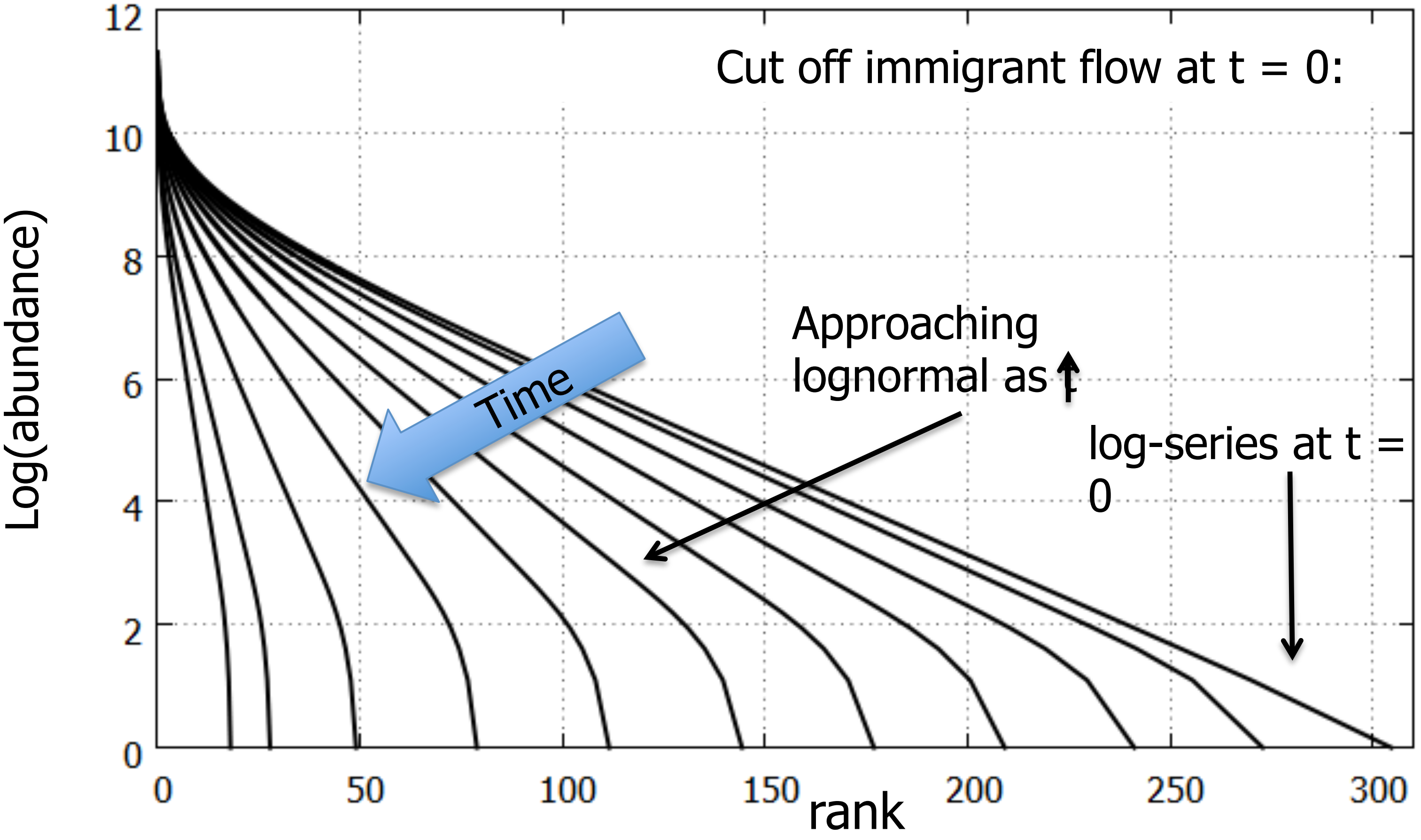
Putting this all together:

$$G(N,t+1) - G(N,t) = -[r_b(N,t) + r_d(N,t) + m_0]G(N,t)$$
$$+[r_b(N-1) + m_0]G(N-1,t) + r_d(N+1,t)G(N+1,t)$$

$$H(E,t+1) - H(E,t) = w_{\mathrm{ont}}(E-1,t)H(E-1,t)$$
$$+w_{\mathrm{death}}(E+1,t)H(E+1,t) - w_{\mathrm{ont}}(E,t)H(E,t) -$$
$$w_{\mathrm{death}}(E,t)H(E,t)$$

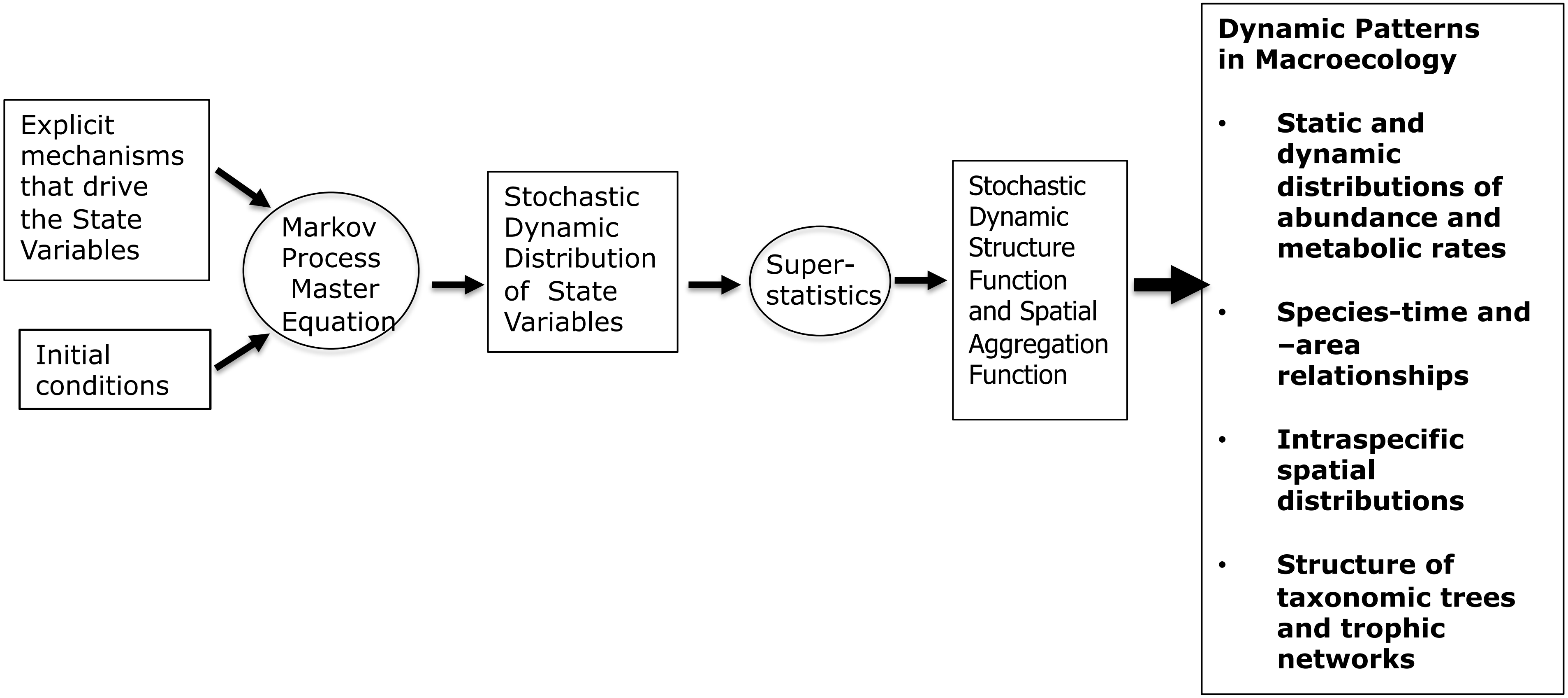
$$F(S,t+1) - F(S,t) =$$
$$-(\lambda_0 S + m(S) + \sigma(S,t))F(S,t)$$
$$+(\lambda_0(S-1) + m(S-1))F(S-1,t)$$
$$+\sigma(S+1,t)F(S+1,t)$$

