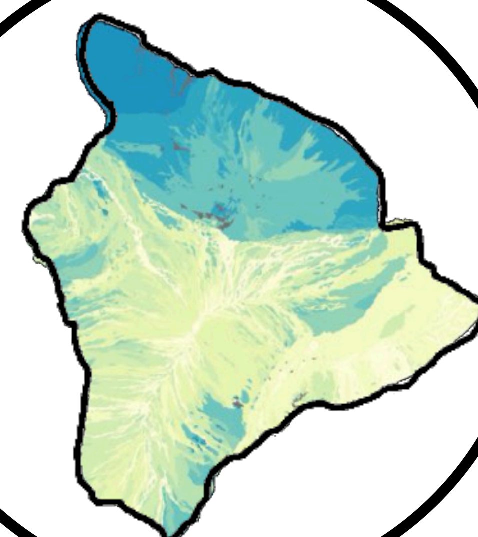
Using mathematics to gain insight into the biodiversity of Hawai‘i

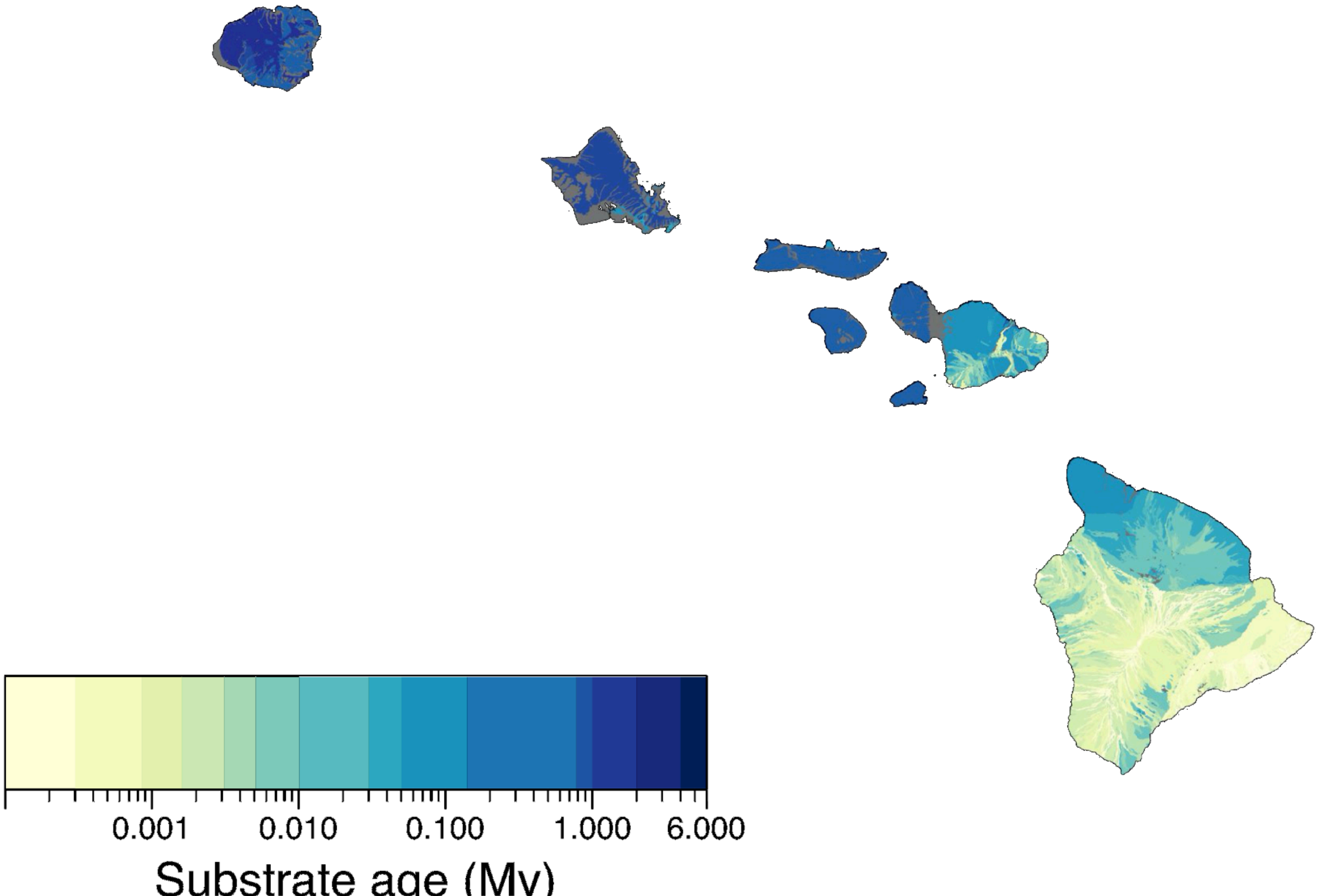
$$\frac{1}{Z}e^{-(\lambda_1 n+\lambda_2 n\varepsilon)}$$

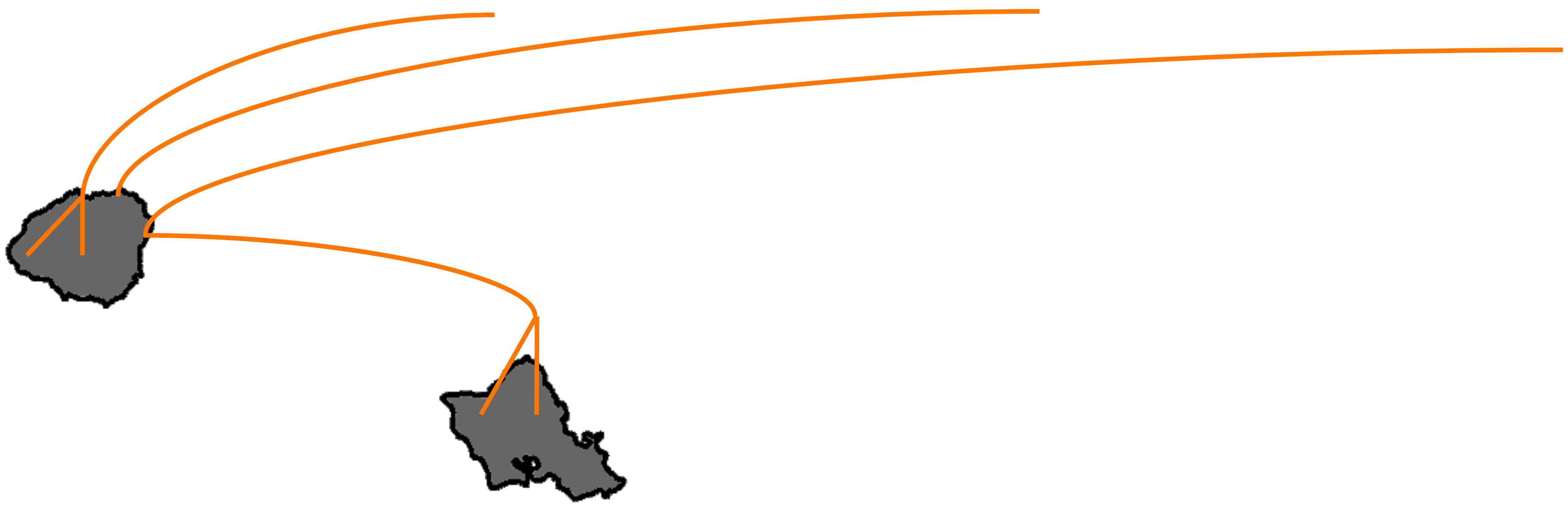


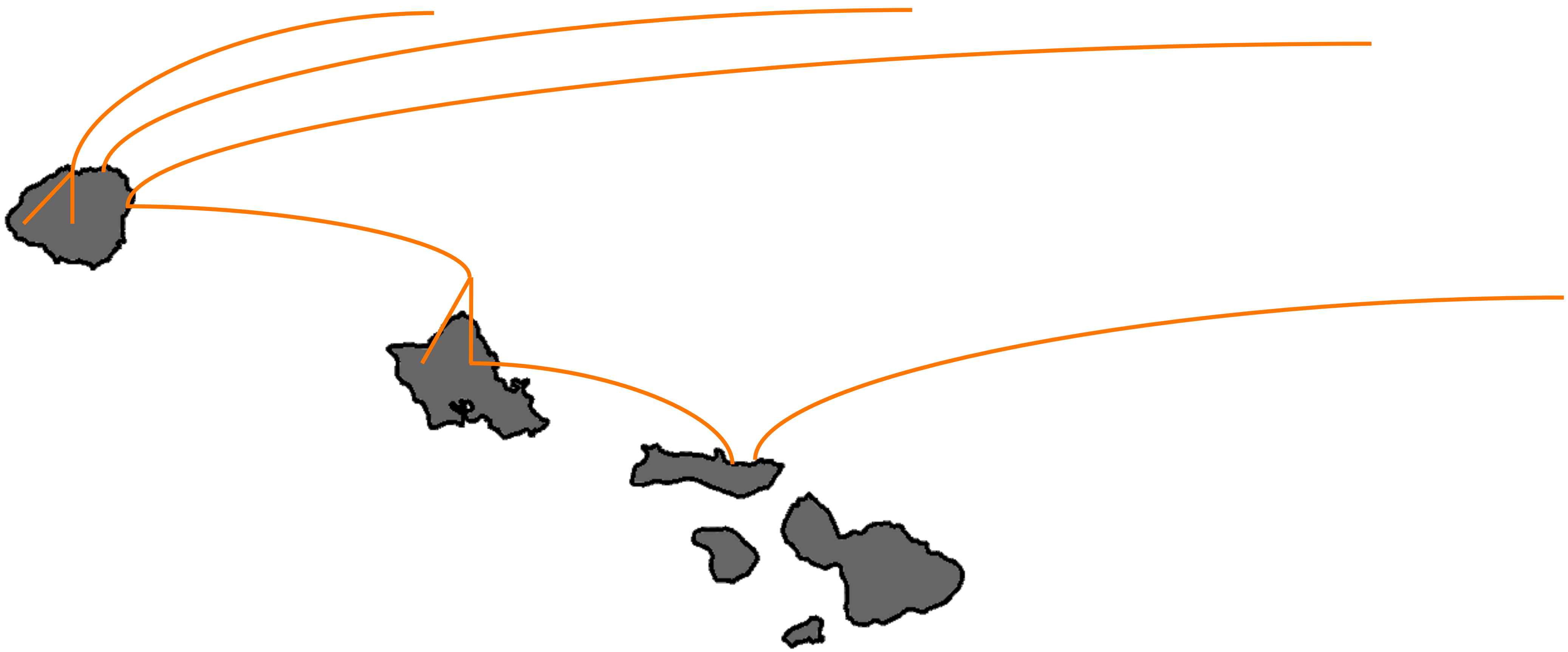
Andy Rominger

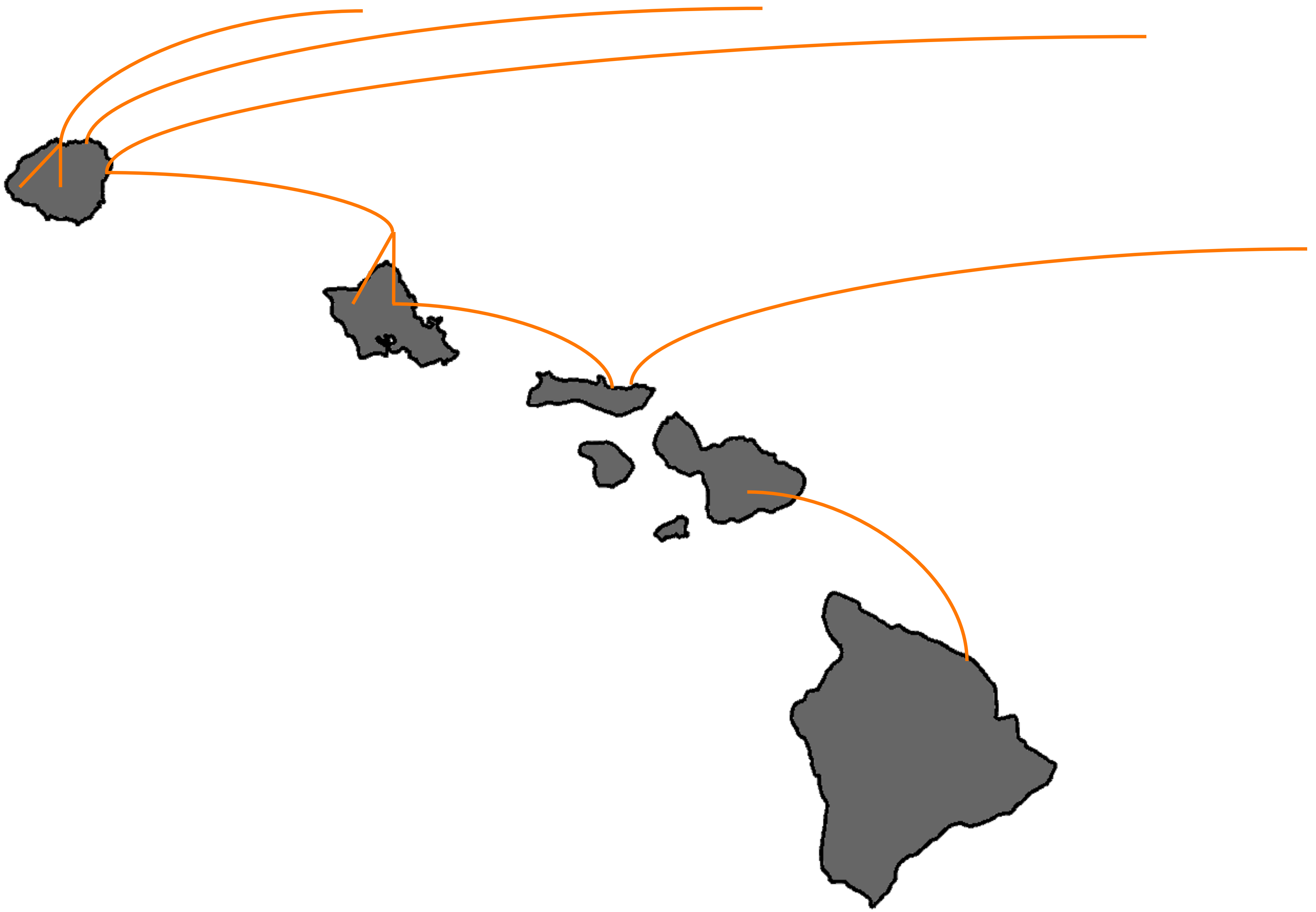
ajrominger.github.io

University of Hawai‘i at Hilo • 22 April 2020


0.001
0.010
0.100
1.000
6.000
Substrate age (My)

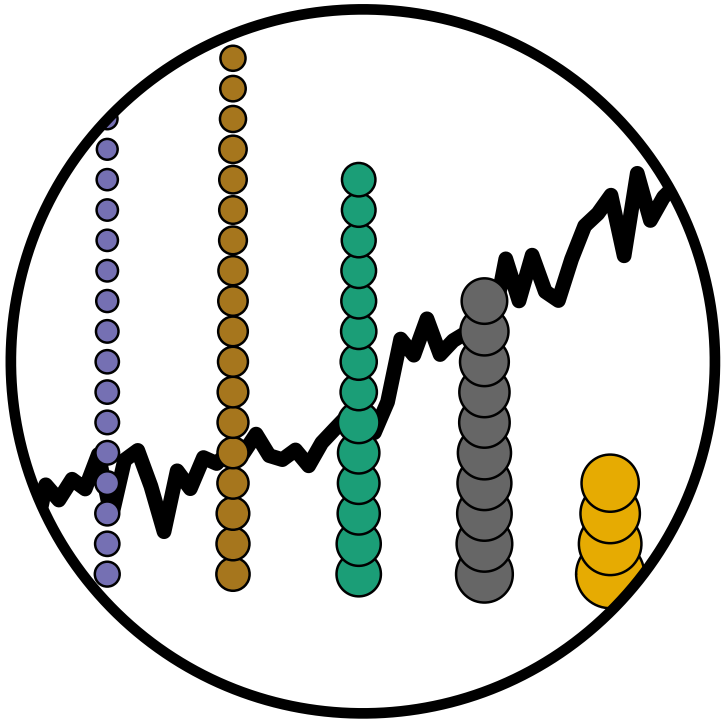




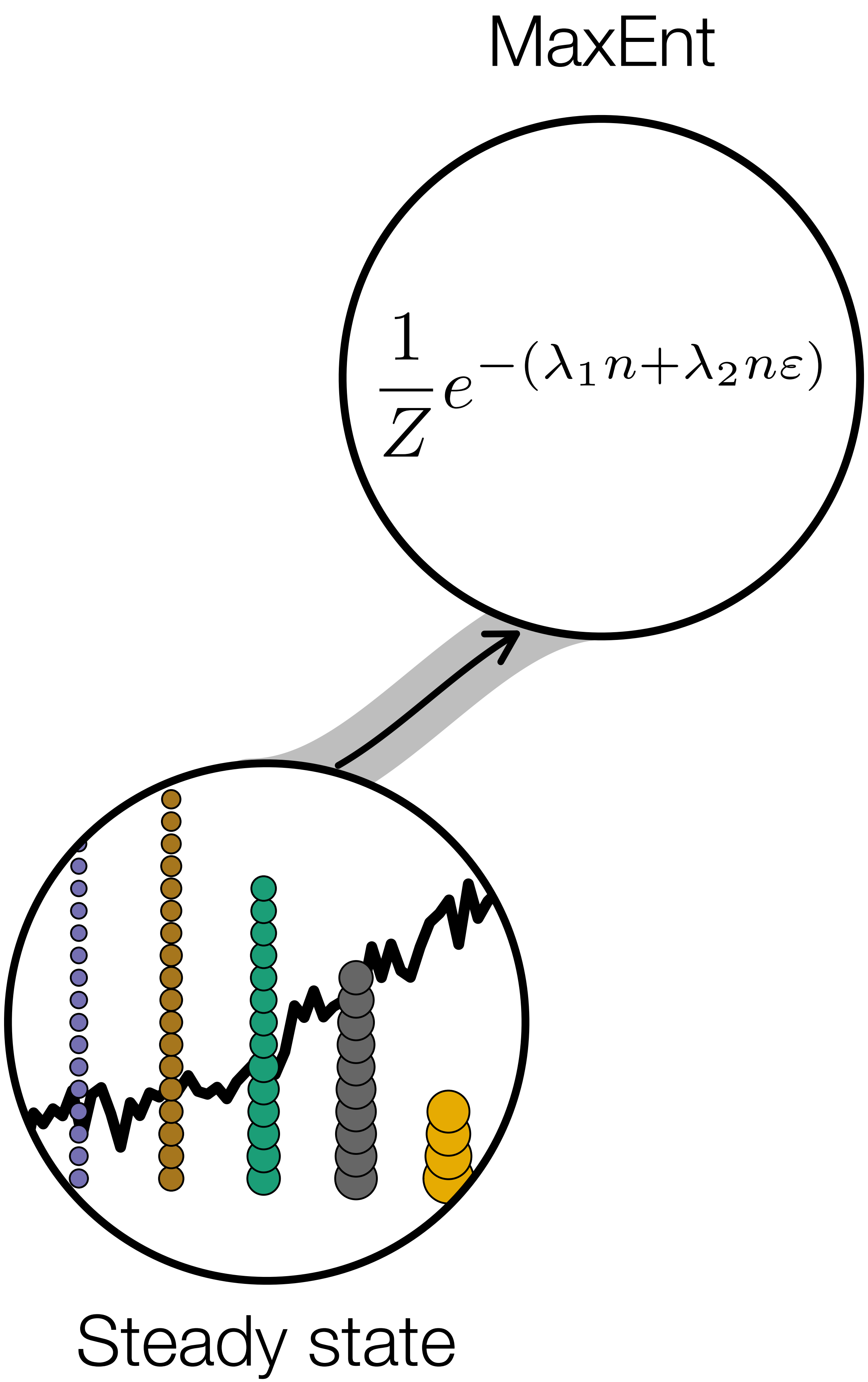



How could we represent these processes with mathematics?

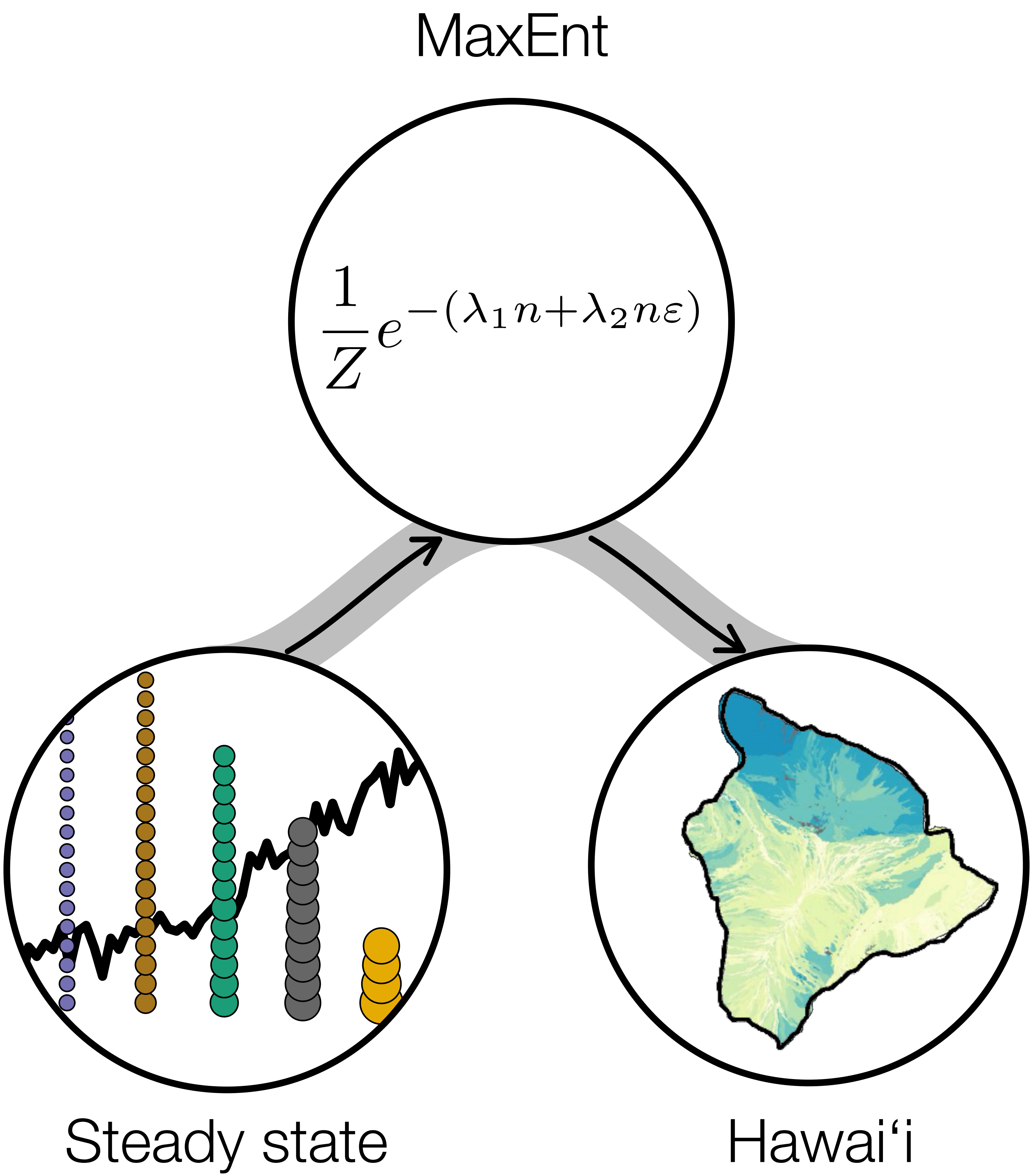

Do biological communities reach steady state?
If so why?
If not why?

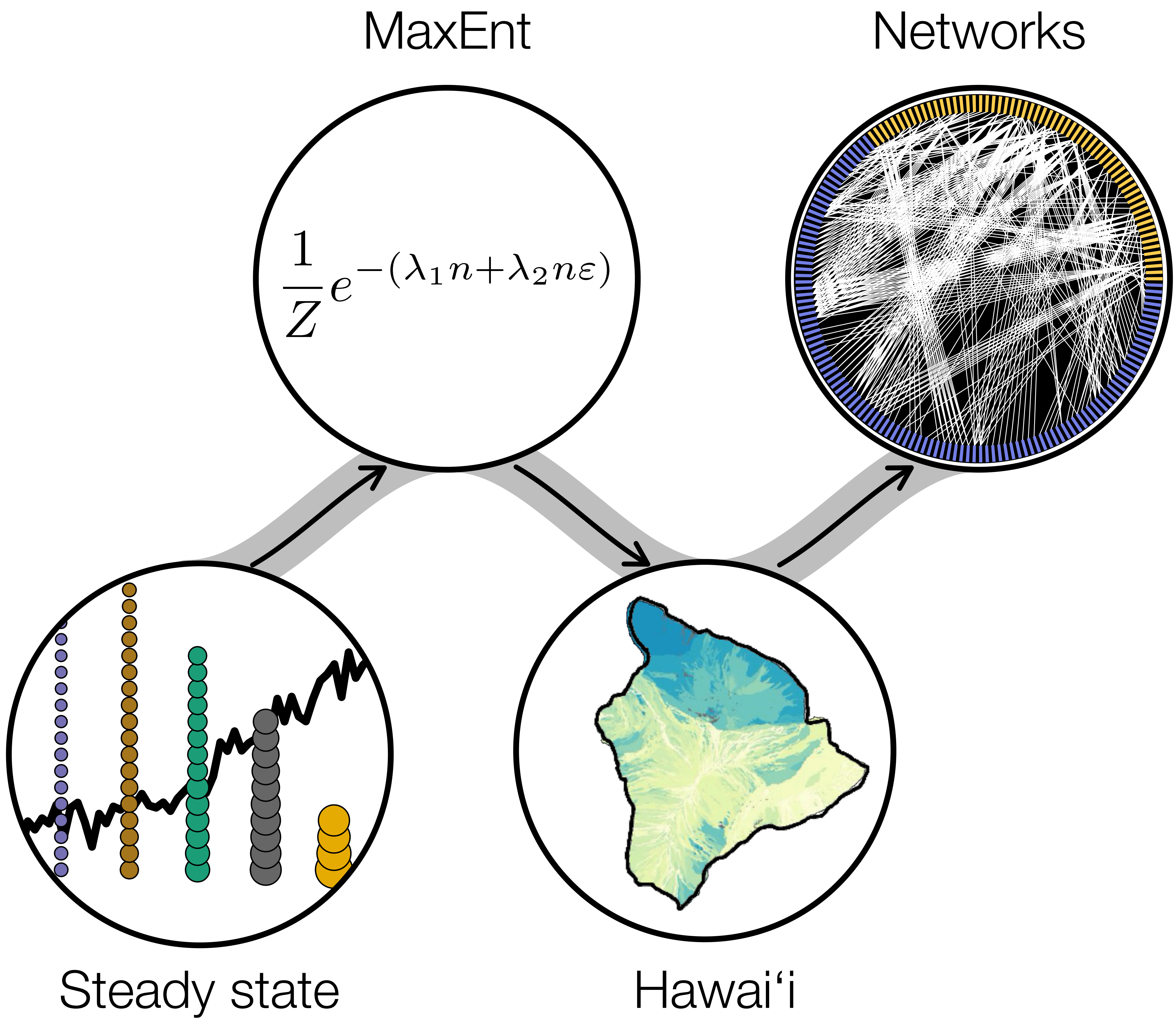


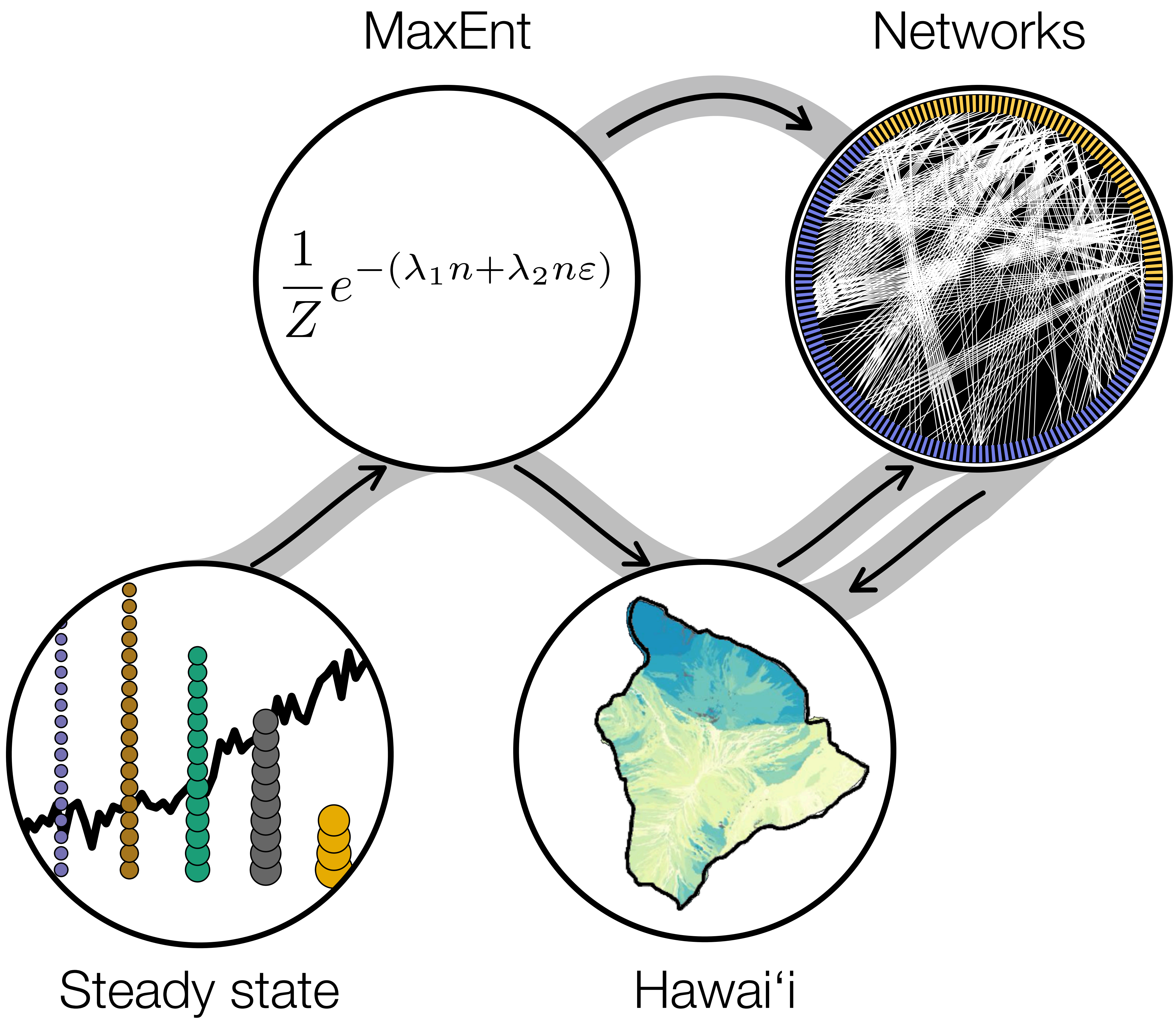
Steady state

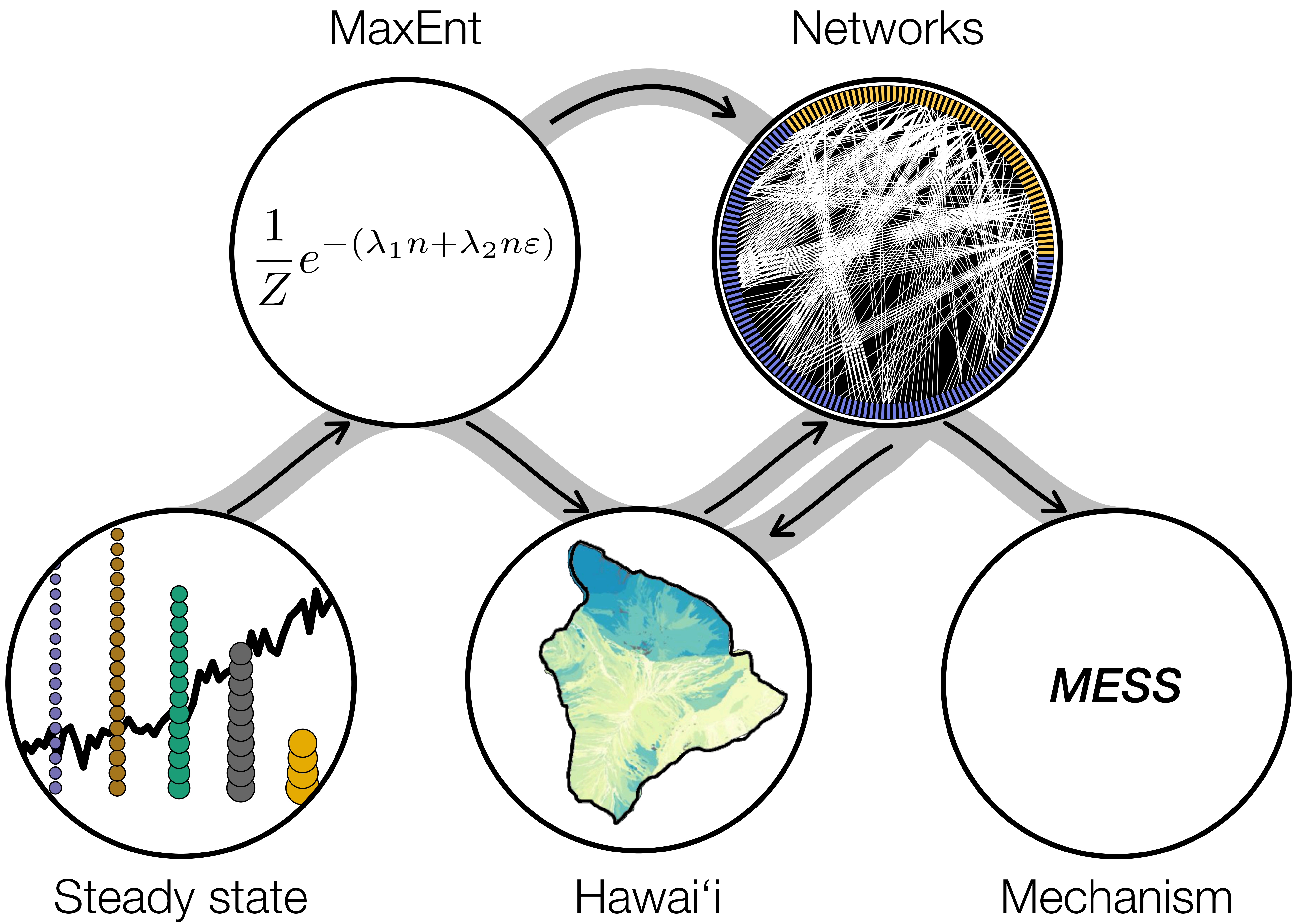


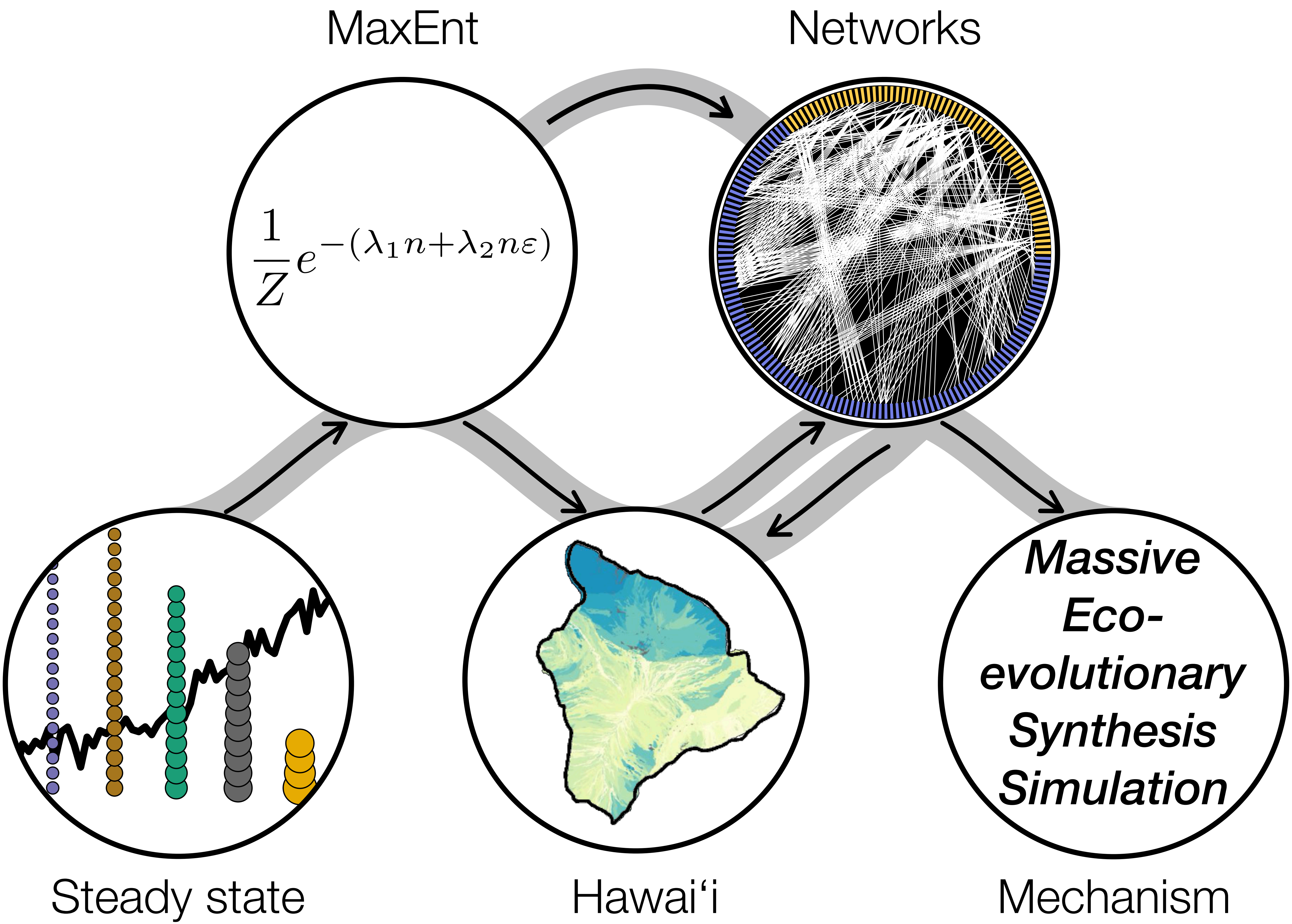
MaxEnt
$$\frac{1}{Z}e^{-(\lambda_1 n+\lambda_2 n\varepsilon)}$$
Steady state

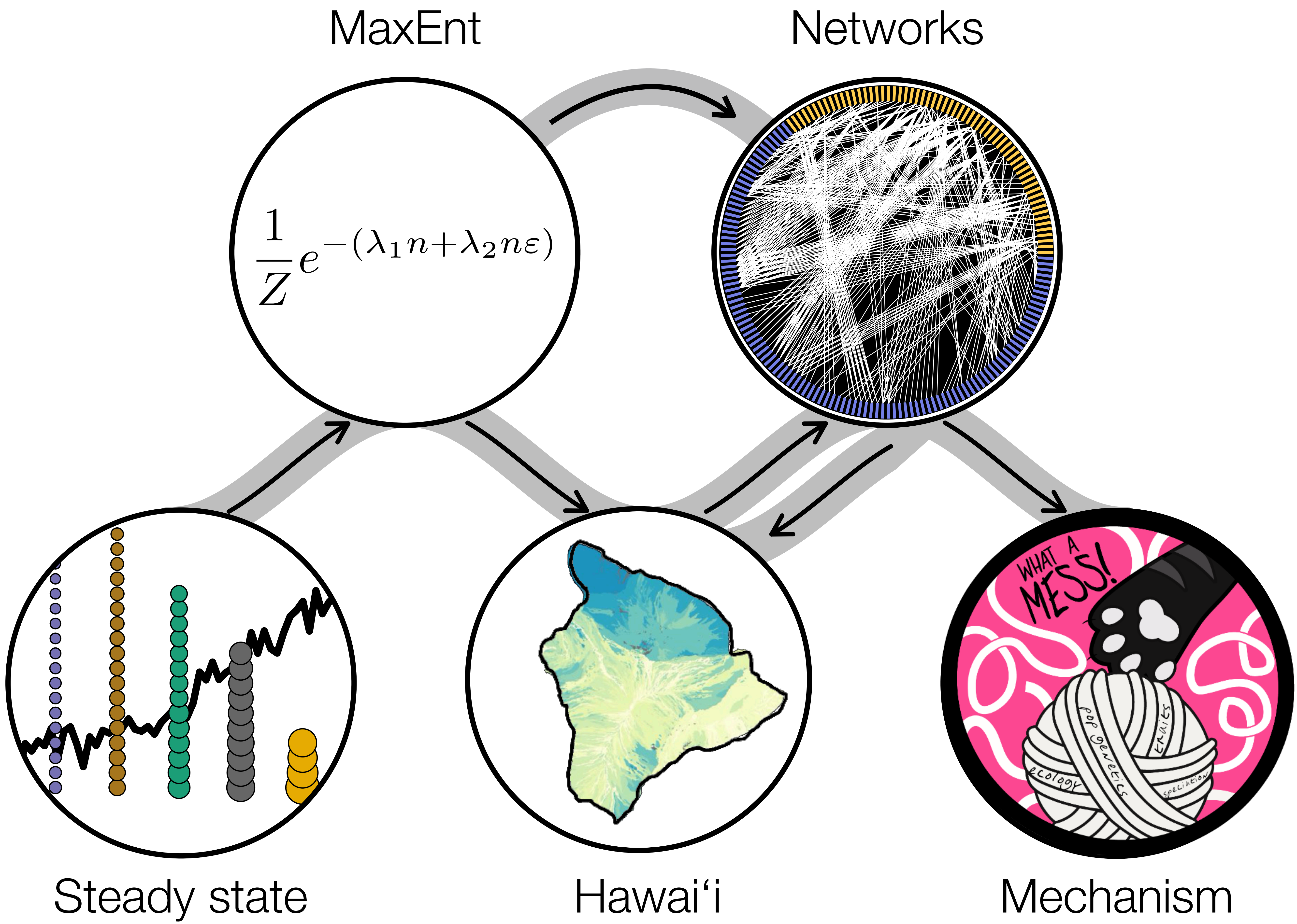

MaxEnt
$\frac{1}{Z}e^{-(\lambda_1 n+\lambda_2 n\varepsilon)}$
Steady state
Hawai‘i


MaxEnt
Networks
$\frac{1}{Z}e^{-(\lambda_1 n+\lambda_2 n\varepsilon)}$
Steady state
Hawai‘i


MaxEnt
Networks
$\frac{1}{Z}e^{-(\lambda_1 n+\lambda_2 n\varepsilon)}$
Steady state
Hawai‘i


MaxEnt
Networks
$\frac{1}{Z}e^{-(\lambda_1 n+\lambda_2 n\varepsilon)}$
MESS
Steady state
Hawai‘i
Mechanism


MaxEnt
Networks
$\frac{1}{Z}e^{-(\lambda_1 n+\lambda_2 n\varepsilon)}$
Massive
Eco-
evolutionary
Synthesis
Simulation
Steady state
Hawai‘i
Mechanism


MaxEnt
Networks
$\frac{1}{Z}e^{-(\lambda_1 n+\lambda_2 n\varepsilon)}$
WHAT A MESS!
pop genetics
traits
ecology
speciation
Steady state
Hawai‘i
Mechanism

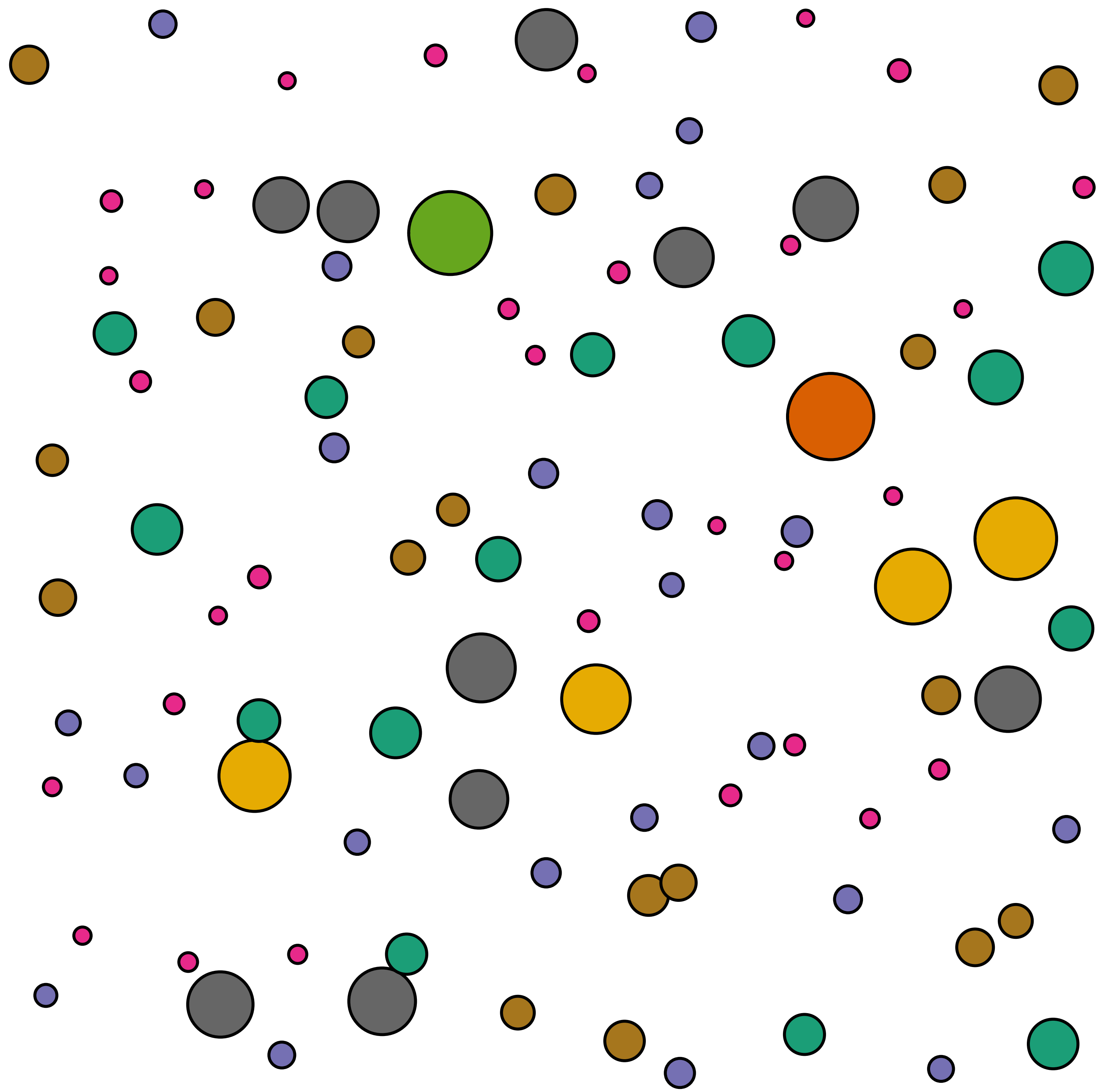

Do biological communities reach steady state?
If so why?
If not why?