An over-simplified version of this predicts a shift in the species abundance distribution from log-series toward lognormal when a system is isolated from its surrounding source of immigrants.



This looks like what is happening at Barro Colorado

SUMMARY:
A hybrid theory (MaxEnt + Mechanism)
Is a promising approach to predicting
The static and dynamic patterns of macroecology



Explicit mechanisms that drive the State Variables

Initial conditions

Markov Process Master Equation

Stochastic Dynamic Distribution of State Variables

Super-statistics

Stochastic Dynamic Structure Function and Spatial Aggregation Function

Dynamic Patterns in Macroecology

- **Static and dynamic distributions of abundance and metabolic rates**
- **Species-time and –area relationships**
- **Intraspecific spatial distributions**
- **Structure of taxonomic trees and trophic networks**

What about Mechanism? (Where’s the meat in METE?)

Four mutually compatible responses:

1. **The state variables embody the mechanisms.**
   Success of METE does not imply mechanism does not matter; the mechanisms are incorporated into the values of the state variables. No further mechanistic assumptions are needed.

2. **The role of mechanism identified when MaxEnt fails.**
   Analogy with PV = nRT and van der Walls force.
   E.G., fixing the energy equivalence rule with evolutionary history.

3. **Mechanism is an emergent property of entropy maximization**
   E.G., birth and death rate dependence upon population size emerges from maximizing $S = k\log(W)$ .. (Zhang & Harte, Theor. Pop. Biology 2015)

4. **Where’s the mechanism in Statistical Mechanics or Quantum Mechanisms?**

MaxEnt is not mechanism-less; rather it avoids having to pre-judge what mechanisms and traits are important.

What if MaxEnt gives a poor prediction?

1. You made a mathematical error in working out the solutions to the equations.

2. Your prior knowledge (in the form of constraints) is not factually correct. If one of your constraints is the mean value over the sought distribution, and you mis-measured it or the values of your state variables were mis-estimated, your predictions will suffer as a consequence.

3. Your constraints may not provide enough adequate information to determine a good answer. If you neglect information, MaxEnt will do its best for you. But it may not be good enough. YOUR ANSWER WILL BE THE BEST POSSIBLE, GIVEN THE CONSTRAINTS THAT YOU USE.

A note on alternative measures of entropy

Tsallis entropy: $S_{\mathrm{Tsallis}} = (q-1)^{-1} (1 - \Sigma_n [p(n)]^q)$

$\rightarrow S_{\mathrm{shannon}}$ as $q \rightarrow 1$

The *Π* distribution describing spatial aggregation only matches data for q ~ 1

Moreover, maximum Tsallis entropy distributions can be construed as resulting from a “superstatistical distribution” of maximum Shannon entropy distributions:

MEP: an example (Paltridge, 1975, 78)

Consider the heat budget of the earth:

Q=heat convection from equator to pole

$$\Omega_e = \sigma T_e^4 + 2Q$$

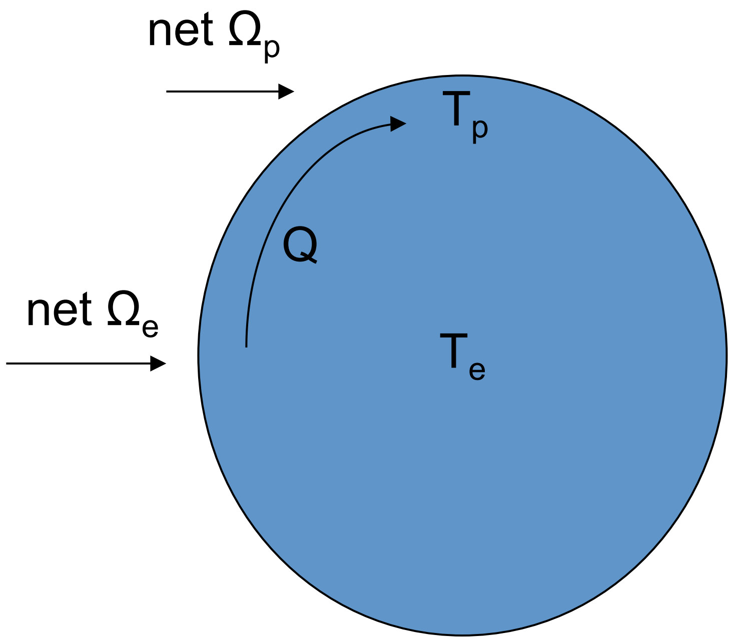
$$\Omega_p + Q = \sigma T_p^4$$ (*can add IR-absorbing atmosphere*)

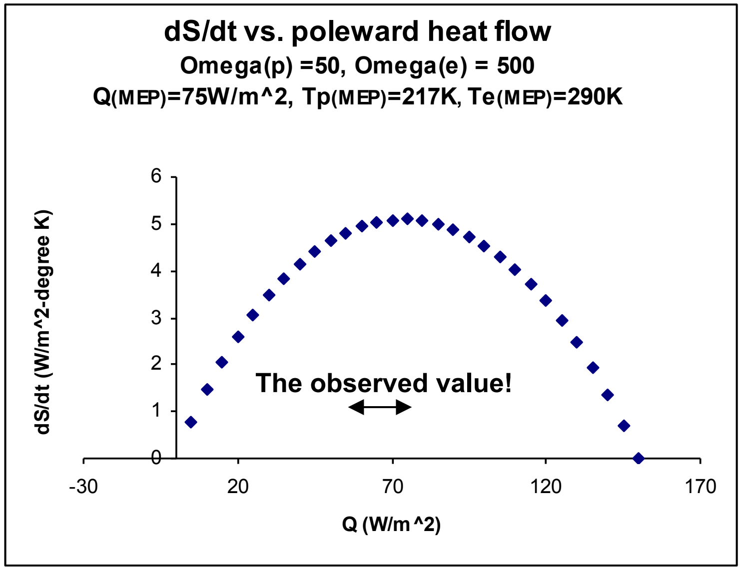
$$\mathbf{dS/dt} = \frac{\mathbf{Q}}{\mathbf{Tp}} - \frac{\mathbf{Q}}{\mathbf{Te}} =$$

$$\frac{Q}{[(\Omega_p + Q)/\sigma]^{1/4}} - \frac{Q}{[(\Omega_e - 2Q)/\sigma]^{1/4}}$$

Now maximize by setting

d(dSdt) /dQ = 0 and solve for Q, T_p, T_e





dS/dt vs. poleward heat flow

Omega(p) =50, Omega(e) = 500

Q(MEP)=75W/m^2, Tp(MEP)=217K, Te(MEP)=290K

MEP works on Earth, Mars, Titan!

Summary

The MaxEnt principle and specification of a few state variables predicts, with no adjustable parameters, realistic expressions for:

- species-level spatial abundance distributions
- relative abundance distribution (Fisher log-series), collector's curve
- species-area and endemics-area relationships
- intra-specific and inter-specific metabolic rate distributions
- distribution of linkages across nodes in plant-pollinator & other food webs
- the scale collapse of all species-area relationships and thus species richness at large scale from knowledge of it at small scale
- metabolism-abundance relationship, with unexpected yet realistic dependence on the structure of the taxonomic tree.

Thanks:

To my Collaborators:

Erin Conlisk
Justin Kitzes
Erica Newman
Adam Smith
Andrew Rominger
Tommaso Zillio
Xiao Xiao
Ethan White
Yu Zhang
Mark Wilber
Chloe Lewis
Wenyu Zhang

To Other Sources of Data:

J. Green
R. Condit
D. Gruner
R. Krishnamani
P. Harnik
J. Goddard
J. Godinez
K. Cherukumilla
STRI
David Storch
W. Kunin
E. White
D. Bartholomew
Brian Enquist
Cory Merow
Yves Basset

To the Funders:

NSF, Miller Foundation, Gordon and Betty Moore Foundation

To my Hosts during the development of METE:

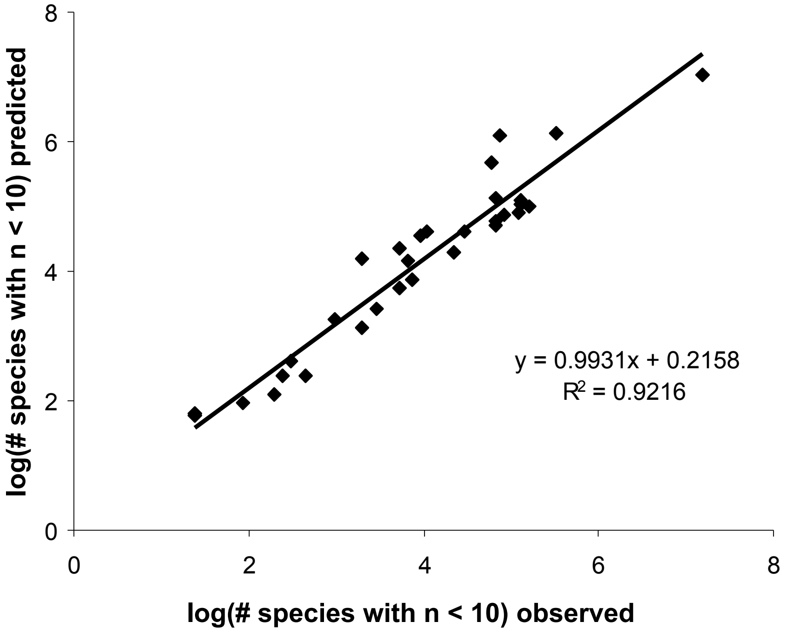
Santa Fe Institute, Rocky Mountain Biological Laboratory, NCEAS, The Chilean Ecological Society, Charles University, University of Padua

A finer look at the species-abundance distribution

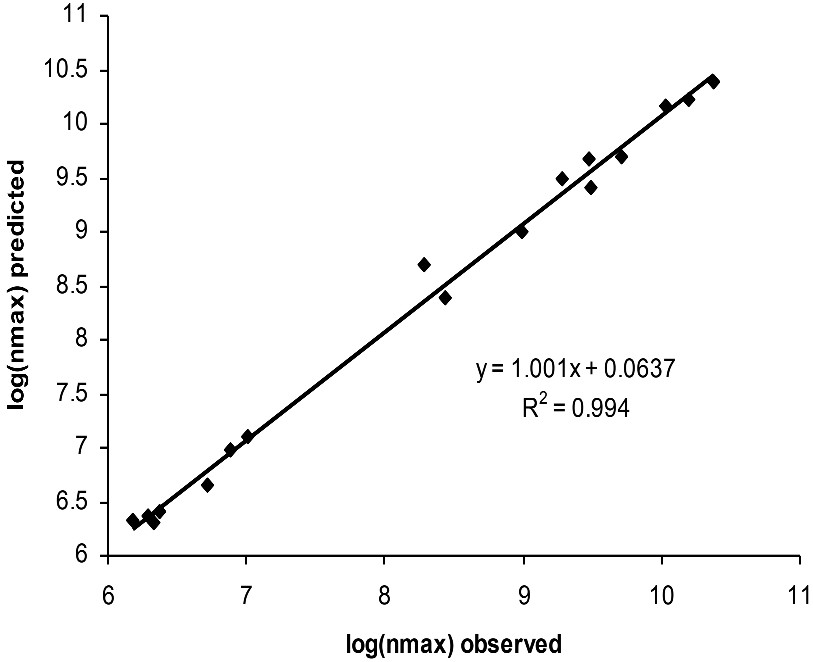
$$\boxed{\Phi(n \mid S_0, N_0) = \frac{e^{-\beta n}}{n \ln(1/\beta)}}$$