How to quantify steady state?

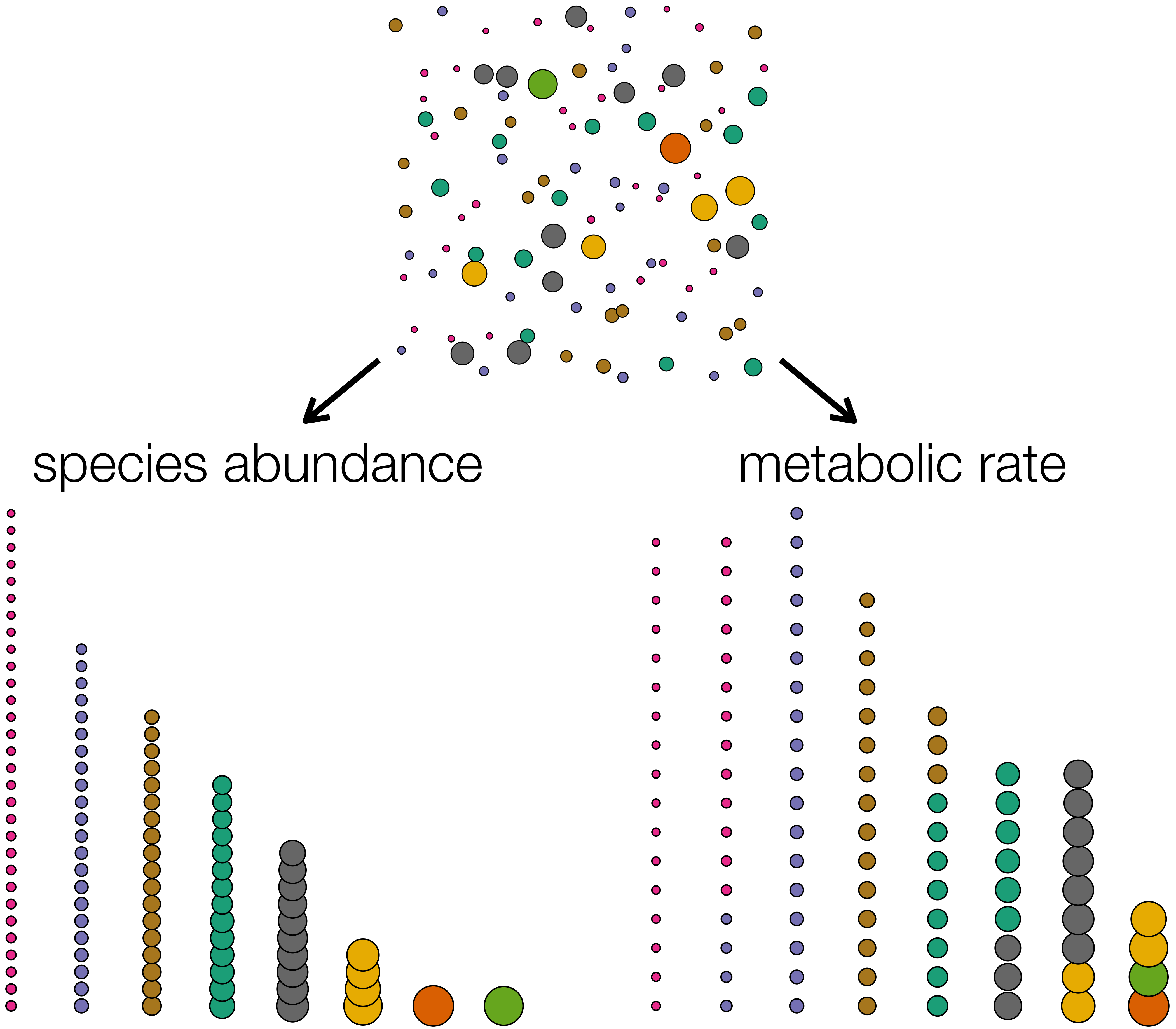
What metrics do we measure?

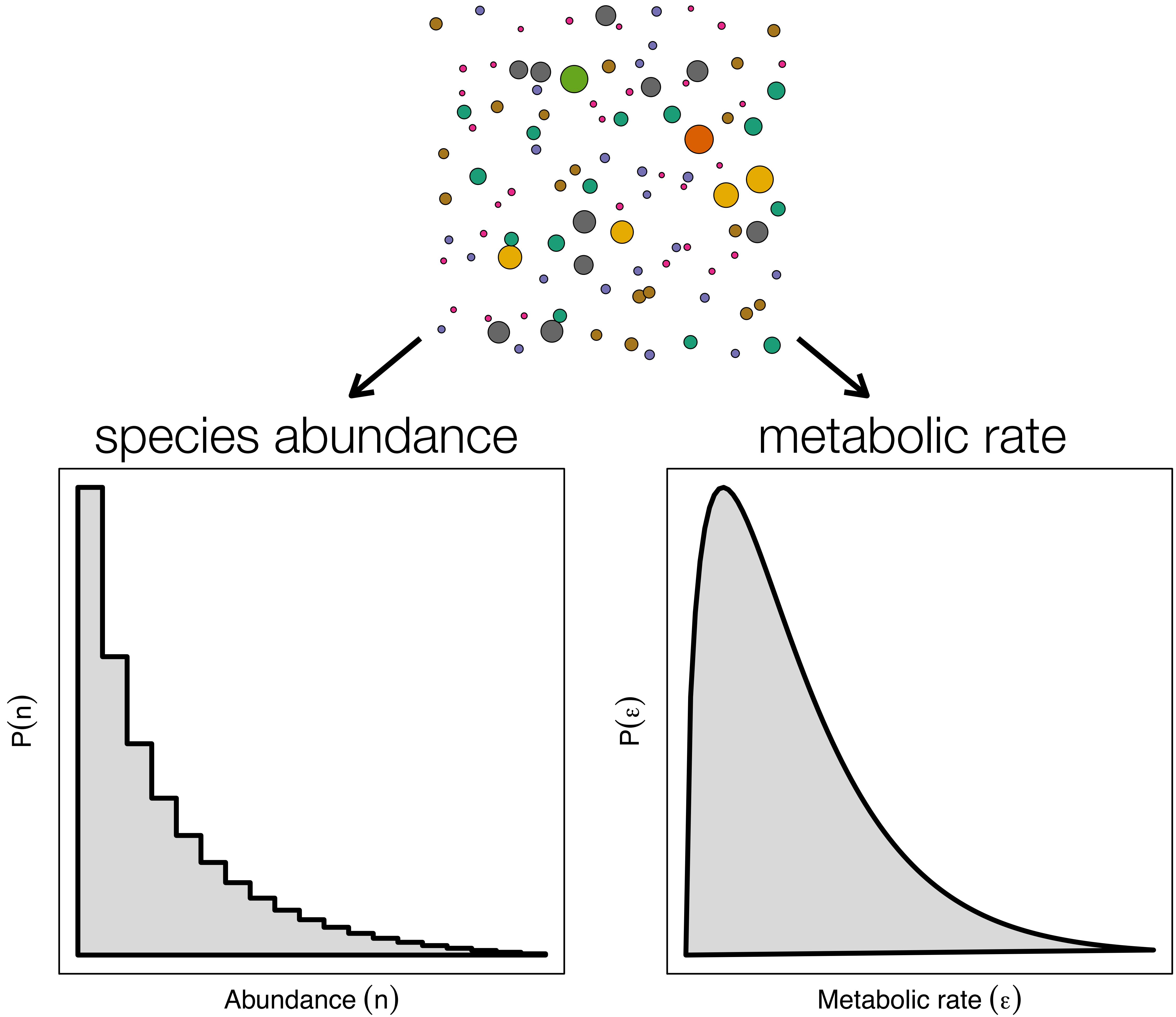
How to quantify steady state?

<u>What metrics do we measure?</u>



color = species size = metabolism


species abundance
metabolic rate


species abundance
metabolic rate
P(n)
Abundance (n)
P(ε)
Metabolic rate (ε)

How to quantify steady state?

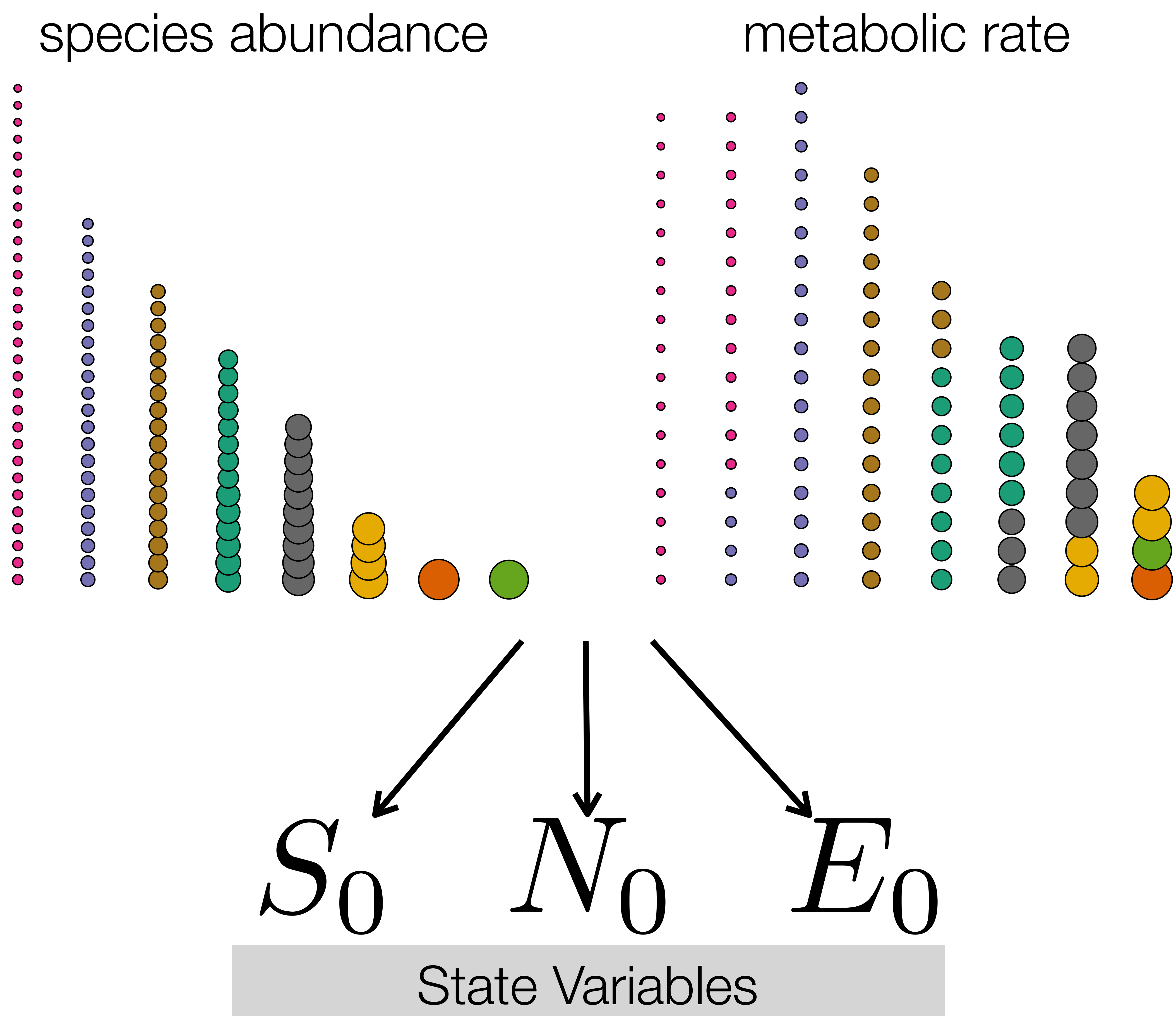
What metrics do we measure?

- species abundance distribution
- metabolic rate distribution

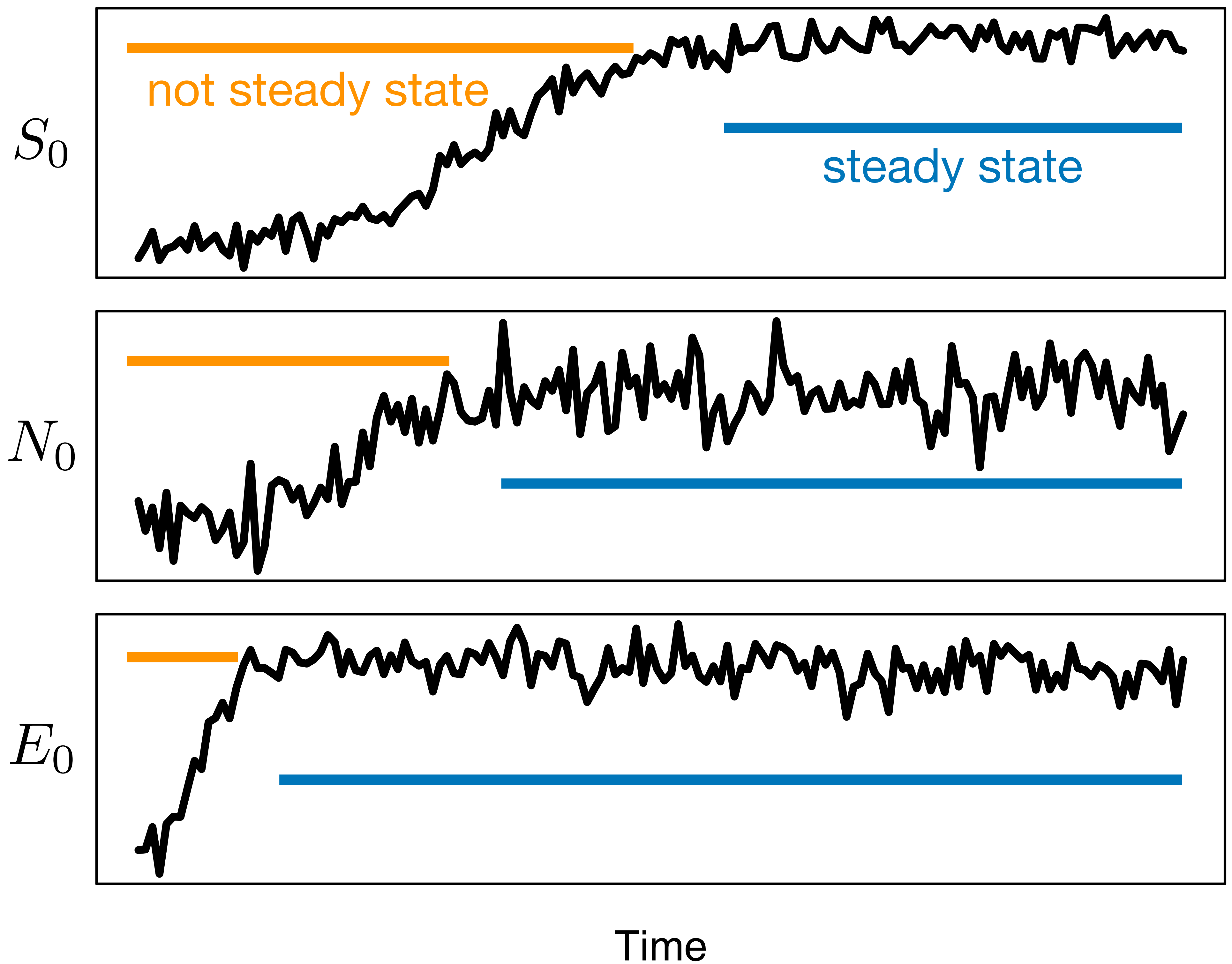
How to quantify steady state?

What metrics do we measure?

- species abundance distribution
- metabolic rate distribution

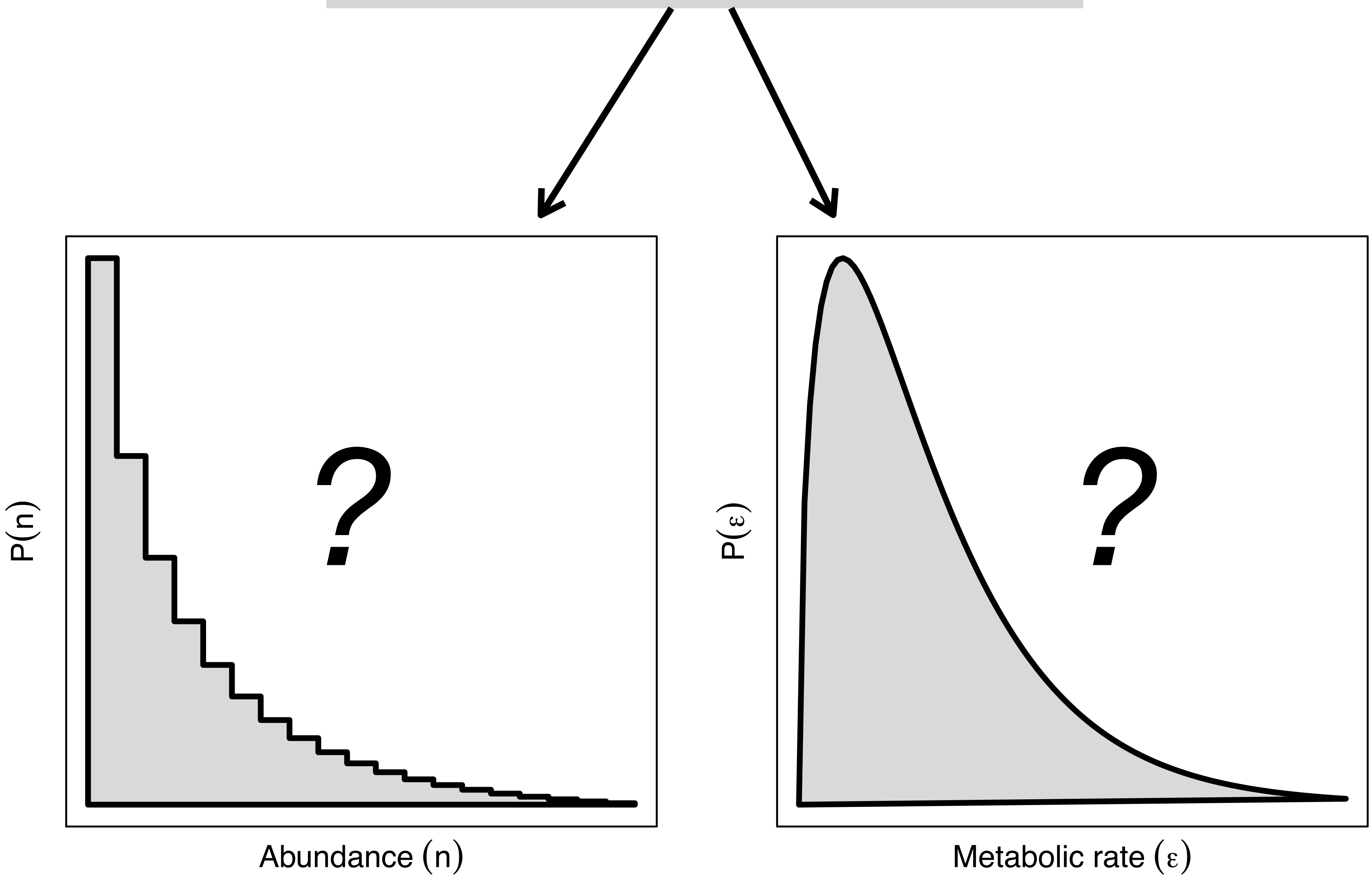

species abundance
metabolic rate
S_0
N_0
E_0
State Variables

Steady state means stationary state variables

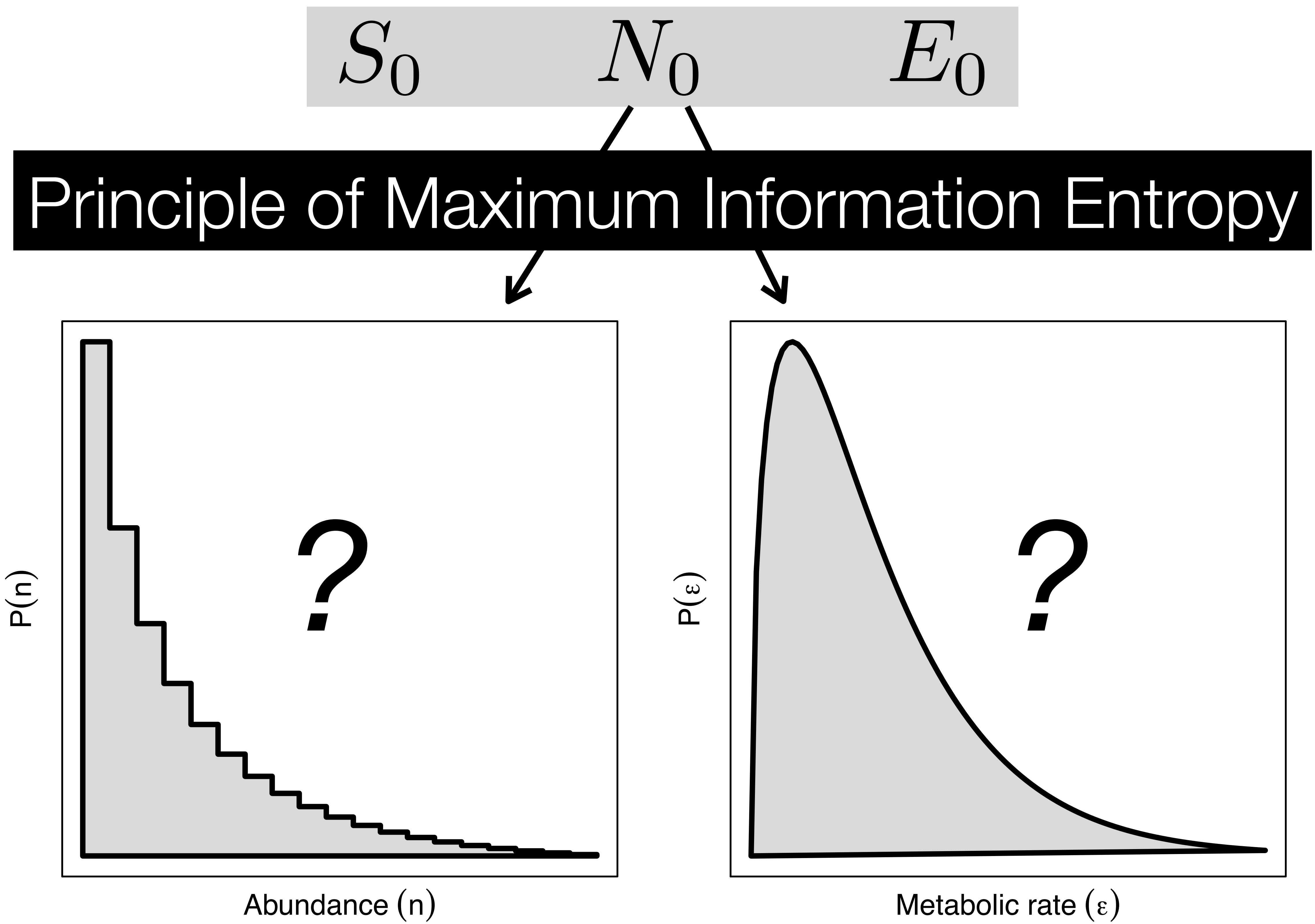


How do we derive *P(n)* and *P(ε)*?

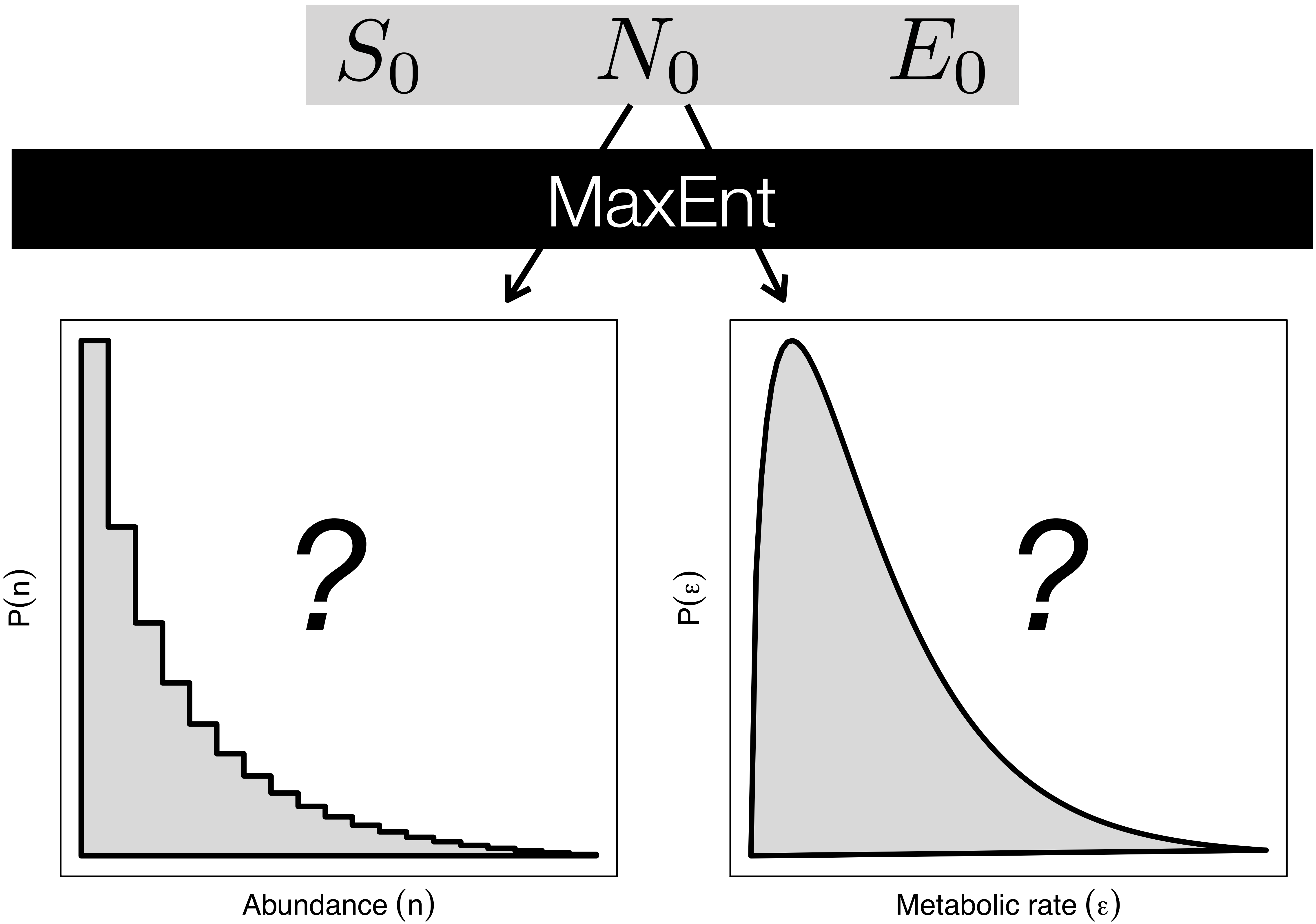
$S_0 \qquad N_0 \qquad E_0$



How do we derive *P(n)* and *P(ε)*?



How do we derive *P(n)* and *P(ε)*?



MaxEnt

Known info

State Variables

MaxEnt

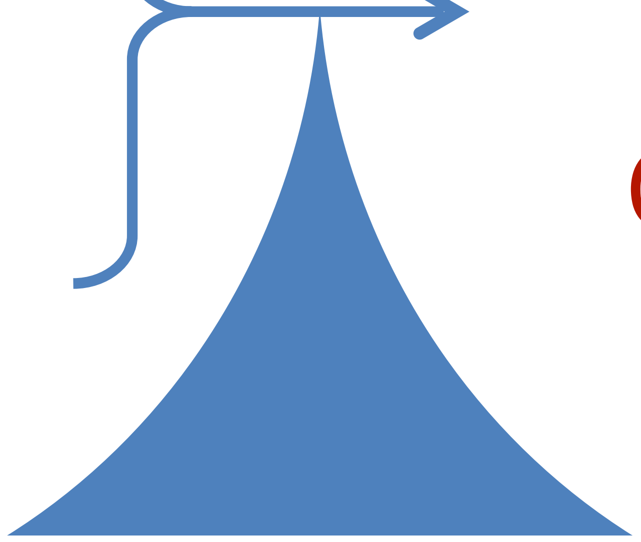
Known info
State Variables

System
probability
distributions

MaxEnt

Known info
State Variables

System
probability
distributions

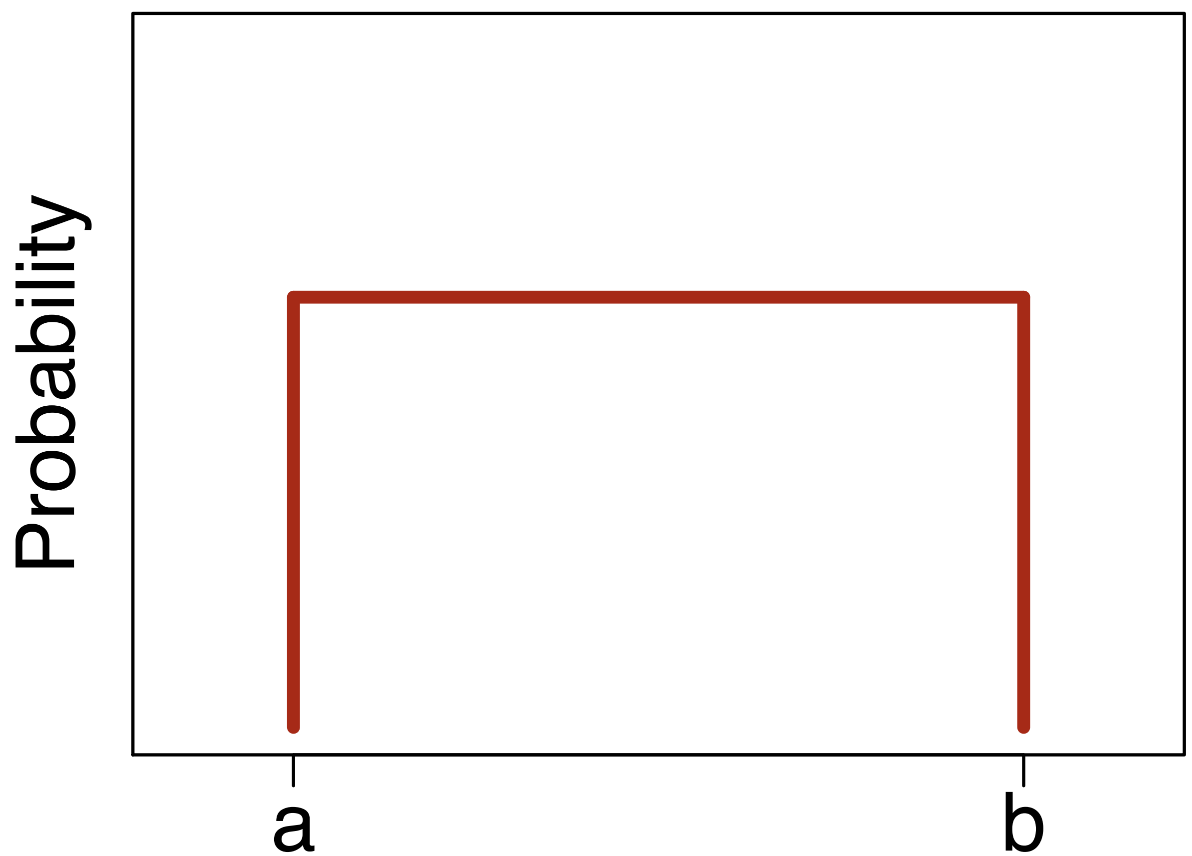


with *no other* info!

MaxEnt

Known info
State Variables

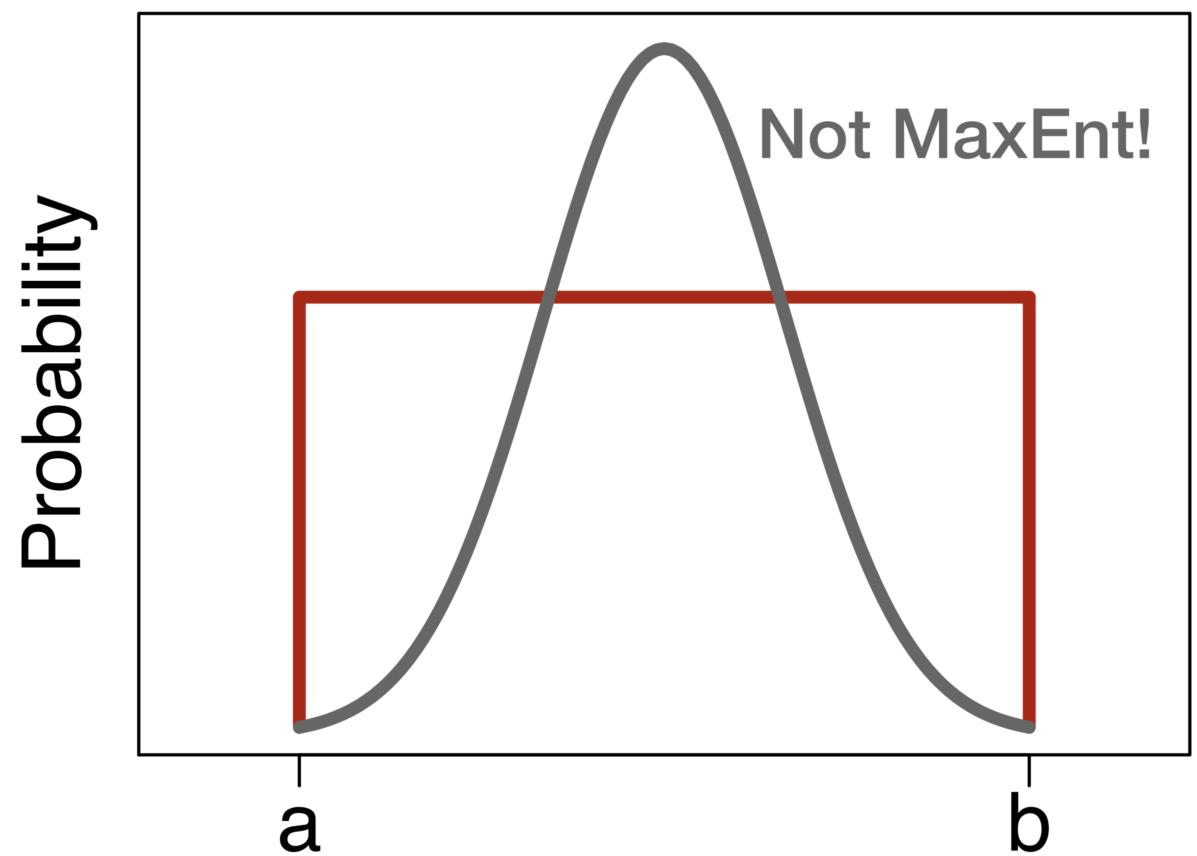
System
probability
distributions



MaxEnt

Known info
State Variables

System
probability
distributions



MaxEnt

Information entropy: $H_P(X) = -\sum_{x \in \mathcal{X}} P(x) \log P(x)$

MaxEnt

Information entropy: $$H_P(X) = -\sum_{x \in \mathcal{X}} P(x) \log P(x)$$

0 1 1 0 0 1 0 1 $$H = -\left(\frac{1}{2}\log\left(\frac{1}{2}\right) + \frac{1}{2}\log\left(\frac{1}{2}\right)\right)$$

MaxEnt

Information entropy: $H_P(X) = -\sum_{x \in \mathcal{X}} P(x) \log P(x)$

0 1 1 0 0 1 0 1 $H = 1 \text{ bit}$

MaxEnt

Information entropy: $H_P(X) = -\sum_{x \in \mathcal{X}} P(x) \log P(x)$

0 1 1 0 0 1 0 1 $\quad H = 1 \text{ bit}$

1 1 1 1 1 0 1 1 $\quad H = -\left(\frac{7}{8}\log\left(\frac{7}{8}\right) + \frac{1}{8}\log\left(\frac{1}{8}\right)\right)$

MaxEnt

Information entropy: $H_P(X) = -\sum_{x \in \mathcal{X}} P(x) \log P(x)$

0 1 1 0 0 1 0 1 $H = 1 \text{ bit}$

1 1 1 1 1 0 1 1 $H = 0.543 \text{ bit}$

MaxEnt

Information entropy: $H_P(X) = -\sum_{x \in \mathcal{X}} P(x) \log P(x)$

0 1 1 0 0 1 0 1 $H = 1 \text{ bit}$ more uncertain

1 1 1 1 1 0 1 1 $H = 0.543 \text{ bit}$ less uncertain

MaxEnt

Information entropy: $H_P(X) = -\sum_{x \in \mathcal{X}} P(x) \log P(x)$

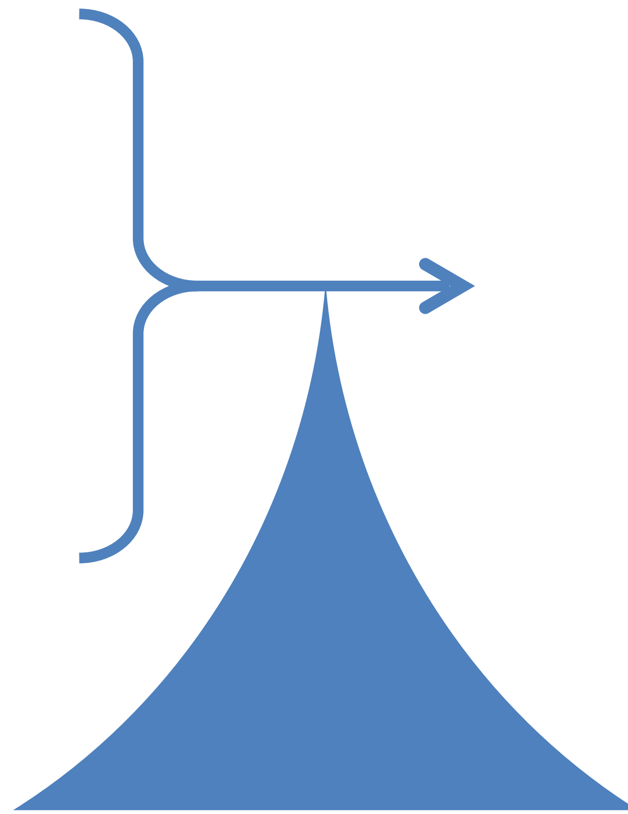
0 1 1 0 0 1 0 1 $H = 1 \text{ bit}$ less info

1 1 1 1 1 0 1 1 $H = 0.543 \text{ bit}$ more info

MaxEnt

Information entropy: $H_P(X) = -\sum_{x \in \mathcal{X}} P(x) \log P(x)$

Known info
State Variables



System
probability
distributions

with *no other* info!

MaxEnt

Information entropy: $H_P(X) = -\sum_{x \in \mathcal{X}} P(x) \log P(x)$

Known info

State Variables

System

probability

distributions

$\boxed{\text{Find } P(x) \text{ that maximizes } H_P(X)}$

MaxEnt…how?

Express state variables as constraints

$$\sum f_1(x)P(x) = c_1$$

$$\sum f_2(x)P(x) = c_2$$

$$\cdots$$

$$\sum f_k(x)P(x) = c_k$$

MaxEnt…how?

Express state variables as constraints

$$\sum f_1(x)P(x) = c_1$$

$$\sum f_2(x)P(x) = c_2$$

$$\dots$$

$$\sum f_k(x)P(x) = c_k$$

Using the method of Lagrange multipliers…

$$P(x) = \frac{1}{Z} e^{-\sum_{i=1}^{k} \lambda_i f_i(x)}$$, λ_i are Lagrange multipliers

MaxEnt and biodiversity

We're interested in a joint distribution

$$R(n, \varepsilon)$$

species abundance   individual metabolic rate

Harte et al. (2008) *Ecology*

MaxEnt and biodiversity

We're interested in a joint distribution

$$R(n, \varepsilon)$$

with constraints

$$\sum_{n \in \mathcal{N}} \int_{\mathcal{E}} R(n, \varepsilon) d\varepsilon = 1$$

$$\sum_{n \in \mathcal{N}} \int_{\mathcal{E}} nR(n, \varepsilon) d\varepsilon = \frac{N_0}{S_0}$$

$$\sum_{n \in \mathcal{N}} \int_{\mathcal{E}} n\varepsilon R(n, \varepsilon) d\varepsilon = \frac{E_0}{S_0}$$

MaxEnt and biodiversity

We're interested in a joint distribution

$$R(n, \varepsilon) = \frac{1}{Z} e^{-\lambda_1 n - \lambda_2 n \varepsilon}$$

with constraints

$$\sum_{n \in \mathcal{N}} \int_{\mathcal{E}} R(n, \varepsilon) d\varepsilon = 1$$

$$\sum_{n \in \mathcal{N}} \int_{\mathcal{E}} n R(n, \varepsilon) d\varepsilon = \frac{N_0}{S_0}$$

$$\sum_{n \in \mathcal{N}} \int_{\mathcal{E}} n \varepsilon R(n, \varepsilon) d\varepsilon = \frac{E_0}{S_0}$$

MaxEnt and biodiversity

We're interested in a joint distribution

$$R(n, \varepsilon) = \frac{1}{Z} e^{-\lambda_1 n - \lambda_2 n \varepsilon}$$

To find Lagrange multipliers solve

$$\frac{\delta \log(Z)}{\delta \lambda_1} = -\frac{N_0}{S_0}$$

$$\frac{\delta \log(Z)}{\delta \lambda_2} = -\frac{E_0}{S_0}$$

Giving: $\frac{N_0}{S_0} = \frac{1}{Z\lambda_2} \sum_{n \in \mathcal{N}} e^{-n(\lambda_1 + \lambda_2)}$ and $\lambda_2 = \frac{S_0}{E_0 - N_0}$

MaxEnt and biodiversity

From this joint distribution

$$R(n, \varepsilon) = \frac{1}{Z} e^{-\lambda_1 n - \lambda_2 n \varepsilon}$$

we derive our distributions of interest

Harte et al. (2008) *Ecology*

MaxEnt and biodiversity

From this joint distribution

$$R(n, \varepsilon) = \frac{1}{Z} e^{-\lambda_1 n - \lambda_2 n \varepsilon}$$

we derive our distributions of interest

$$\boxed{\begin{aligned} P(n) &= \int_{\mathcal{E}} R(n, \varepsilon) d\varepsilon \\ &= \frac{1}{\lambda_2 Z} \frac{e^{-n(\lambda_1 + \lambda_2)}}{n} \end{aligned}}$$

Harte et al. (2008) *Ecology*

MaxEnt and biodiversity

From this joint distribution

$$R(n, \varepsilon) = \frac{1}{Z} e^{-\lambda_1 n - \lambda_2 n \varepsilon}$$

we derive our distributions of interest

$$\boxed{\begin{aligned} P(n) &= \int_{\mathcal{E}} R(n, \varepsilon) d\varepsilon \\ &= \frac{1}{\lambda_2 Z} \frac{e^{-n(\lambda_1 + \lambda_2)}}{n} \end{aligned}}$$

$$\boxed{\begin{aligned} P(\varepsilon) &= \frac{S_0}{N_0} \sum_{n \in \mathcal{N}} n R(n, \varepsilon) \\ &= \frac{S_0}{N_0 Z} \frac{e^{-(\lambda_1 + \lambda_2 \varepsilon)}}{\left(1 - e^{-(\lambda_1 + \lambda_2 \varepsilon)}\right)^2} \end{aligned}}$$

Harte et al. (2008) *Ecology*

Package *meteR*

Input | **Core Functions** | **Predictions** | **Analysis**

Species (S_0)
Individuals (N_0)
Energy (E_0)
Area (A_0)

Individual Data

Taxon	Size	x	y
sp1	1.4	1.4	7.3
sp1	2.4	1.1	7.1
sp2	5.6	1.3	7.8
sp3	7.1	1.7	7.4
sp3	9.3	2.3	6.1

Aggregated Data

Taxon	#	Size	x	y
sp1	2	1.9	1	7
sp2	1	10.7	1	7
sp3	1	43.1	1	7
sp3	1	34.1	2	6

Ecosystem Structure Function $R(n, \varepsilon)$

Spatial Structure Function $\Pi(n)$

Probability distributions

$\Phi(n)$
$\Psi(\varepsilon)$
$\Theta(\varepsilon)$
…

Relationships

$S(A)$
$E(A)$
$\varepsilon(n)$
…

```
print
plot
q,r,p,d
logLik
residuals
mse
```

```
print
plot
residuals
mse
```