$\beta = \lambda_1 + \lambda_2$

MaxEnt predicts the number of rare species



... and the abundance of the most abundant species



Harte et al., Ecology, 2009

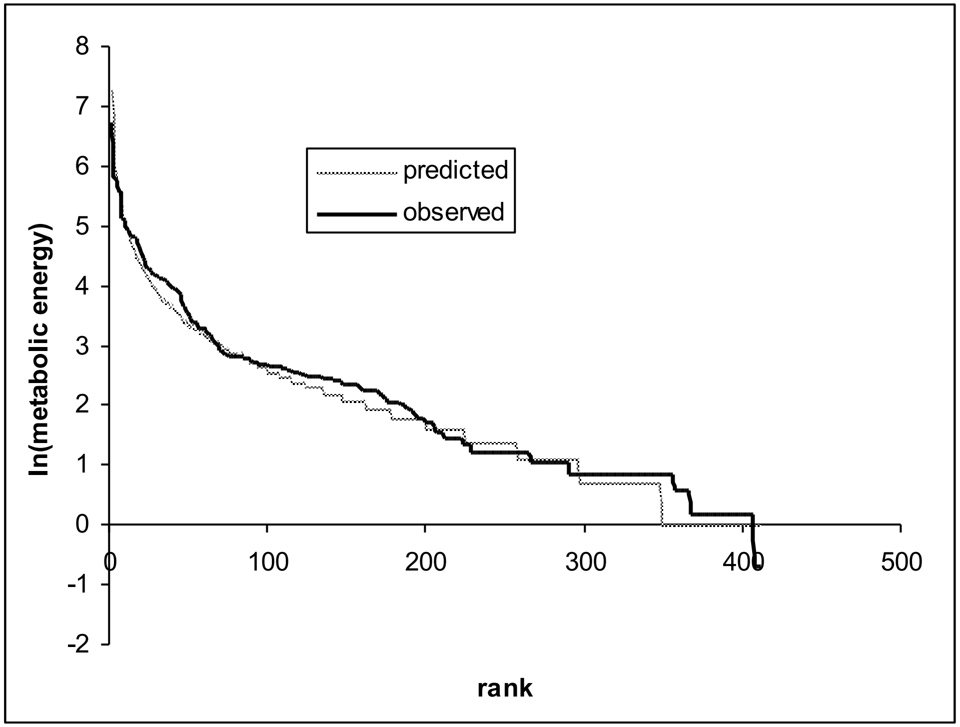
Harte, Oxford U. Press, 2011

A finer look at the distribution of metabolic rates over individuals

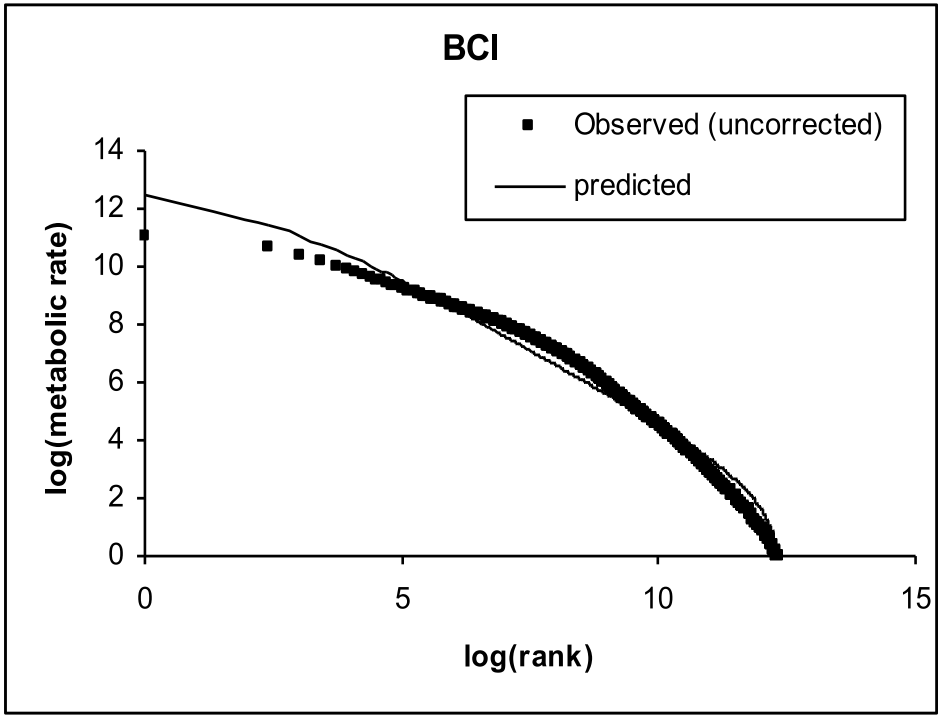
$$\Psi(\varepsilon \mid S_0, N_0, E_0) \approx \lambda_2 \cdot \beta \cdot \frac{e^{-\gamma}}{(1-e^{-\gamma})^2}$$