Example R Code

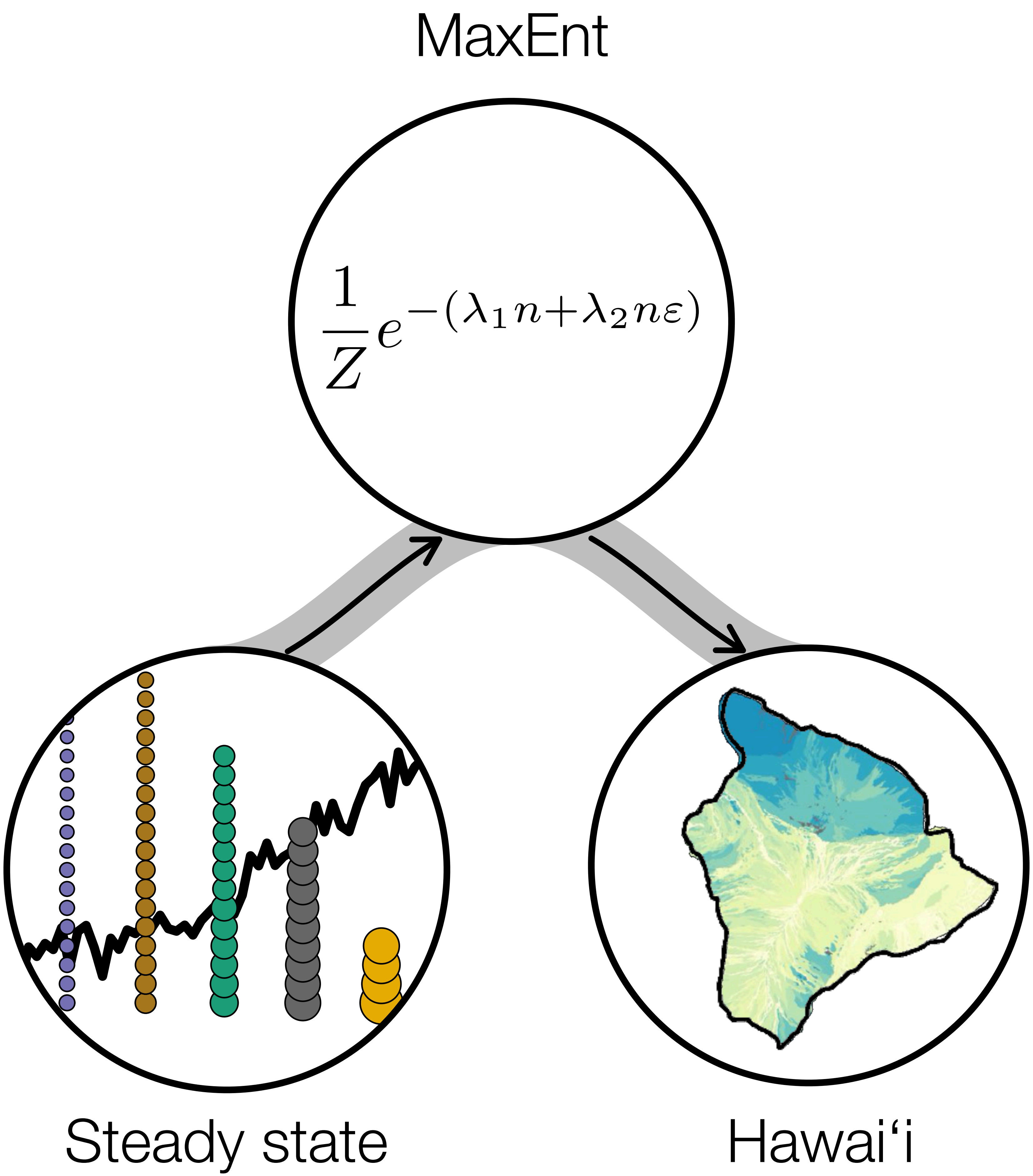
```
data(arth)
data(anbo)
```

```
esf<-meteESF(spp,
        abund,
        power)
Ssf<-meteSSF(spp,
        abund,
        coords)
```

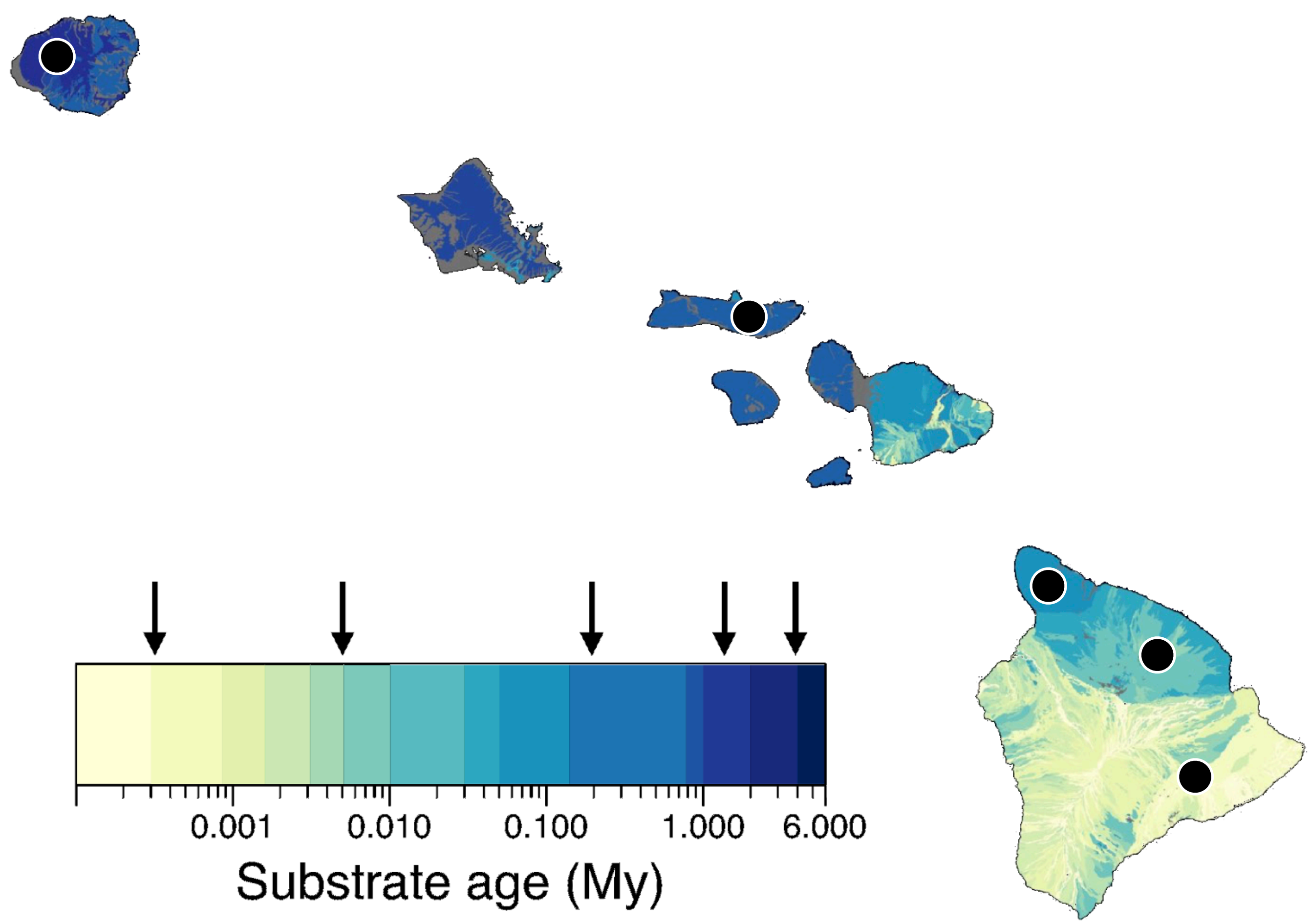
```
sad<-sad(esf)
ipd<-ipd(esf)
sipd<-sipd(esf)
sar1<-downscaleSAR(esf)
sar2<-meteSAR(spp,abund,x,y)
etc.
```

```
plot(sad)
logLik(sad)
mse(sad)
sad$r() #simulation
sad$q() #quantiles
etc.
```

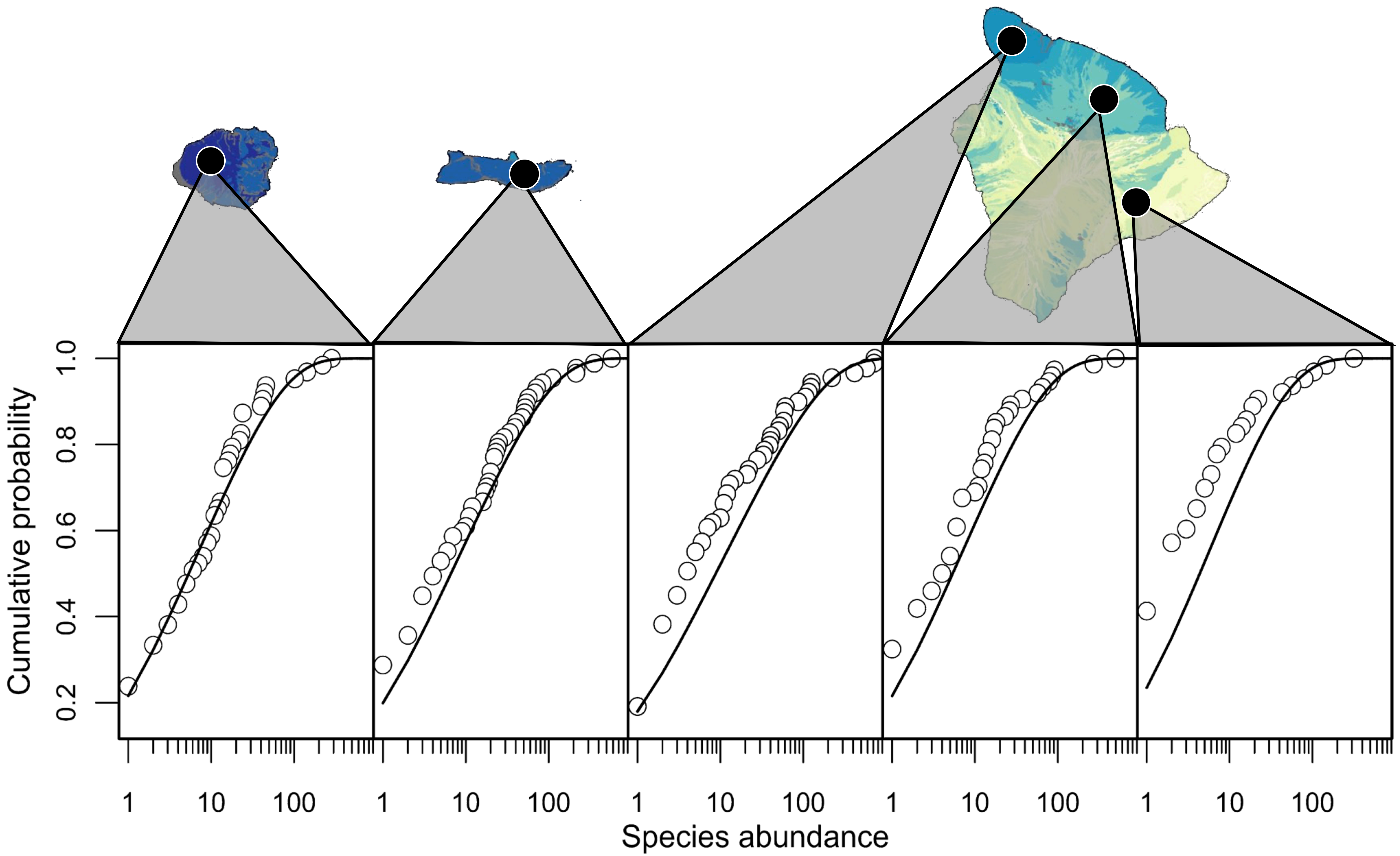
Rominger & Merow (2017) *MEE*


MaxEnt
$\frac{1}{Z}e^{-(\lambda_1 n+\lambda_2 n\varepsilon)}$
Steady state
Hawai‘i

MaxEnt and Hawaiian biodiversity

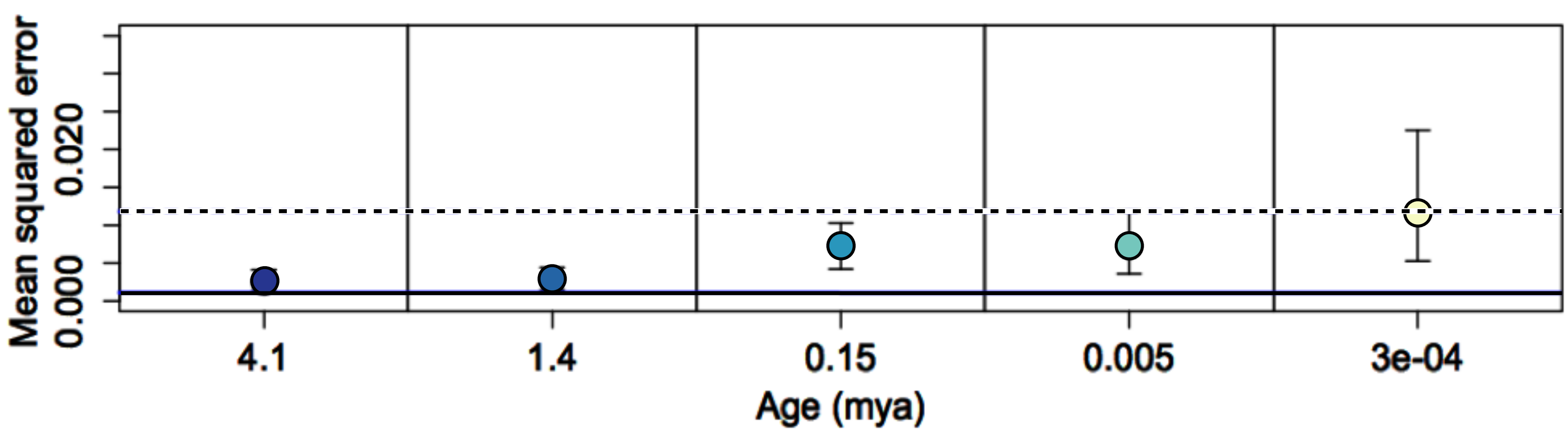


Data from Gruner (2007) Biol. J. Linn. Soc. 90: 551—570

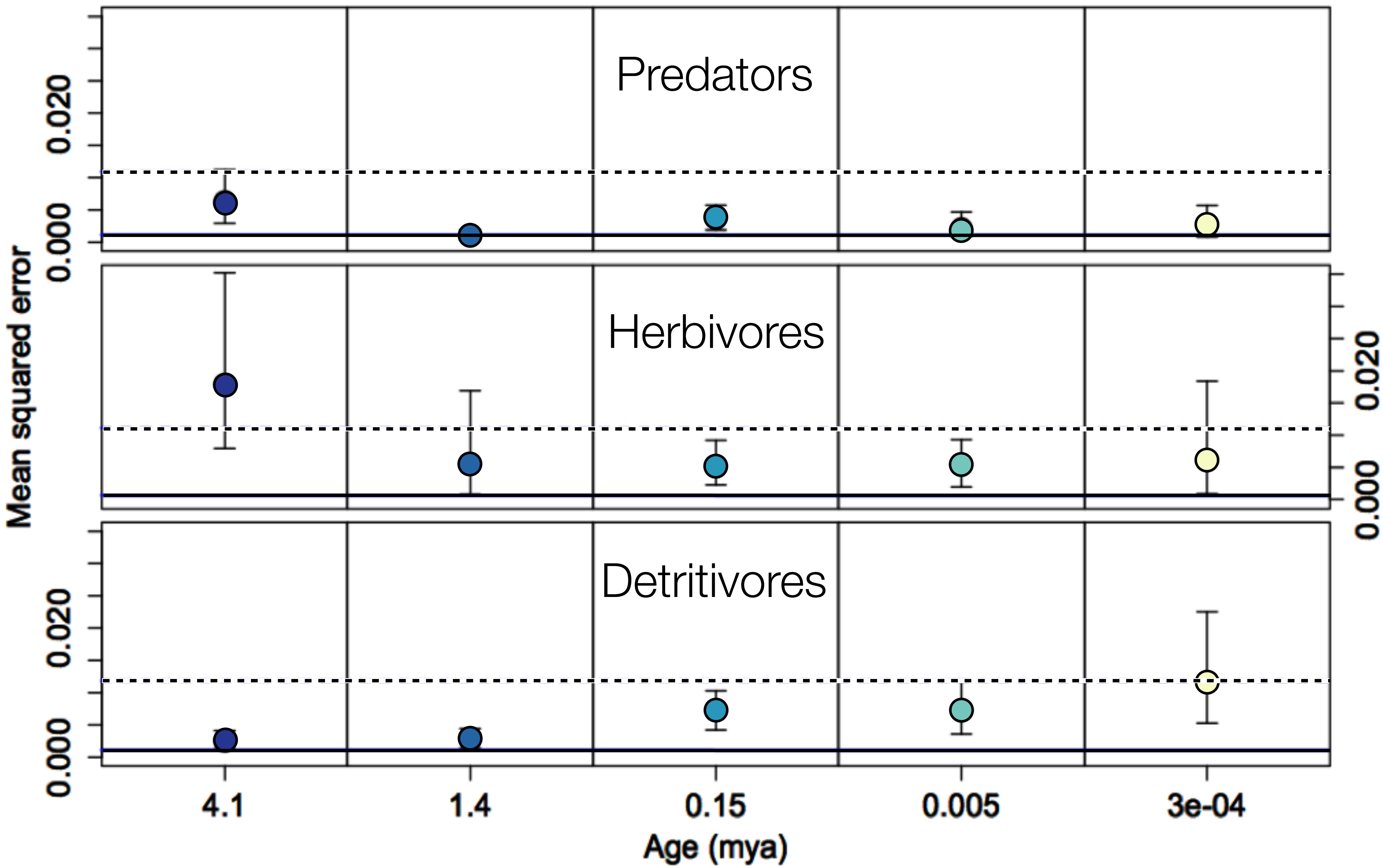


Cumulative probability
0.2
0.4
0.6
0.8
1.0
1
10
100
Species abundance

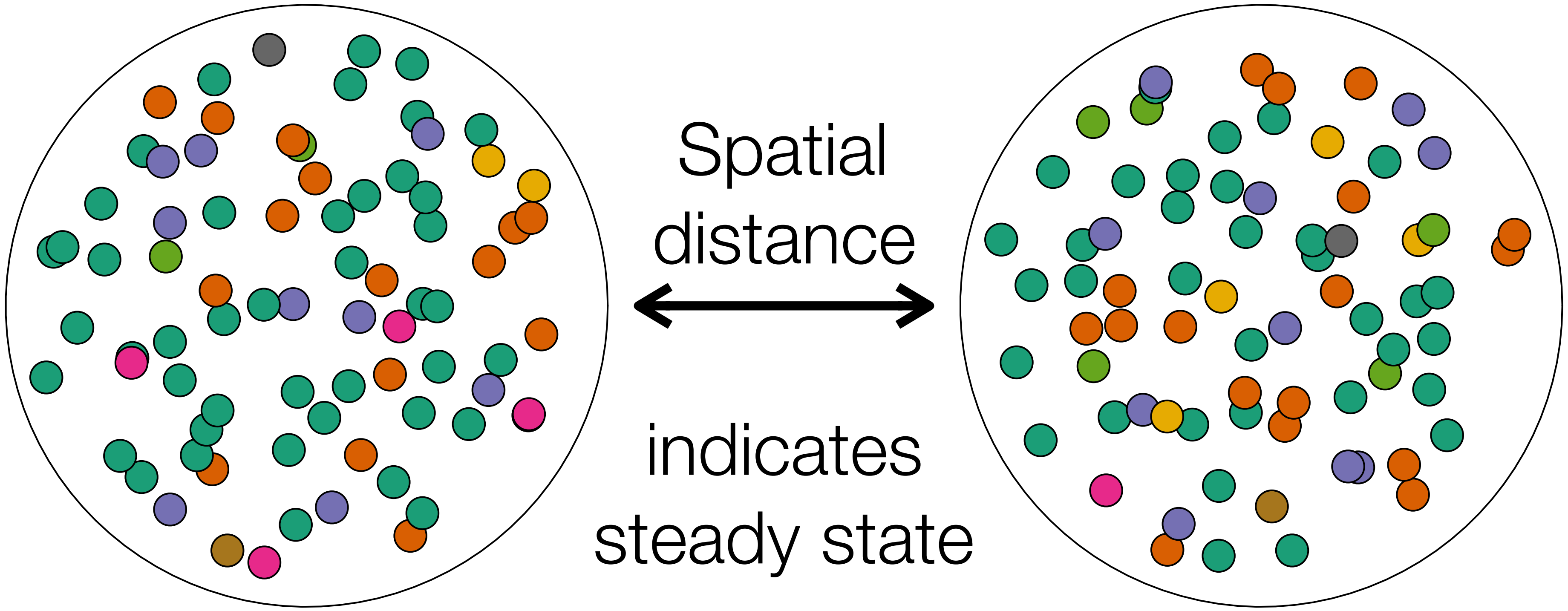
Species abundance $P(n)$



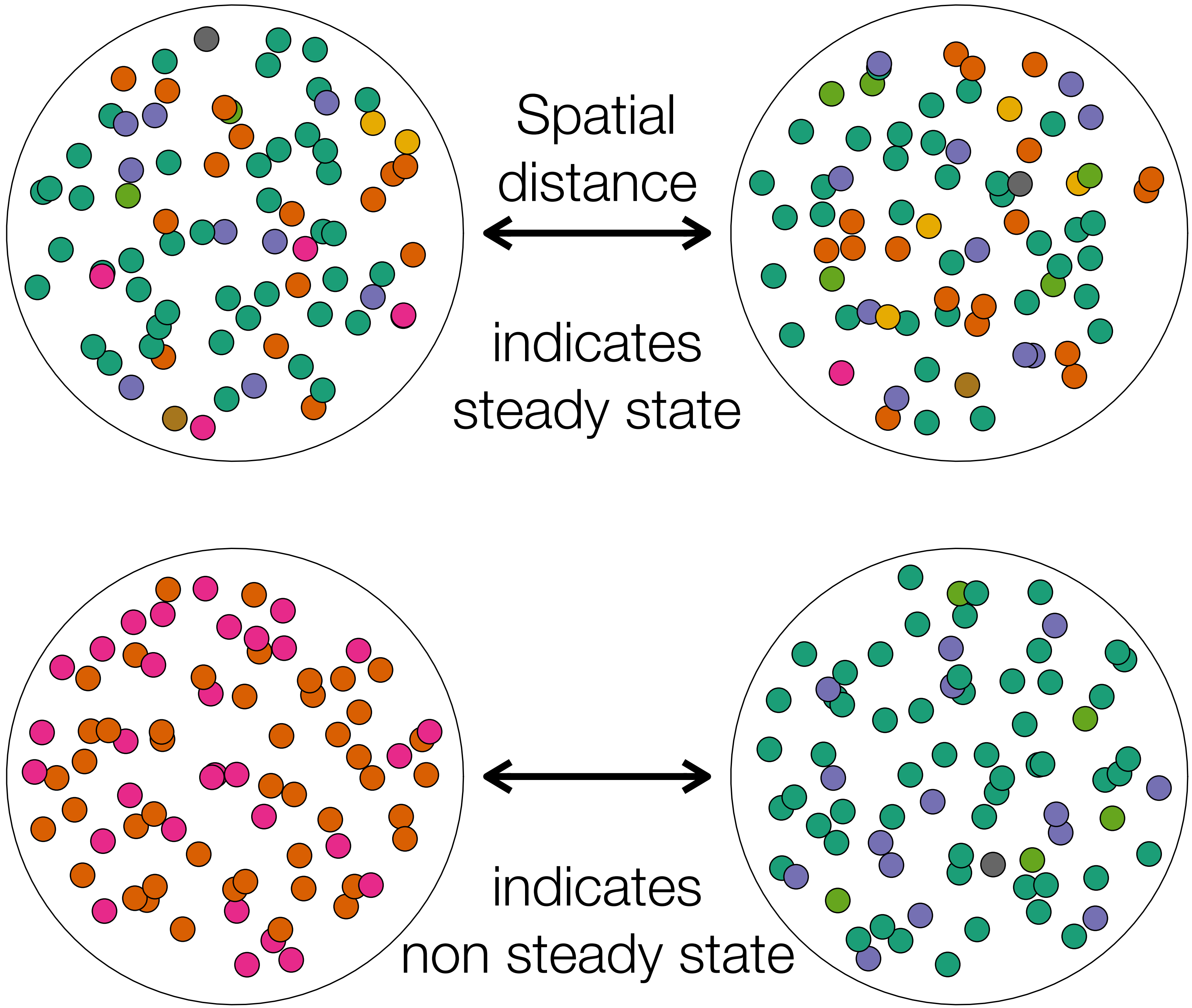
Species abundance *P(n)*



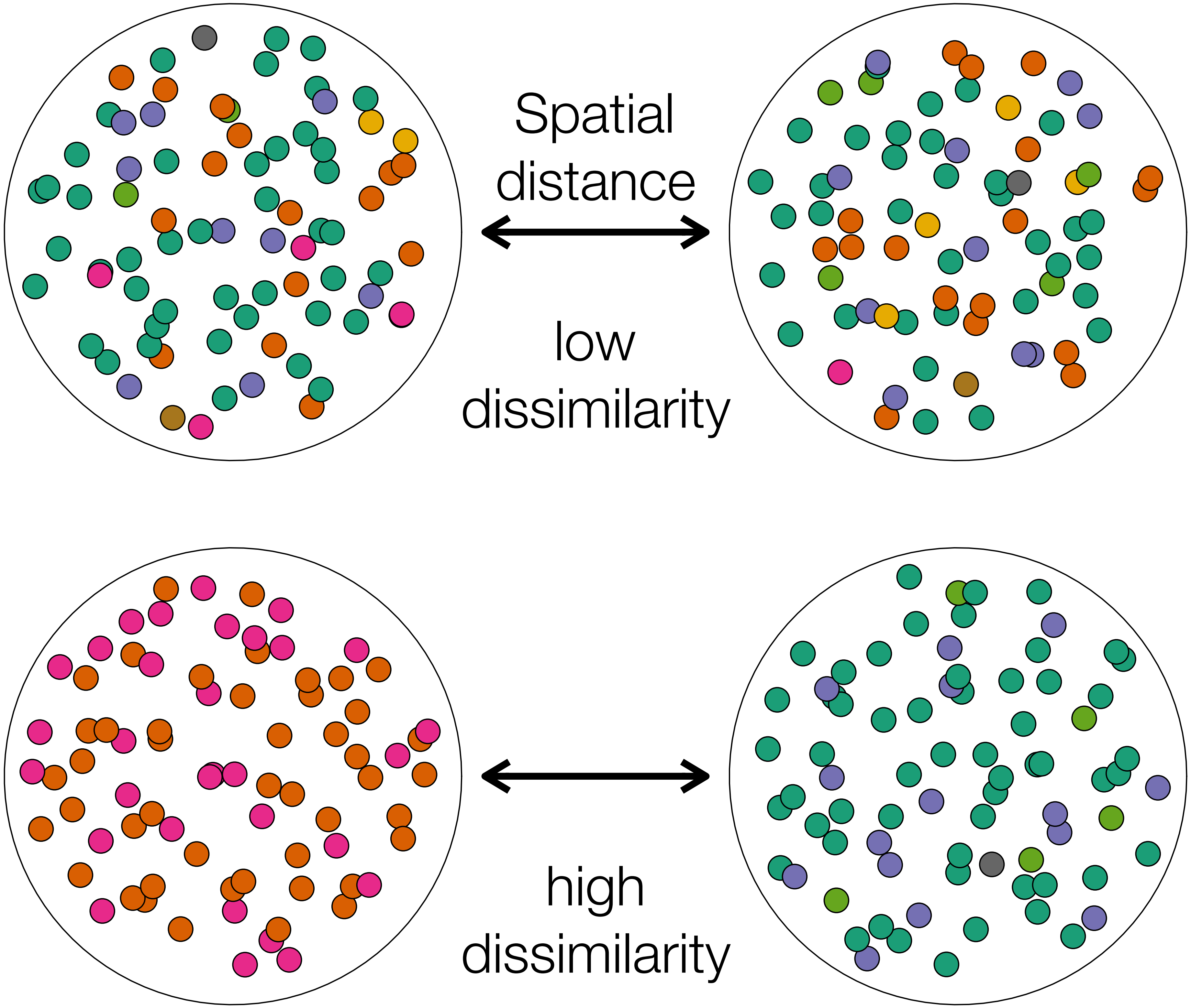
Space-for-time steady state?

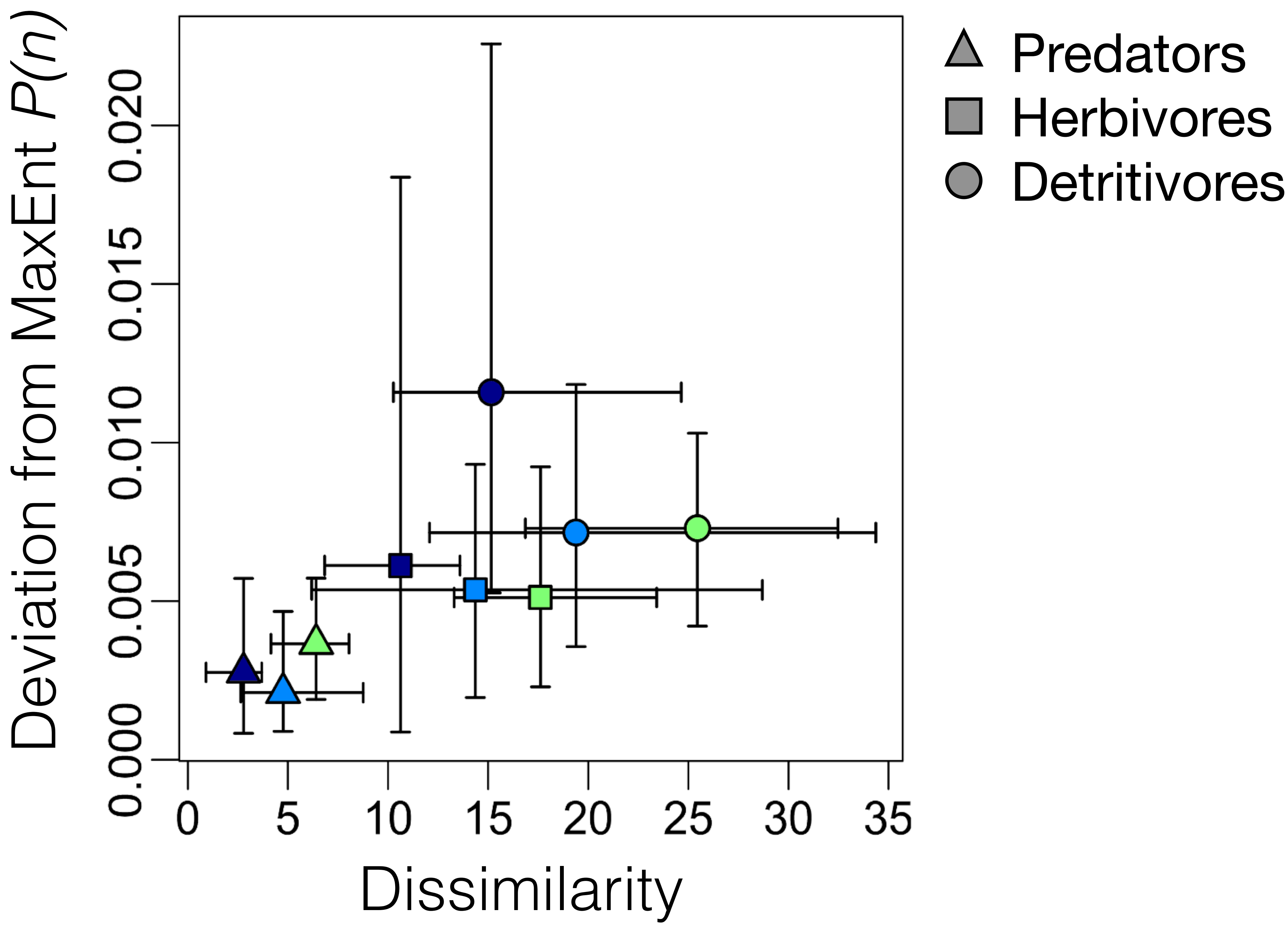


Space-for-time steady state?



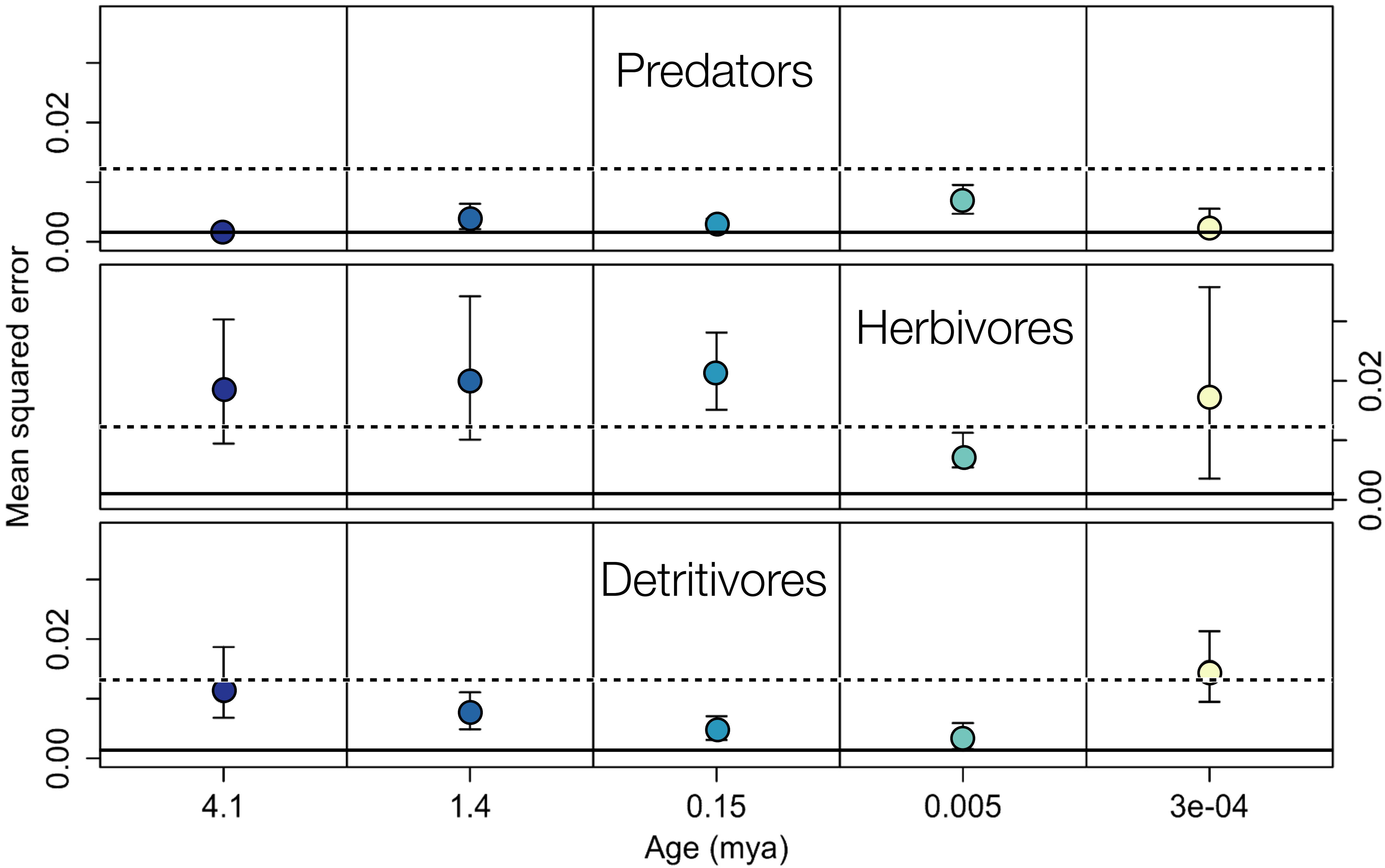
Space-for-time steady state?

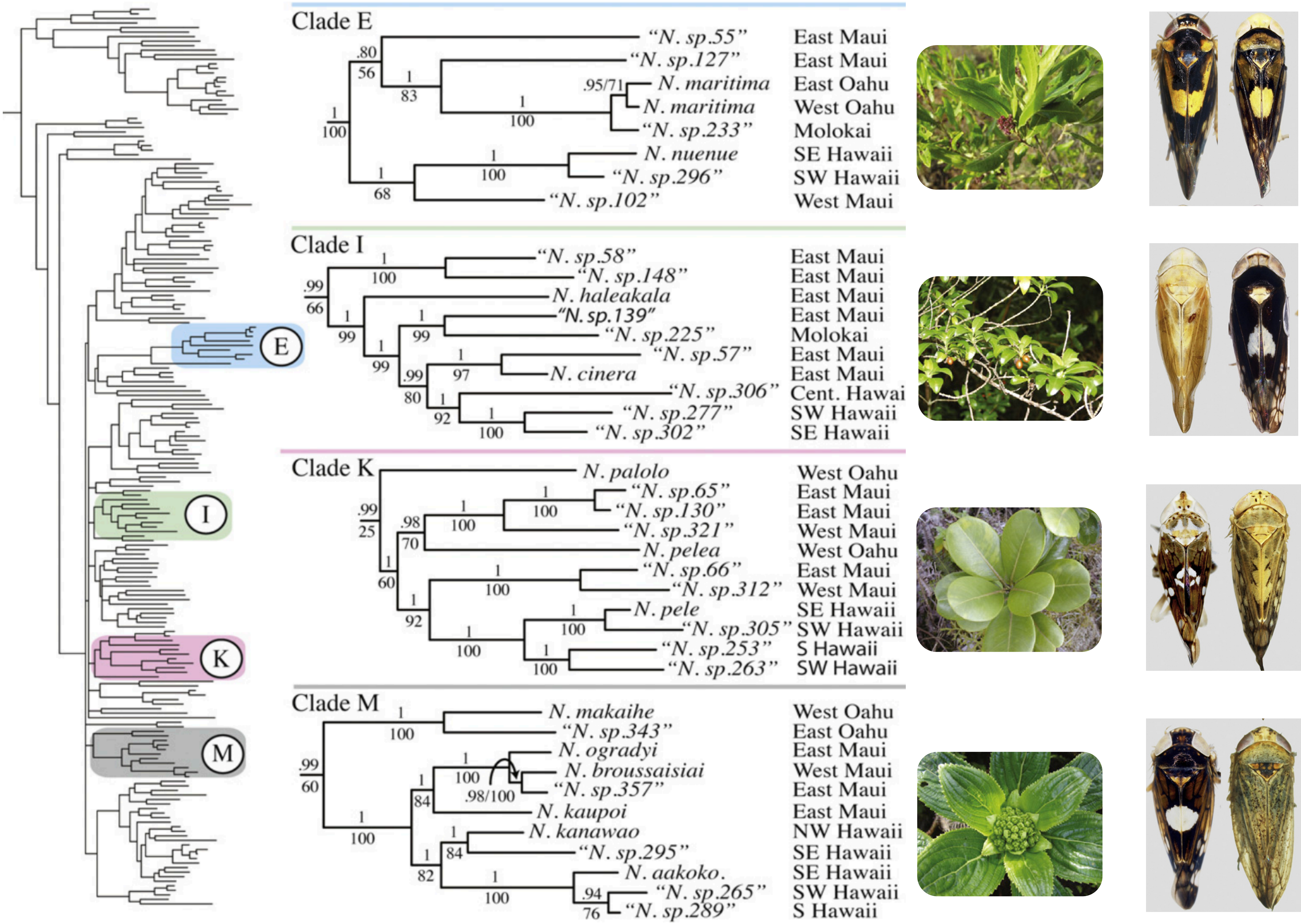




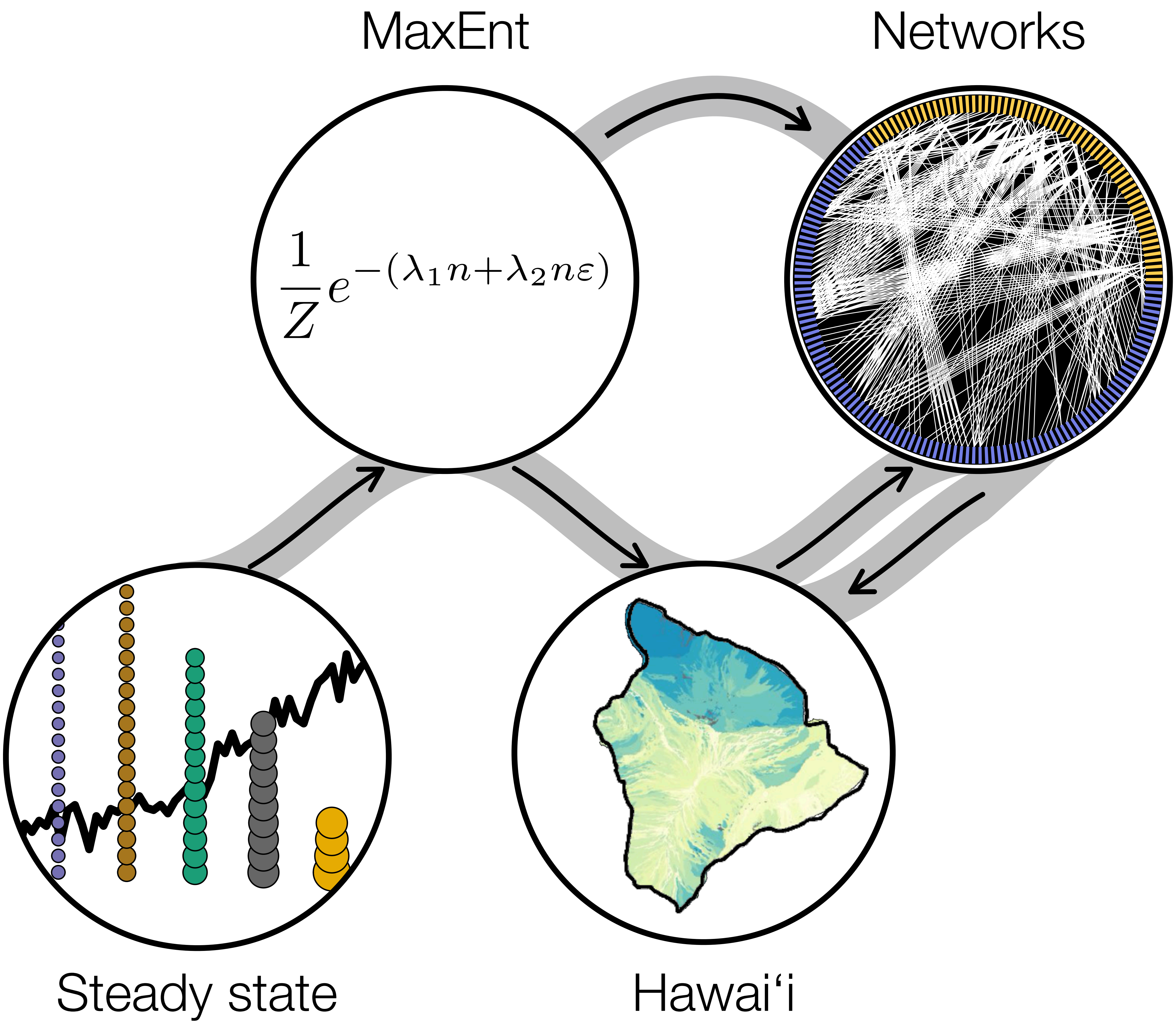
Deviation from MaxEnt P(n)
0.000
0.005
0.010
0.015
0.020
0
5
10
15
20
25
30
35
Dissimilarity
Predators
Herbivores
Detritivores

Metabolic rate $P(\varepsilon)$

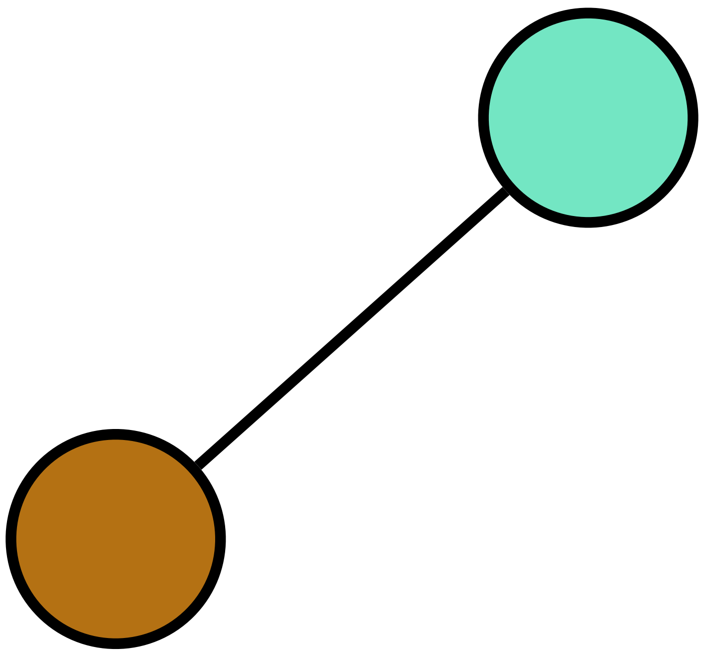




Bennett & O'Grady (2012) *Mol. Phylo. Evol.*

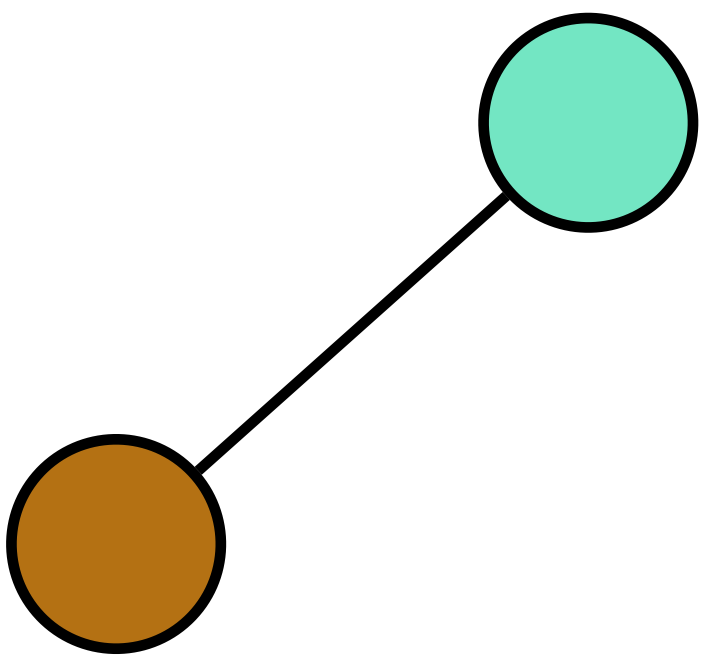

MaxEnt
Networks
$\frac{1}{Z}e^{-(\lambda_1 n+\lambda_2 n\varepsilon)}$
Steady state
Hawai‘i

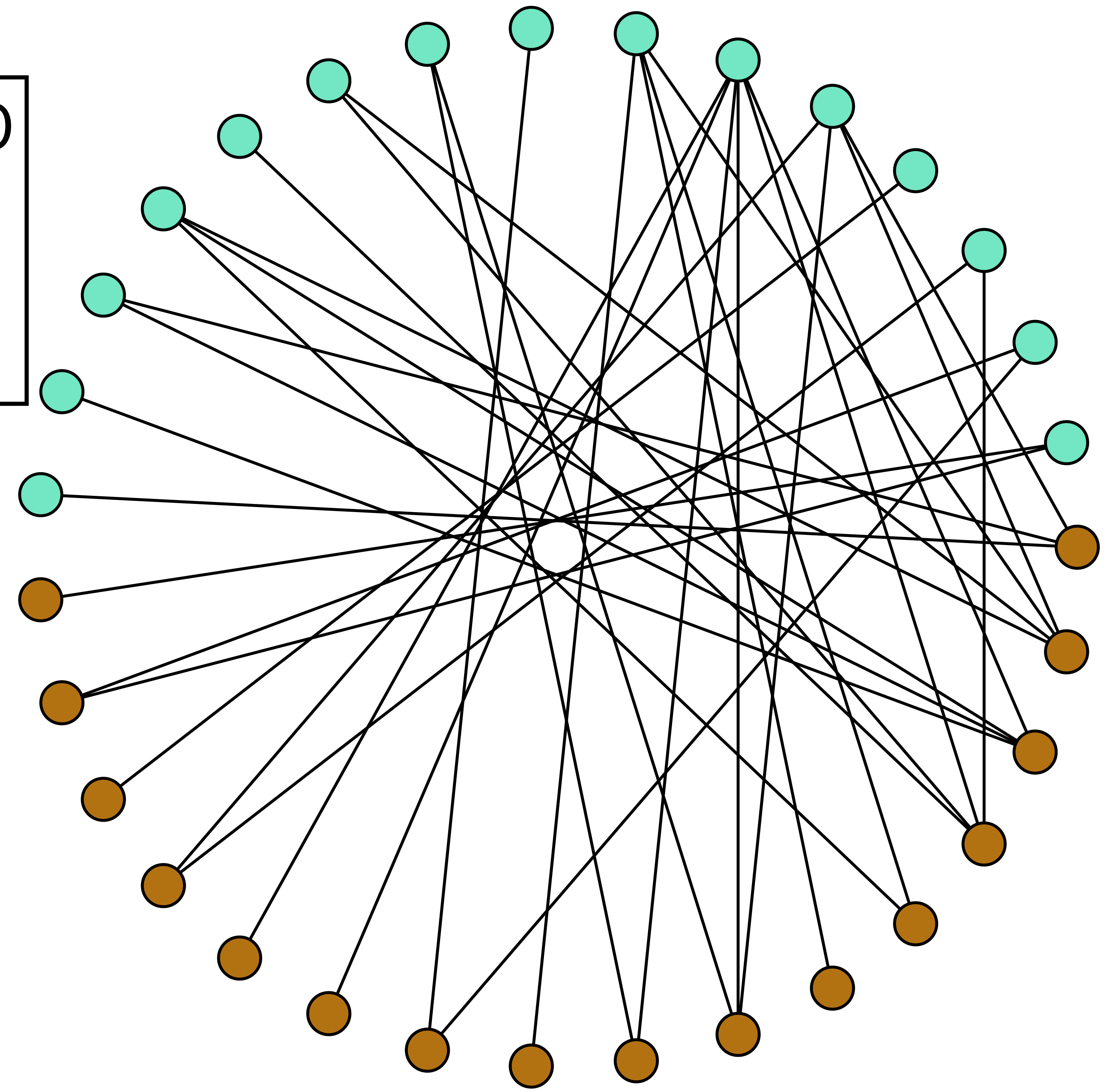
HDIM6643; laupLSAG_10
Nesophrosyne pelea
09:25; 2012-10-01
Beating Melicope sp.