$\gamma = \lambda_1 + \lambda_2\varepsilon$

pit-trapped insects; assume $\varepsilon \sim m^{3/4}$

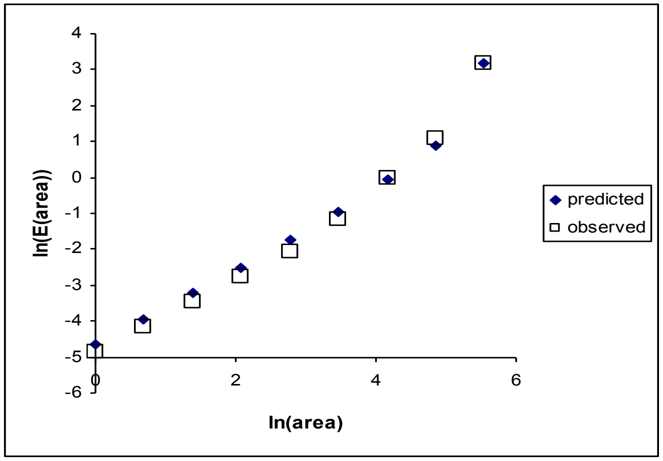


BCI forest; assume $\varepsilon \sim$ basal area

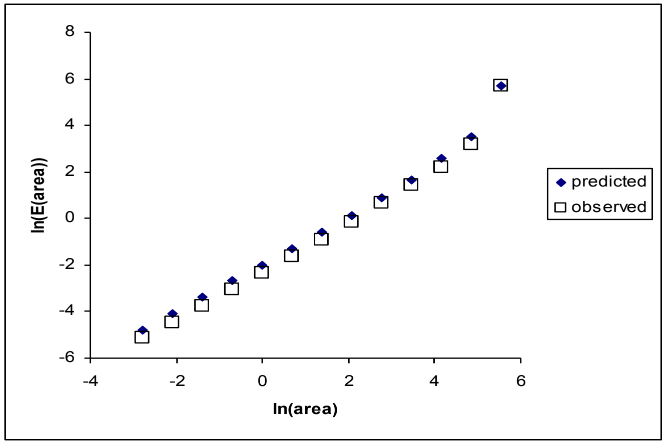


The Endemics-Area Relationship

$$E(A) = S_0 \sum_{n_0=1}^{N_0} \Phi(n_0) * \mathit{\Pi} \ (n_0 \mid n = n_0, A, A_0)$$



Serpentine



BCI

Scale collapse of the species-area relainship allows us to predict the slopes of species-area curves at large spatial scales, thereby allowing estimation of:

a. Species Richness at large spatial scales

E.G., 500,000 arthropod species and 18,000 tree species in Amazonia (Harte and Kitzes, 2015)

b. Extinction Rates under habitat loss, taking into account the decrease in slope, z, with increasing spatial scale:

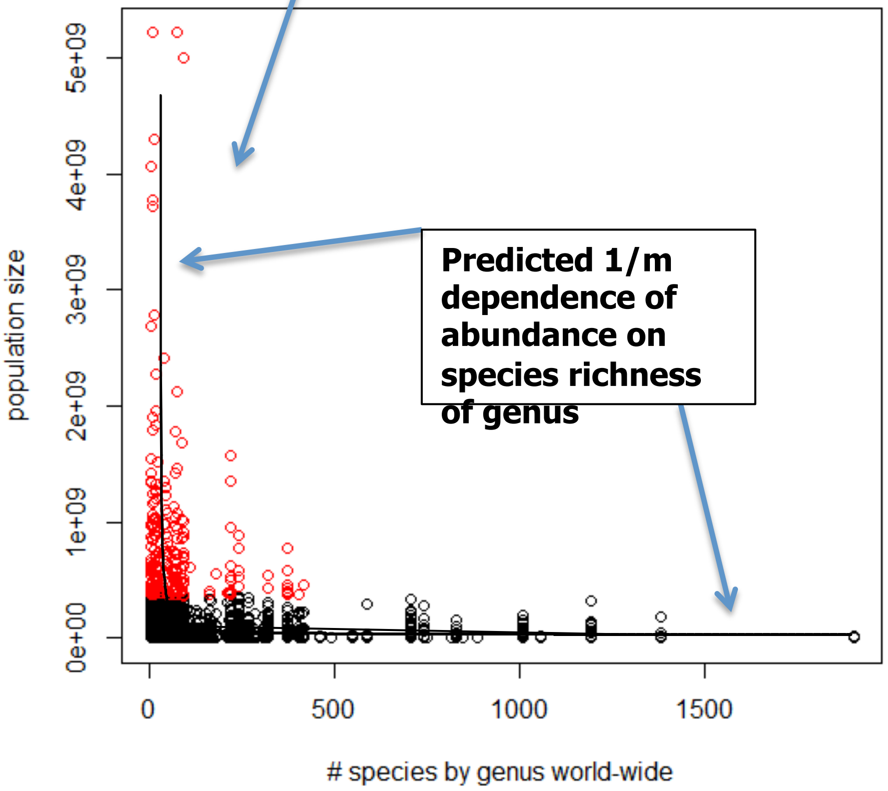
Initial habitat loss results in fewer extinctions than previously predicted with $z = .25$, subsequent habitat loss results in more extinctions than previously predicted (Kitzes and Harte, 2011)

c. The conditions under which an “extinction debt” should occur under habitat loss (Kitzes and Harte, 2015).

The taxonomically extended theory predicts that observed patterns in macroecology depend on species richness of higher taxonomic levels

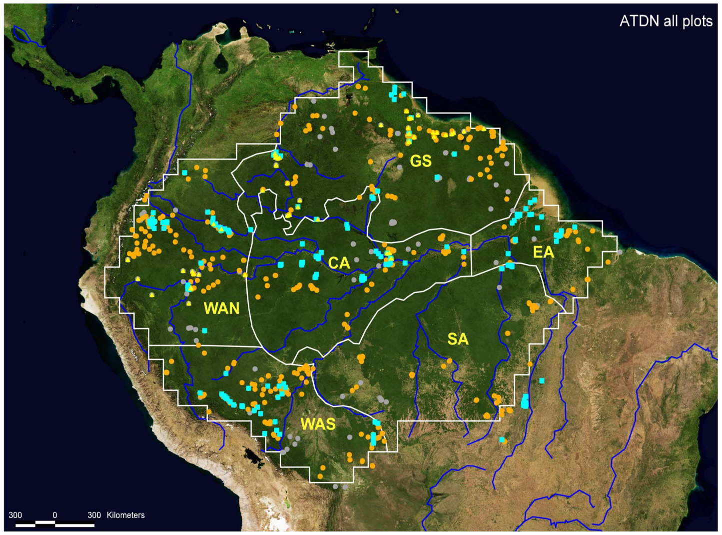
Prediction 1: The most abundant species belong to families or genera that contain relatively few species:

(*e.g., Amazon tree data: ter Steege et al., Science 2013*)



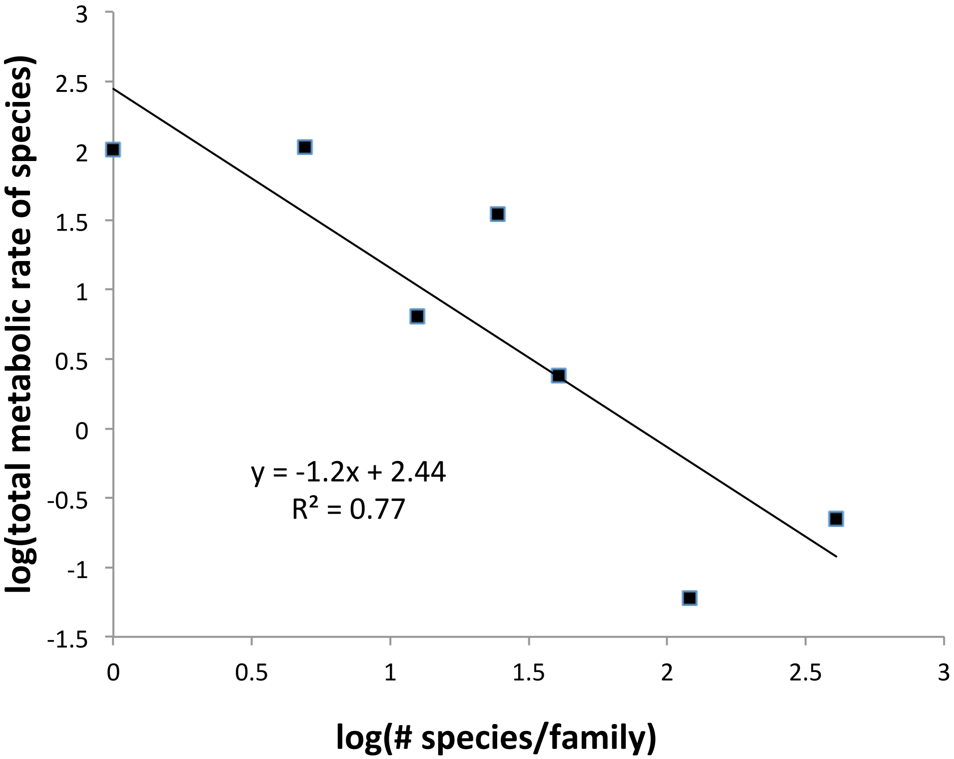
Predicted 1/m dependence of abundance on species richness of genus

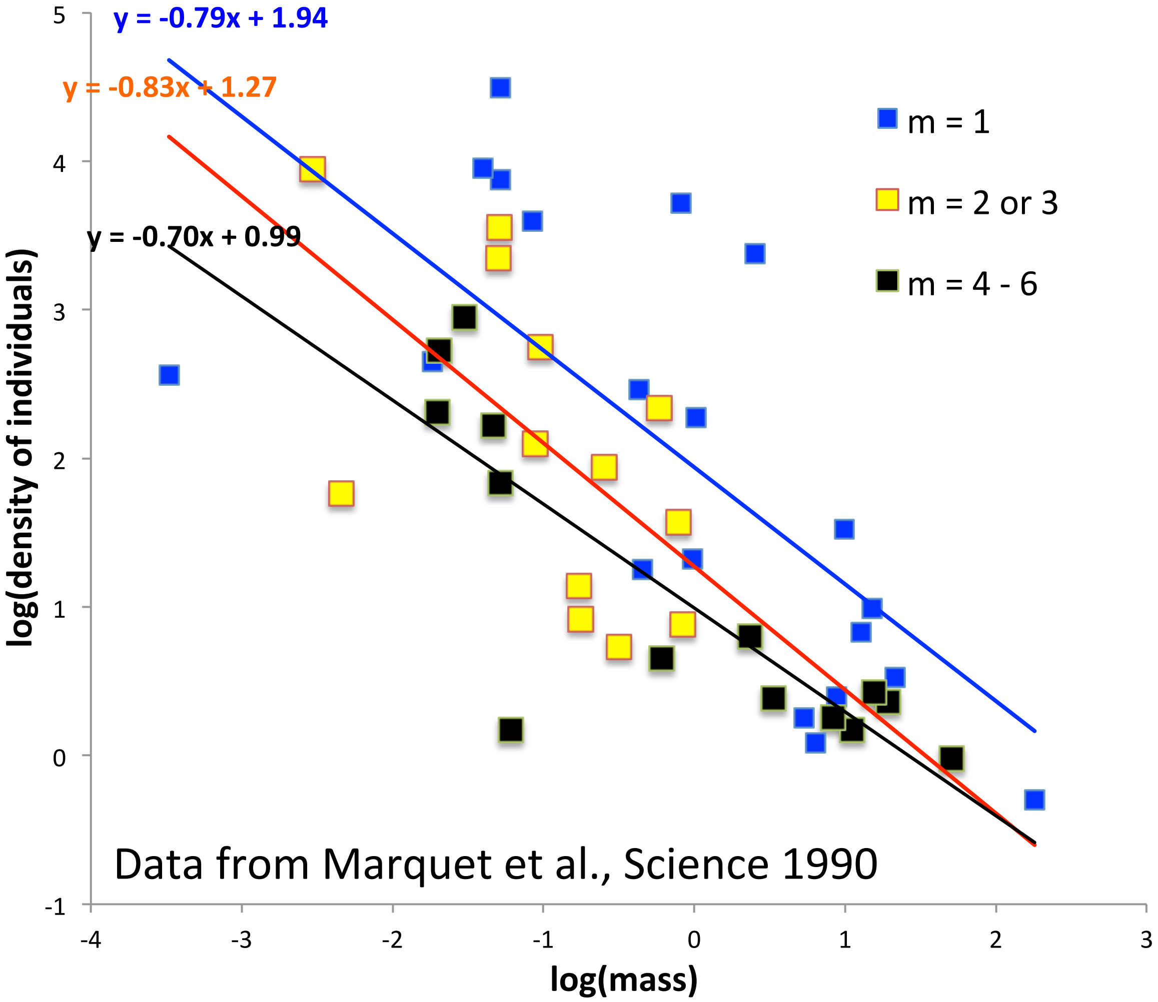
$$< n \mid m > \approx \frac{e^{-\beta m}}{m\left[\frac{G_0}{N_0} - \beta \ln(m)\right]}$$