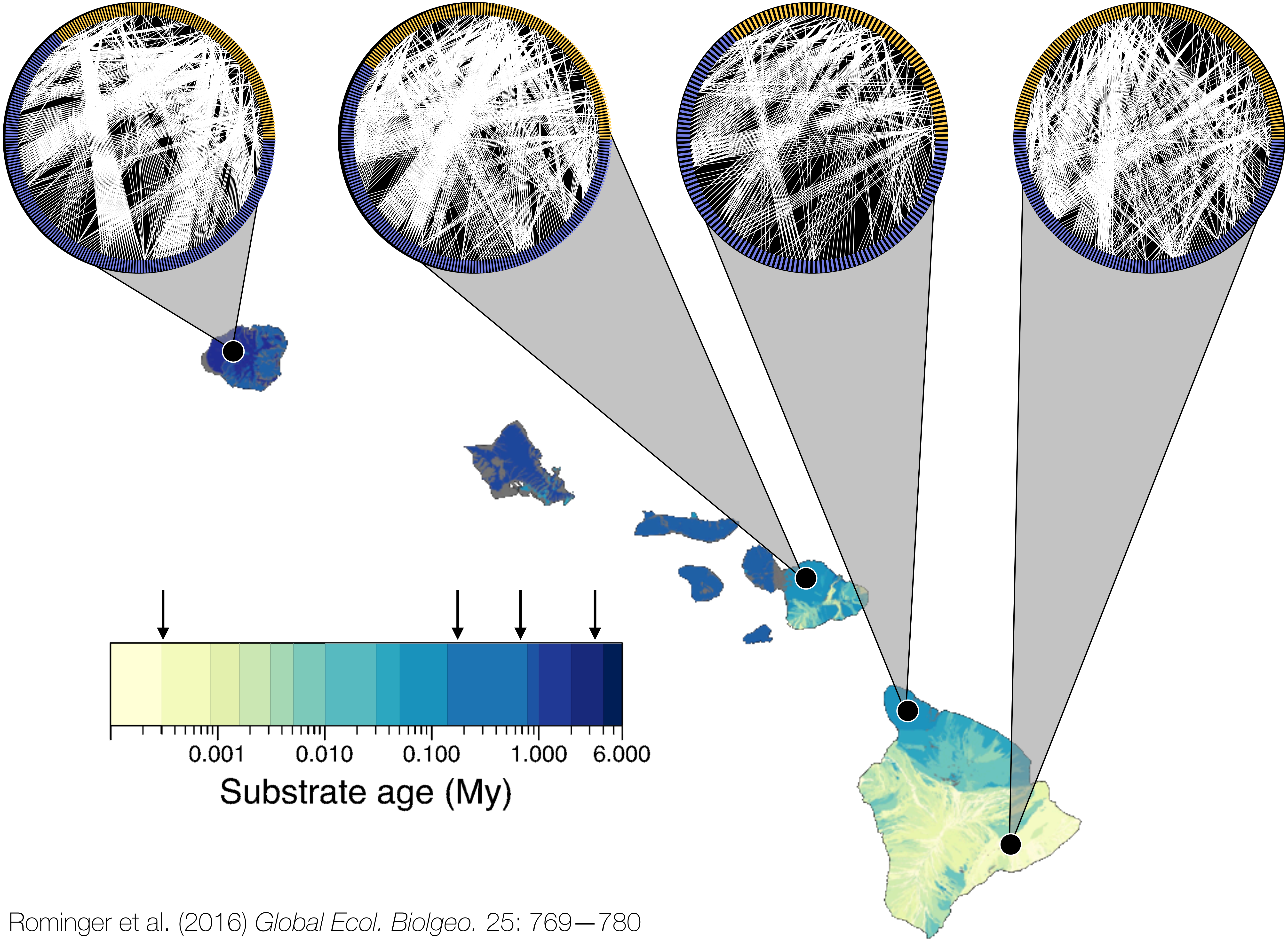
Rominger et al. (2016) *Global Ecol. Biolgeo.* 25: 769—780

HDIM6643; laupLSAG_10
Nesophrosyne pelea
09:25; 2012-10-01
Beating Melicope sp.





Rominger et al. (2016) *Global Ecol. Biolgeo.* 25: 769—780



Rominger et al. (2016) *Global Ecol. Biolgeo.* 25: 769—780

MaxEnt for bipartite networks

$$\frac{\text{Total number of edges}}{\text{Total number of vertices}}$$

Rominger et al. (2016) *Global Ecol. Biolgeo.* 25: 769—780

MaxEnt for bipartite networks

$$\frac{L_0}{S_0} \quad \frac{\text{Total number of edges}}{\text{Total number of vertices}}$$

Rominger et al. (2016) *Global Ecol. Biolgeo.* 25: 769—780

MaxEnt for bipartite networks

$$\sum_{l \in \mathcal{L}} P(l) = 1$$

$$\sum_{l \in \mathcal{L}} lP(l) = \frac{L_0}{S_0}$$

Rominger et al. (2016) *Global Ecol. Biolgeo.* 25: 769—780

MaxEnt for bipartite networks

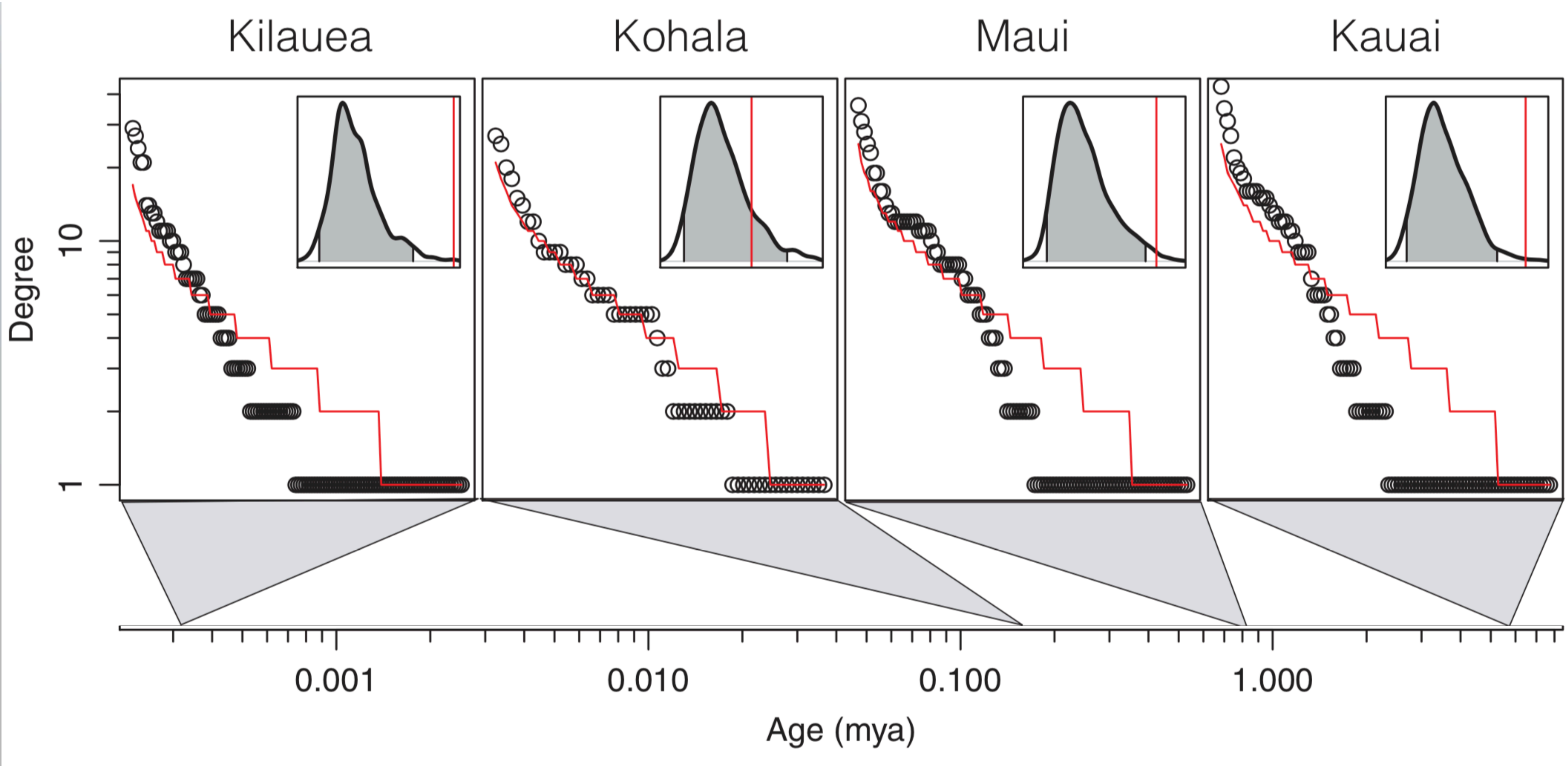
$$\sum_{l \in \mathcal{L}} P(l) = 1$$

$$\sum_{l \in \mathcal{L}} l P(l) = \frac{L_0}{S_0}$$

$$P(l) = \frac{1}{Z} e^{-\lambda l}$$

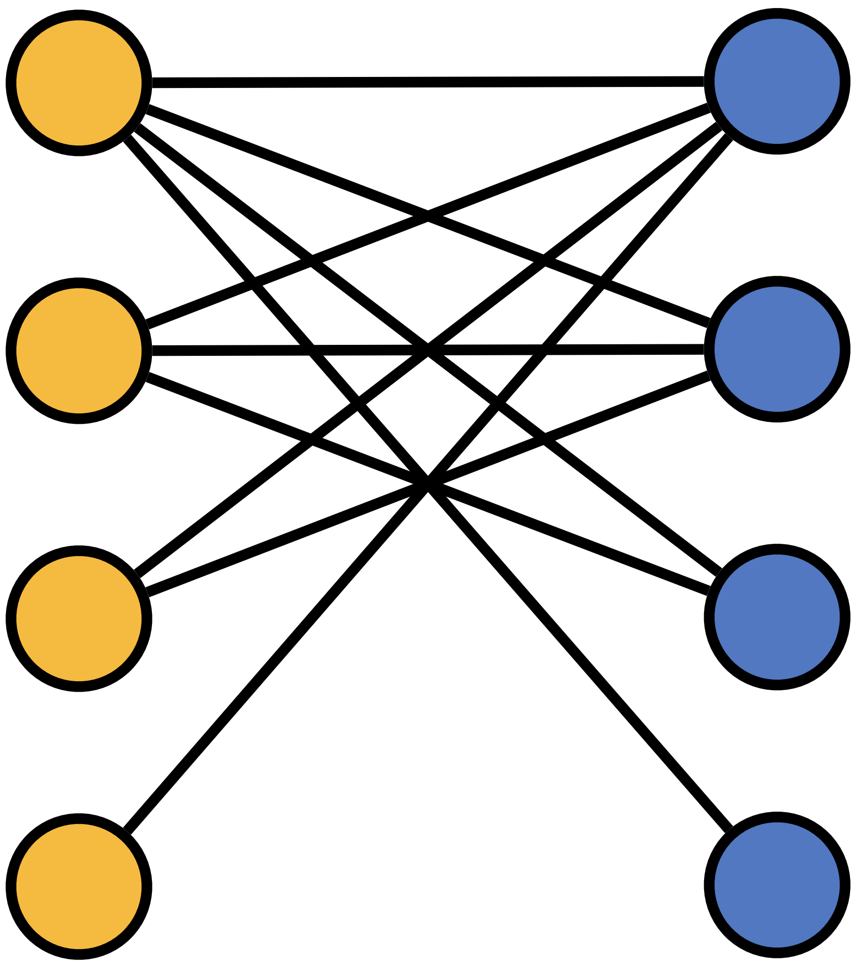
Rominger et al. (2016) *Global Ecol. Biolgeo.* 25: 769—780

Herbivore network

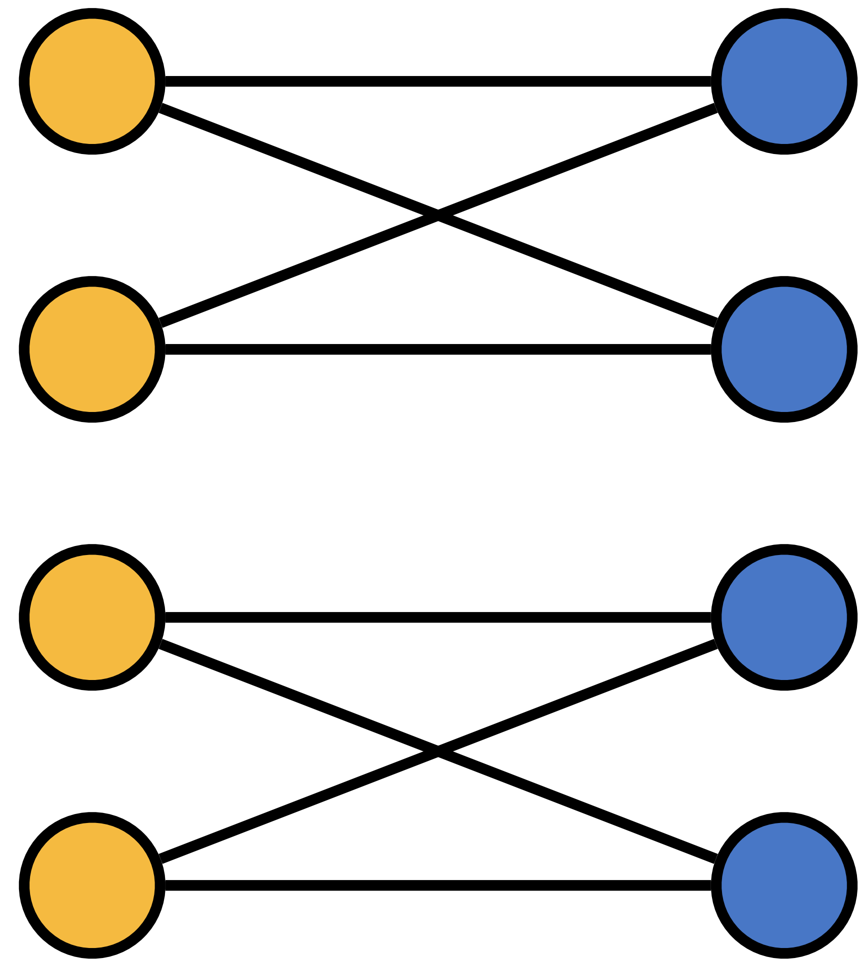


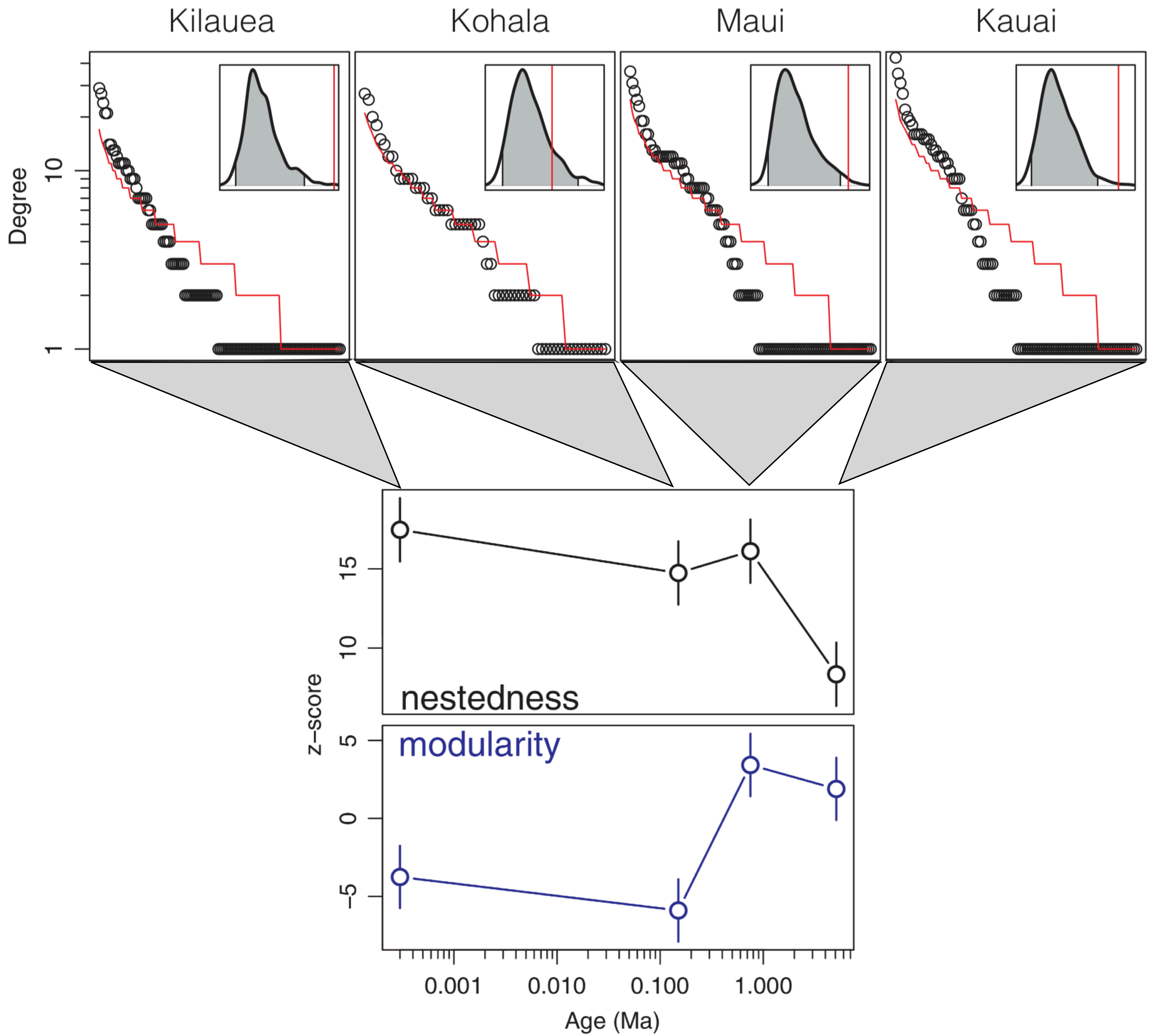
Rominger et al. (2016) *Global Ecol. Biolgeo.* 25: 769—780

Nestedness

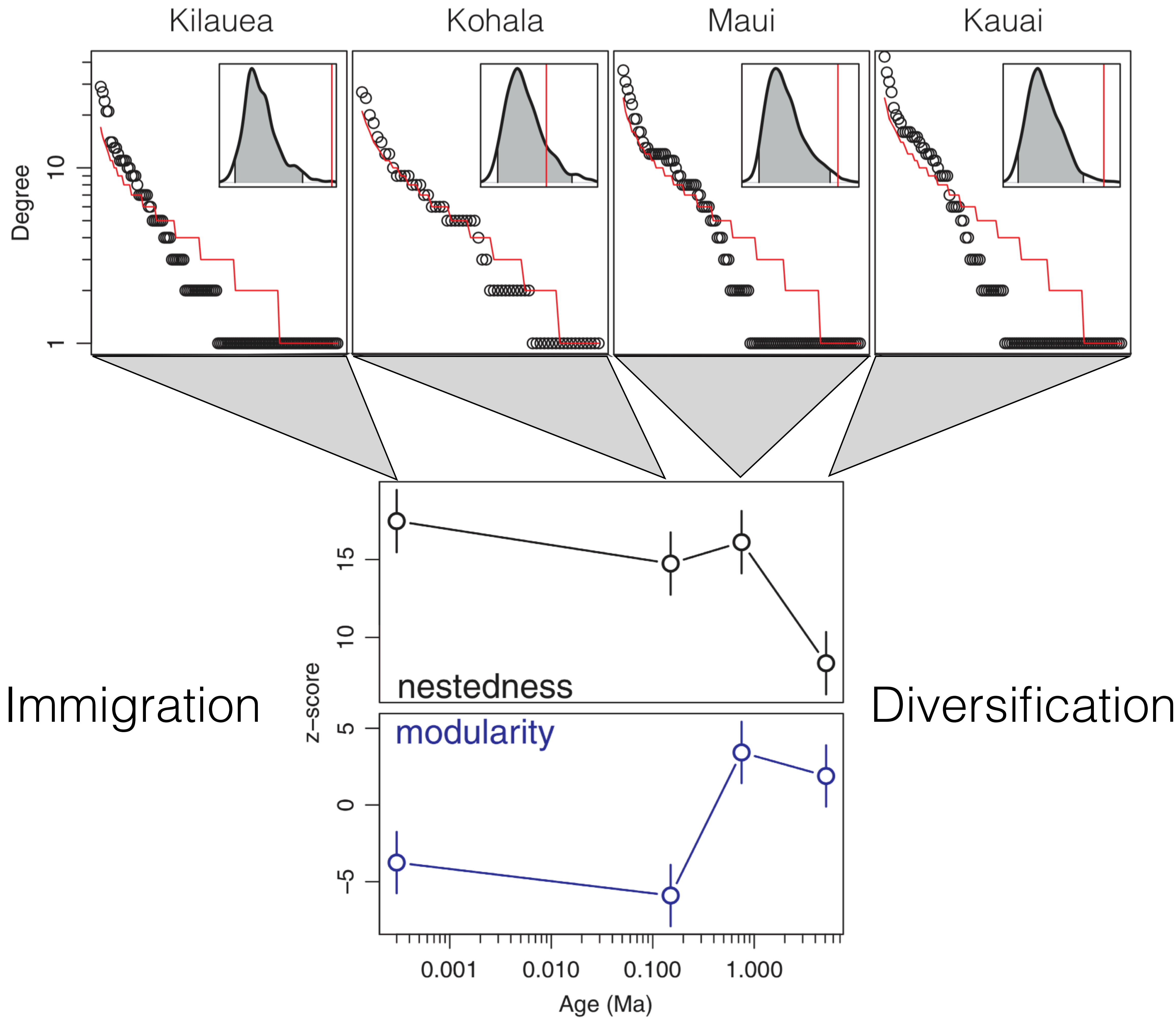


Modularity





Kilauea
Kohala
Maui
Kauai
Degree
10
1
z−score
15
10
nestedness
5
0
−5
modularity
0.001
0.010
0.100
1.000
Age (Ma)