Trees at Barro Colorado Island:
Energy Equivalence and Taxonomy

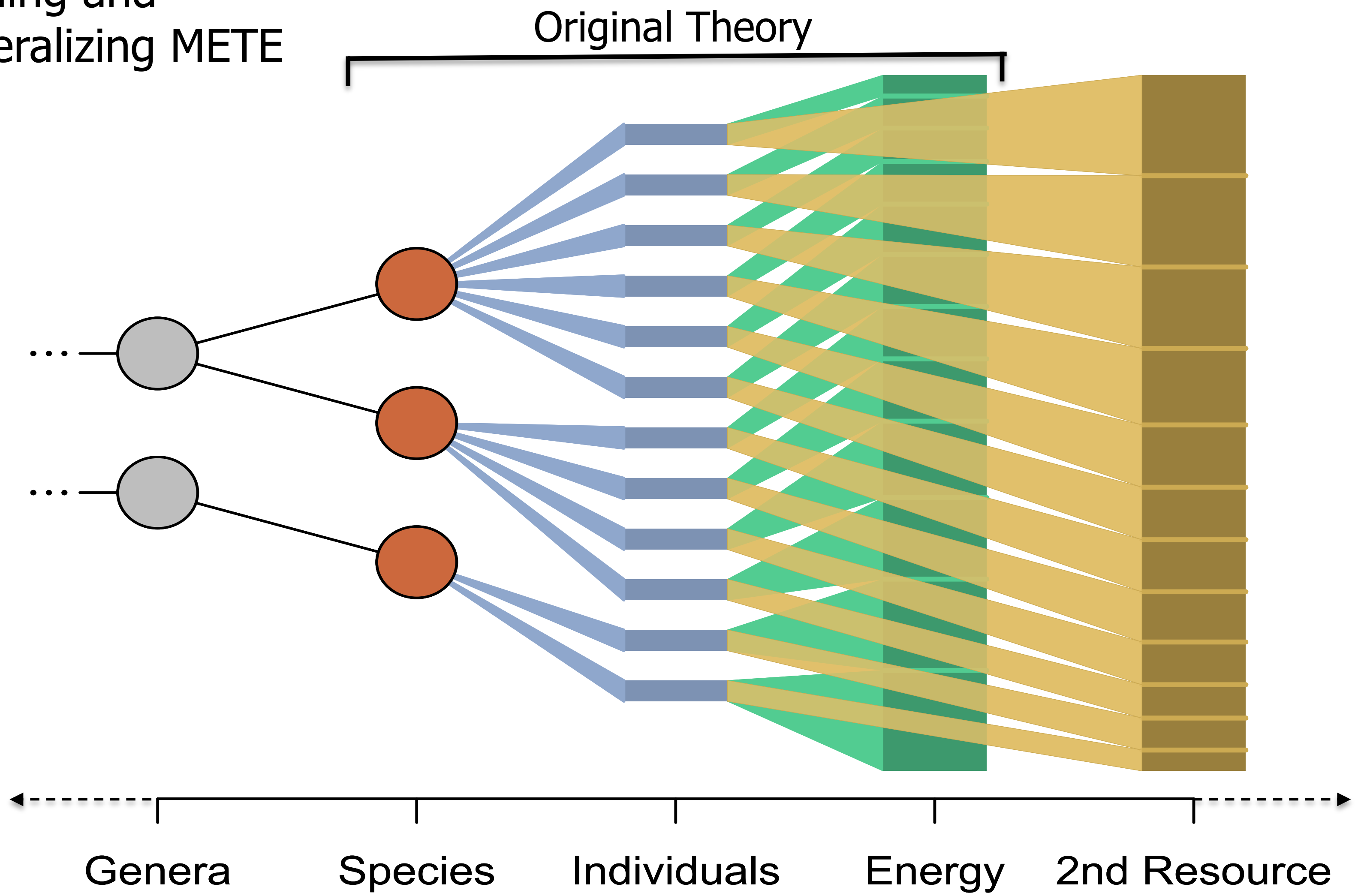
The total metabolic rate of all individuals in each species varies inversely with the species richness of the family that the species is in





Data from Marquet et al., Science 1990

Extending and Generalizing METE



← Alters size-abundance distribution

→ Alters predicted rarity

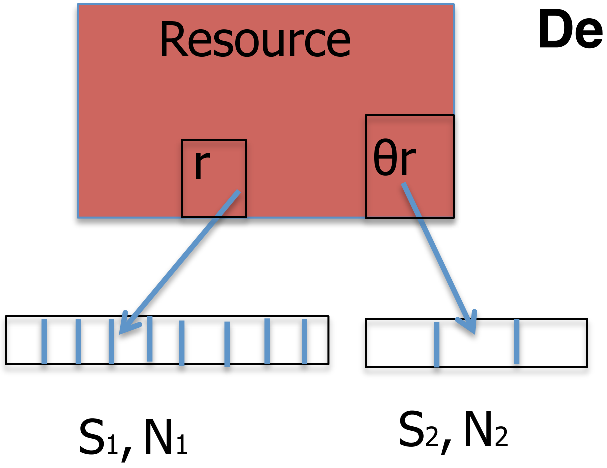
Including additional resource constraints (in addition to energy, *E*)

The log-series SAD becomes:

$$\Phi(n) \sim \frac{e^{-\lambda n}}{n^r}$$

r **- 1** = # additional resources

The inclusion of additional resource constraints predicts increased rarity



Demographics and Species Co-existence Revisited

a. Conventional approach: $b_i - d_i \underline{=} f_i(N_i)$

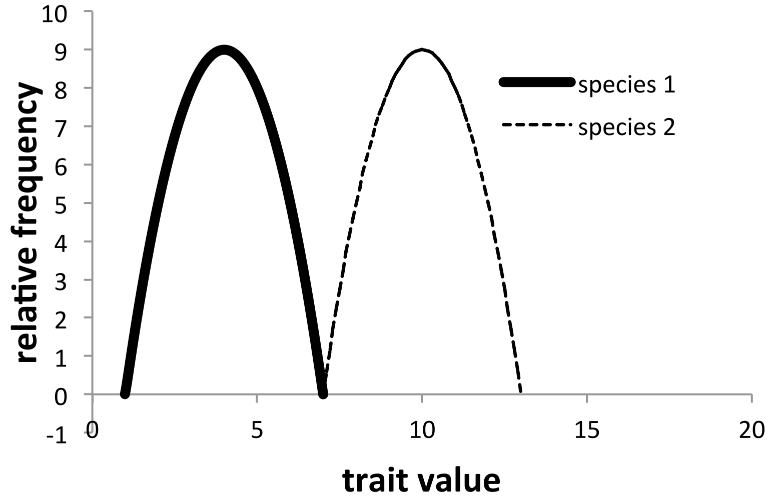
(in Lotka-Volterra eqs, $a_{ii} > a_{ij}$)

b. Zhang and Harte: use **$S = k \log(W)$**

$W = W(D, \theta)$: maximize at each time step

A parameter **θ (surrogate for body size) determines the relative resource requirement per unit time per individual.**

$$D = \frac{\text{distinguishability of individuals within species}}{\text{distinguishability of individuals between species}}$$



W = # of possible allocations of resource units consistent with a macrostate (S species, each with n_i individuals)

$$= W_{\text{between}} \times (W_{\text{within}})^D$$

(W's from combinatorics)

Maximize log($W(D, \theta)$) at successive time steps

RESULTS I

i. $b_i - d_i = f_i(N_i, \theta, D)$; f_i predicted by maximizing log(W)

Density dependence is an emergent property of
S = k log(W), not an imposed assumption.

ii. A pair of species can coexist iff G(θ, D) > 0; G predicted by maximizing W

Under predicted conditions, two species can coexist on one resource because there are more ways to allocate the resource if the species co-exist than if one drives the other to extinction.

Coexistence possible only if:
the intra-specific distinguishability of individuals is sufficiently less than the inter-specific distinguishability of individuals

Results II
Steady state

$$\text{Steady state abundance}_i = \frac{2}{e}(C\theta_i)^{\frac{1}{D_{r,i}-1}}$$

Constraint constant

Individual distinguishability

Resource requirement

Given

- $D_{r,i}$, body size abundance relationship
  - $D_{r,i} = 0$ for all i => energetic equivalence

- $D_{r,l,}$ θ_i, Species abundance distribution (SAD)
  - The higher the D_r, the more skewed the SAD

So what governs ecological complexity?

An extreme, but I think, correct view:

Ecosystems are
what they are
because anything else
is highly improbable

Energy Equivalence:

<metabolic rate> ~ 1/abundance

The ASNE theory predicts this rule:

The within-species distribution of metabolic rates:

$$\Theta(\varepsilon \mid n) = \frac{R}{\Phi} = \lambda_2 n e^{-\lambda_2 n(\varepsilon - 1)}$$

n = abundance
of species

➔ Total energy requirement of a species with abundance n:

$\int n\varepsilon\, \theta(\varepsilon|n)d\varepsilon = n <\varepsilon> = n + 1/\lambda_2 \sim 1/\lambda_2$ **for most species**

➔ Species obey energy equivalence if:

$$n << 1/\lambda_2 = (E_0 - N_0)/S_0 .$$

What else besides DynaMETE is going on?

1. An Ideal Biodiversity Law?

Analogous to PV = nRT, what relationship exists among observed values of State Variables S, N E, …? (Rominger, Merow, Harte)

2. PhyloMaxEnt:

Can the shape (topology and geometry) of phylogenies be derived from MaxEnt? What constitutes useful State Variables? (Rominger)

3. Transition rates from Boltzmann dynamics:

Deriving ontogenic growth rates, demographics, and interaction networks by maximizing S = log(W), where W is the number of resource allocations.

(Zhang and Harte, *Theo. Pop. Biol.* 2015; Zhang PhD thesis 2016)

4. Higher order spatial statistics

Predicting spatial correlations from MaxEnt-derived point processes

Extending METE from Static to Dynamic

Static systems :

***MaxEnt* adequately predicts the form of many of the metrics of macroecology**

Lacking is theory describing Rates of Change in these Metrics during the Processes of:

- **Speciation and Extinction**
- **Succession**
- **Adaptive Responses (e.g., to "global change")**

Possible approaches:

1. **Maximum Entropy Production**
2. **Non-extensive entropy**
3. **Dynamic, stochastic theory of state variables: use "master equation" incorporating dominant mechanisms as transition probabilities**
4. **Maximum resource allocation entropy**

Extension of METE to higher taxonomic levels

Example: inclusion of genus as a category (AGSNE)

State Variables:

G_0 = # genera
S_0 = # species
N_0 = # individuals
E_0 = total metabolic rate

The probability function *Q* replaces *R*

$Q(m,n,\varepsilon|G_0, S_0, N_0, E_0)$, defined as follows:

Pick a genus;
***Q* is the probability it has *m* species**
and if you pick one of those species from that genus,
that it has *n* individuals,
and that if you pick one of those individuals from that species,
that it has metabolic rate ε.

The constraints:

$$< m >= \frac{S_0}{G_0} = \sum_{m,n,\varepsilon} mQ$$

$$< n_G >= \frac{N_0}{G_0} = \sum_{m,n,\varepsilon} mnQ$$

$$< \varepsilon_G >= \frac{E_0}{G_0} = \sum_{m,n,\varepsilon} mn\varepsilon Q$$

Now, with AGSNE, we can predict the “old” metrics that the ASNE theory predicts:

Species abundance distribution

Species-area relationship

Endemics-area relationship

Energy distribution over individuals and species

And also some new metrics:

The distribution of species over genera

The genus-area relationship

Distribution of abundances and metabolic rates over species within a genus with m species.

Solutions of AGSNE:

$$Q(m,n,\varepsilon) = \frac{1}{Z(\lambda_1,\lambda_2,\lambda_3)} e^{-\lambda_1 m} e^{-\lambda_2 mn} e^{-\lambda_3 mn\varepsilon}$$

Master Distribution

$$< m >= \frac{S_0}{G_0} \approx \frac{1}{\lambda_1 \ln(\lambda_1{}^{-1})} \qquad < n_G >= \frac{N_0}{G_0} \approx \frac{1}{\beta \ln(\beta^{-1})} \qquad \lambda_3 = \frac{G_0}{E_0 - N_0}$$

Determining Lagrange multipliers ($\beta=\lambda_1+\lambda_2$)

$$\Gamma(m) \approx \frac{e^{-\lambda_1 m}}{m \log(\lambda_1{}^{-1})}$$

Distribution of Species over Genera

$$\Phi(n) \cong \frac{\lambda_1 \cdot e^{-(\lambda_1+\beta n)}}{n \ln(\beta^{-1})(1-e^{-(\lambda_1+\beta n)})}$$

Distribution of Individuals over Species

$$\Psi(\varepsilon) \approx \frac{\beta\lambda_3 \cdot \ln(\lambda_1 + \gamma(\varepsilon))}{\gamma^2(\varepsilon) \ln(\beta + \lambda_1)}$$

Distribution of Metabolic Rates over individuals

($\gamma(\varepsilon) = \lambda_2 + \lambda_3\varepsilon$)

$$\Theta(\varepsilon \mid m,n) = \lambda_3 mne^{-\lambda_3 mn(\varepsilon-1)}$$

Distribution of metabolic rates in species with n individuals in a genus with m species