Kilauea
Kohala
Maui
Kauai
Degree
10
1
Immigration
Diversification
nestedness
modularity
z−score
15
10
5
0
−5
0.001
0.010
0.100
1.000
Age (Ma)

New opportunities

Student project ideas:

- Exploring different constraints

$$\sum_{n\in\mathcal{N}} \int_{\mathcal{E}} R(n,\varepsilon)d\varepsilon = 1$$

$$\sum_{n\in\mathcal{N}} \int_{\mathcal{E}} nR(n,\varepsilon)d\varepsilon = \frac{N_0}{S_0}$$

$$\sum_{n\in\mathcal{N}} \int_{\mathcal{E}} n\varepsilon R(n,\varepsilon)d\varepsilon = \frac{E_0}{S_0}$$

$$\sum_{n\in\mathcal{N}} \int_{\mathcal{E}} R(n,\varepsilon)d\varepsilon = 1$$

$$\sum_{n\in\mathcal{N}} \int_{\mathcal{E}} nR(n,\varepsilon)d\varepsilon = \frac{N_0}{S_0}$$

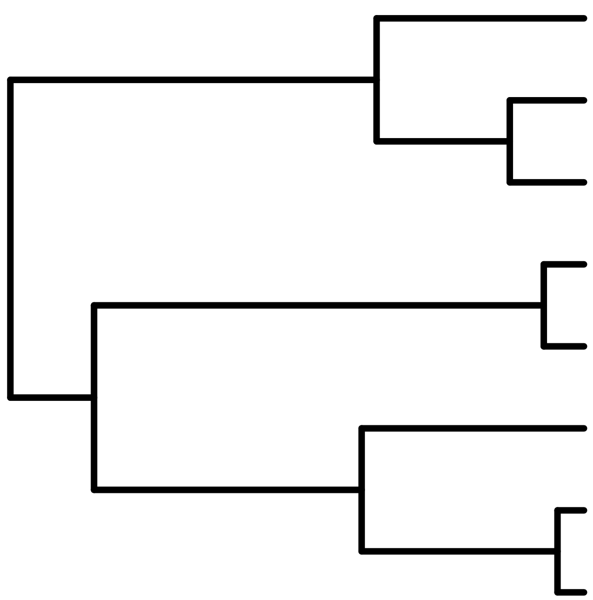
$$\sum_{n\in\mathcal{N}} \int_{\mathcal{E}} \varepsilon R(n,\varepsilon)d\varepsilon = \frac{1}{S_0}\sum_{i=1}^{S_0} \bar{\varepsilon}_i$$

$$\sum_{n\in\mathcal{N}} \int_{\mathcal{E}} n\varepsilon R(n,\varepsilon)d\varepsilon = \frac{E_0}{S_0}$$

New opportunities

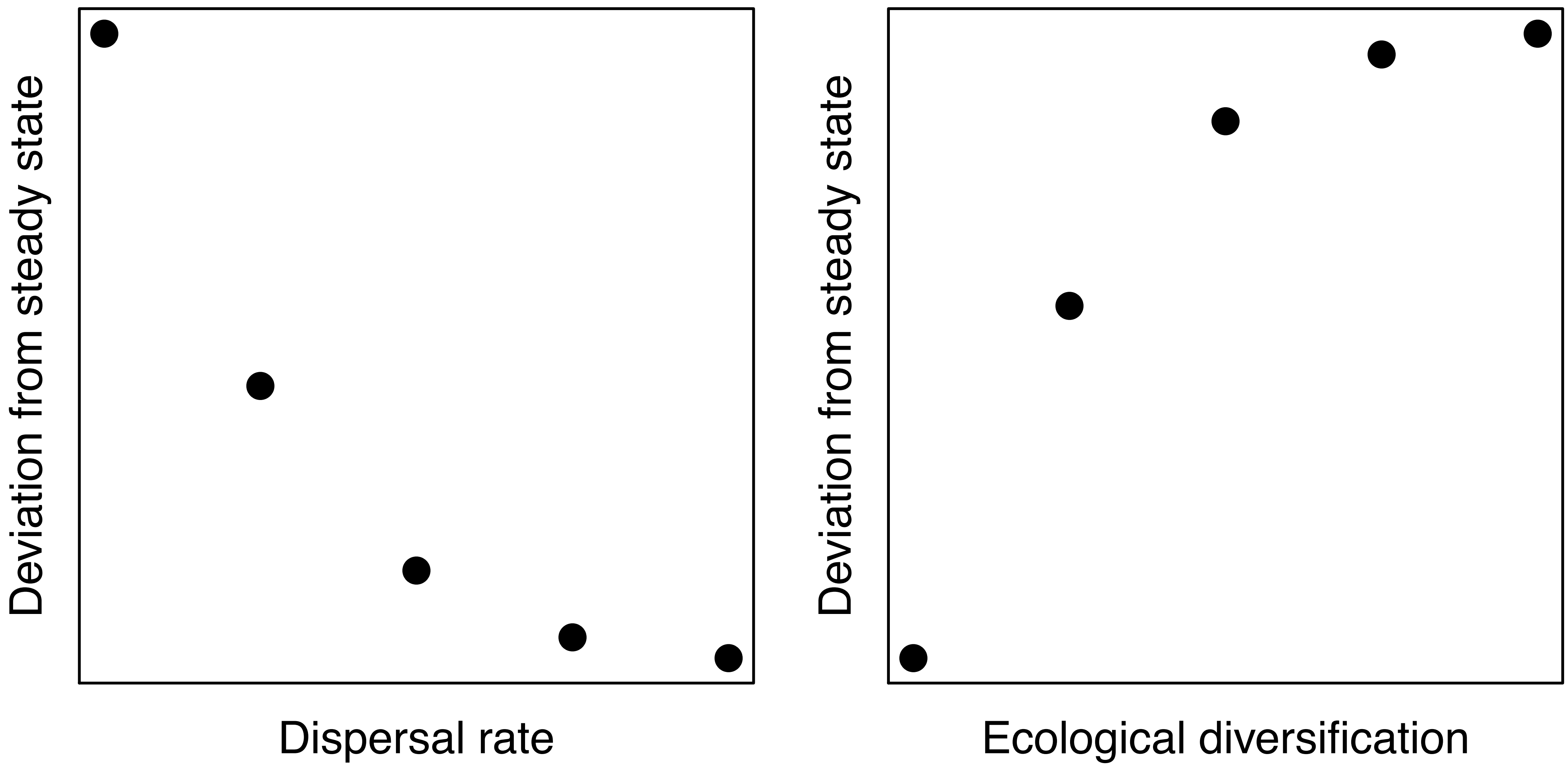
Student project ideas:

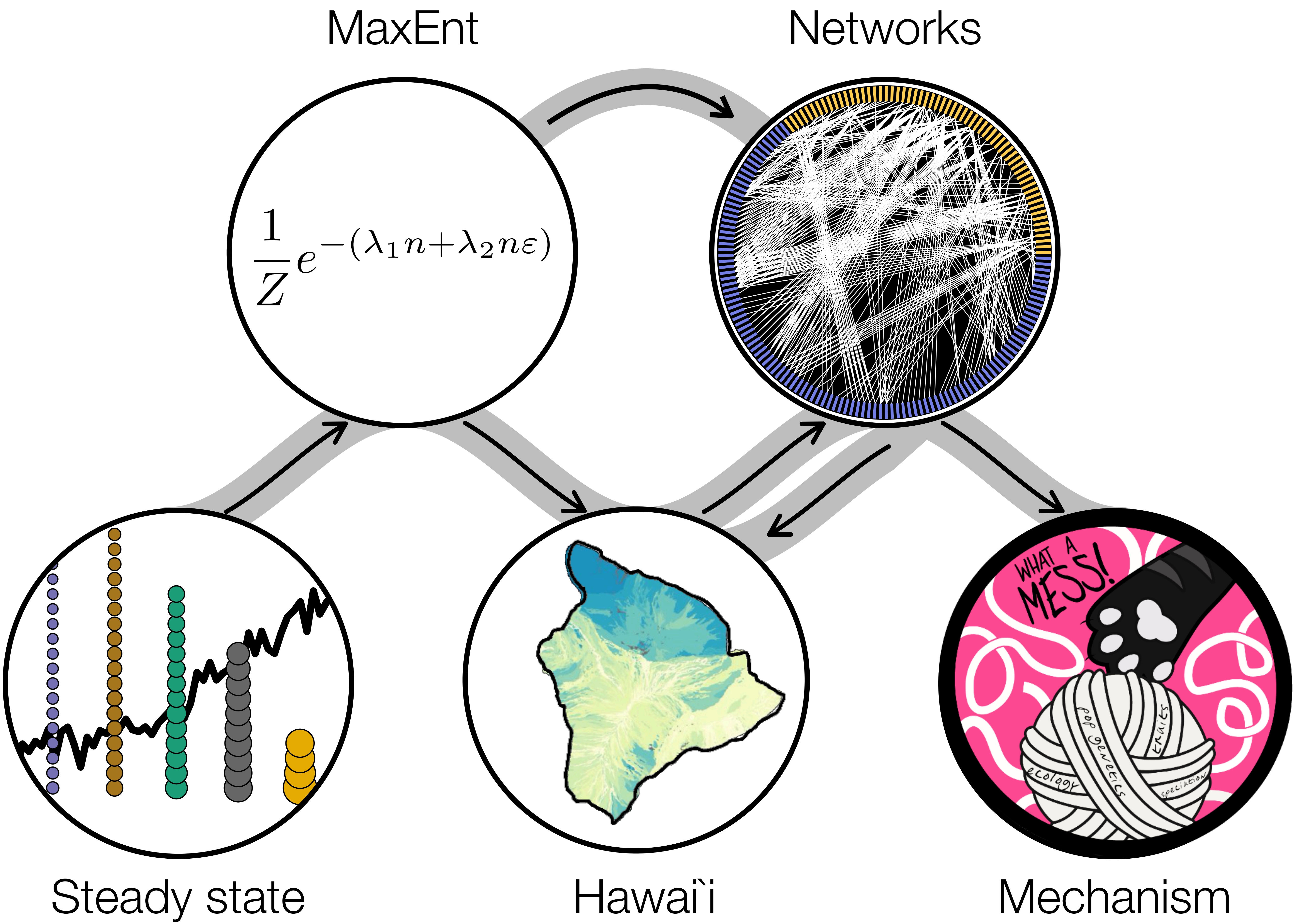
- Exploring different constraints
- Deriving MaxEnt predictions for other interesting biodiversity metrics



CTCTCAGTGC
GTCACAGTGC
GTCACAGTTC
GACACAGTTC
GACAGAGTTC

We're left with hypotheses




MaxEnt
Networks
$\frac{1}{Z}e^{-(\lambda_1 n+\lambda_2 n\varepsilon)}$
WHAT A MESS!
pop genetics
traits
ecology
speciation
Steady state
Hawaì'i
Mechanism

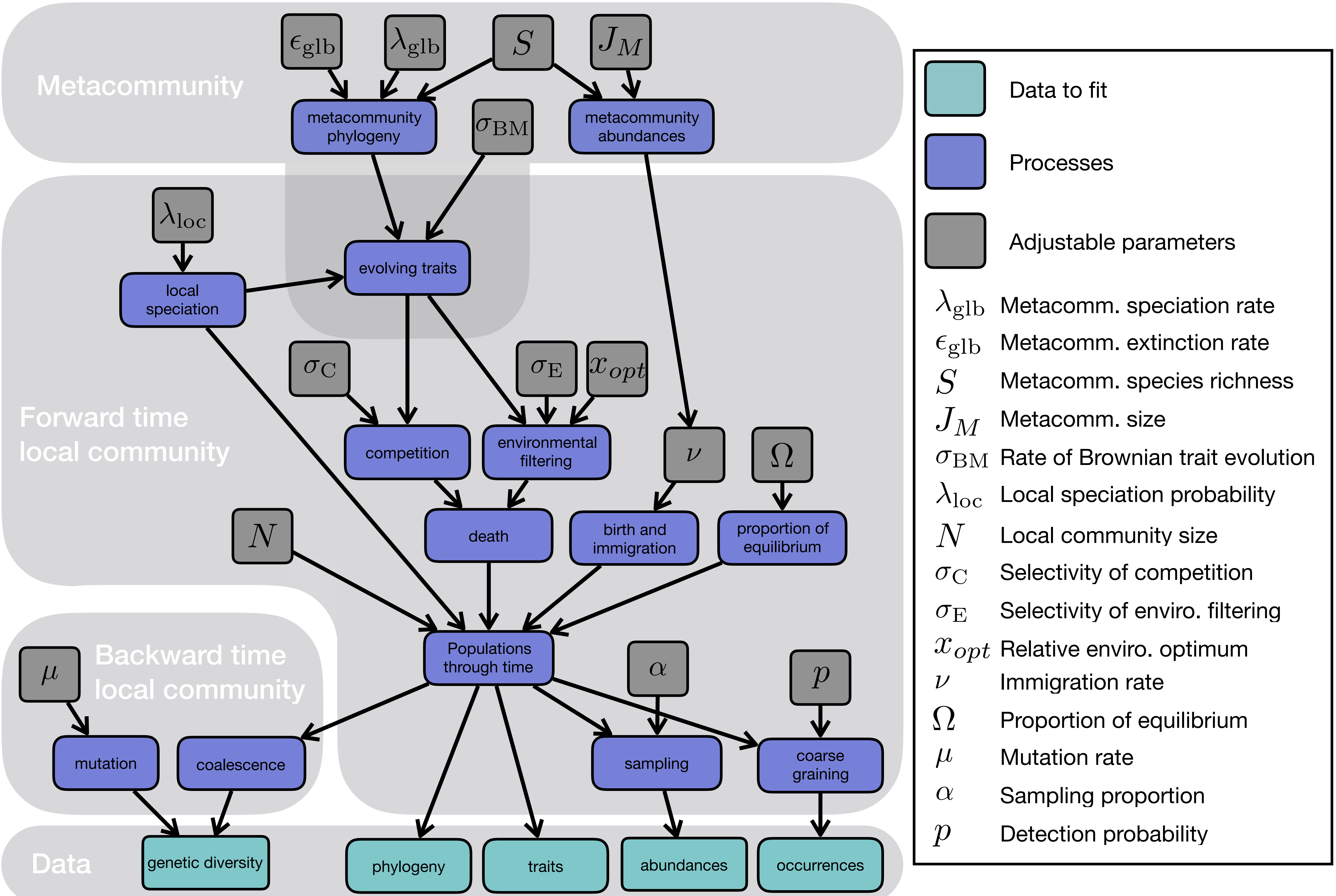
Mechanistic models to test competing hypotheses

Mechanistic models to test competing hypotheses

the Massive Eco-evolutionary Synthesis Simulation

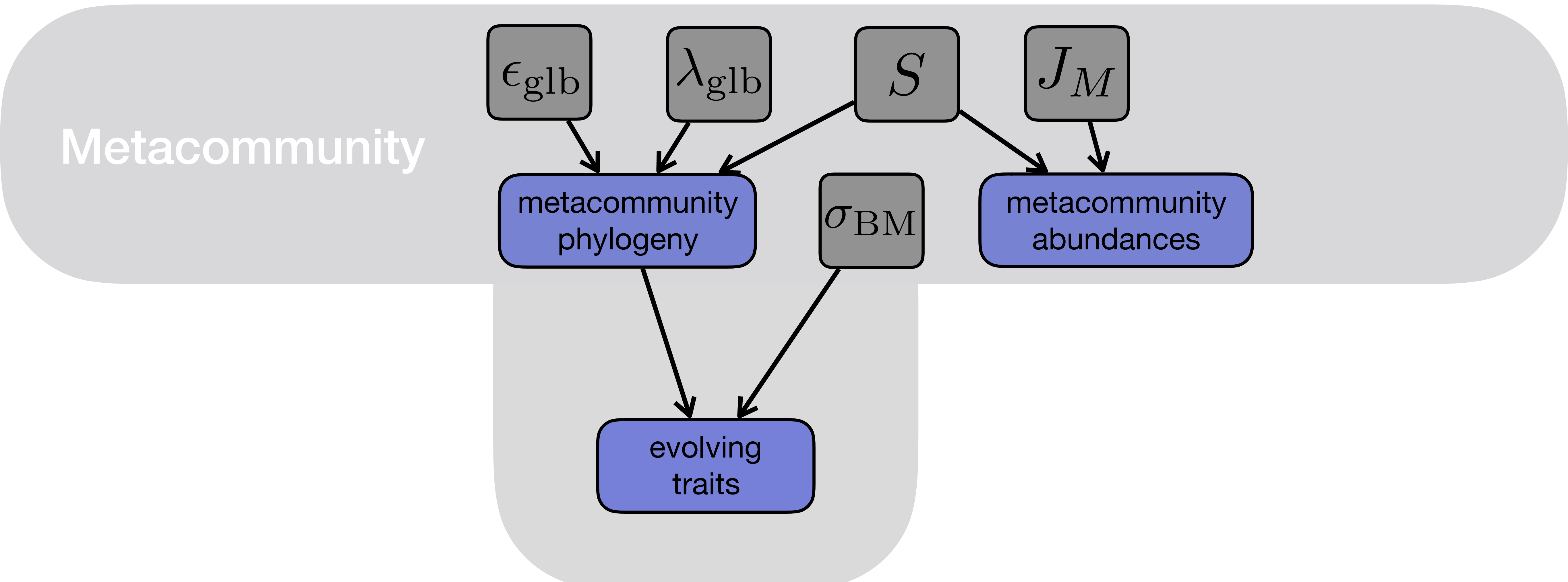
Mechanistic models to test competing hypotheses

the MESS

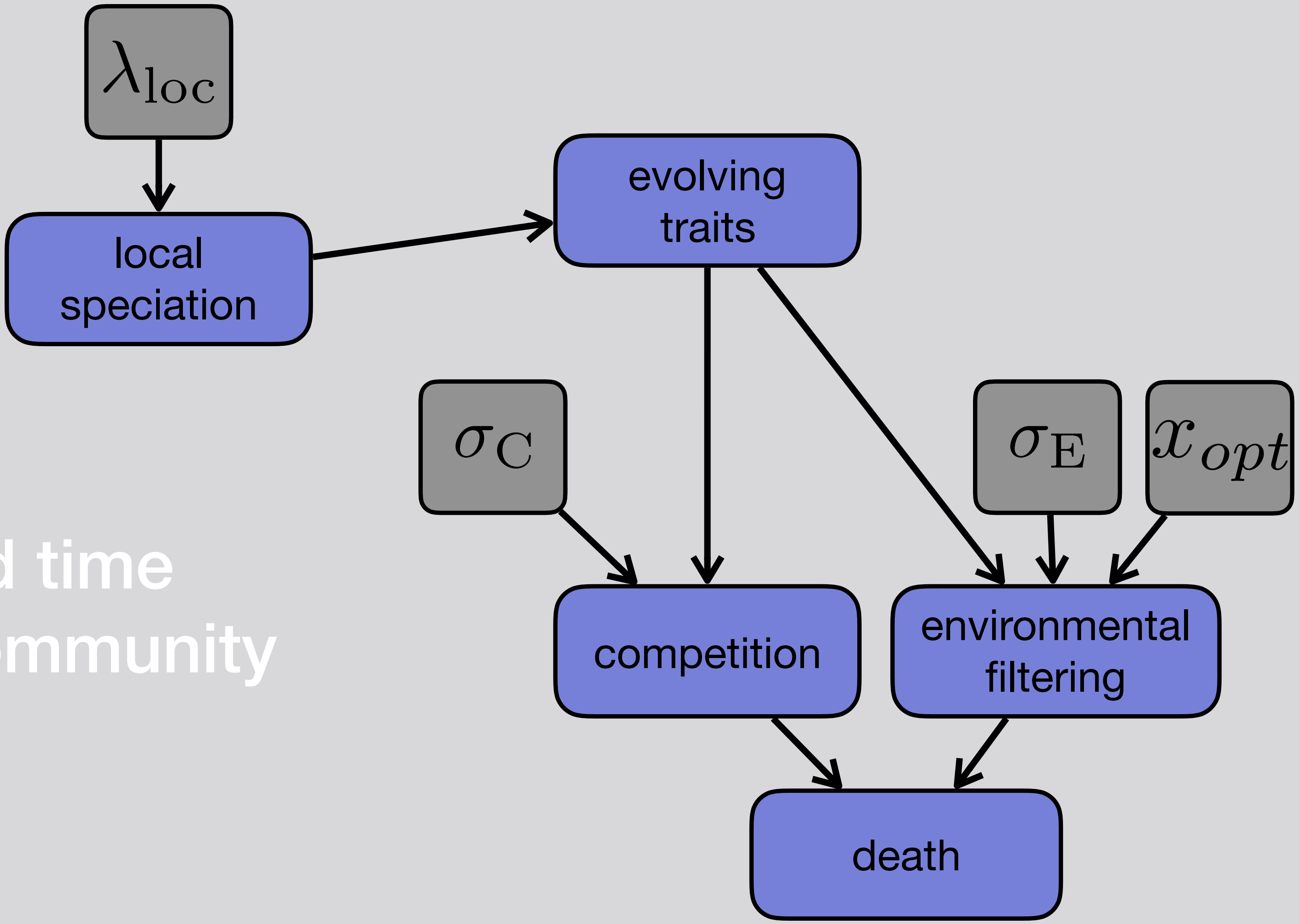


Metacommunity
ϵ_{glb}
λ_{glb}
S
J_M
metacommunity phylogeny
σ_{BM}
metacommunity abundances
Forward time local community
λ_{loc}
local speciation
evolving traits
σ_{C}
σ_{E}
x_{opt}
competition
environmental filtering
ν
Ω
N
death
birth and immigration
proportion of equilibrium
Backward time local community
μ
Populations through time
α
p
mutation
coalescence
sampling
coarse graining
Data
genetic diversity
phylogeny
traits
abundances
occurrences
Data to fit
Processes
Adjustable parameters
λ_{glb} Metacomm. speciation rate
ϵ_{glb} Metacomm. extinction rate
S Metacomm. species richness
J_M Metacomm. size
σ_{BM} Rate of Brownian trait evolution
λ_{loc} Local speciation probability
N Local community size
σ_{C} Selectivity of competition
σ_{E} Selectivity of enviro. filtering
x_{opt} Relative enviro. optimum
ν Immigration rate
Ω Proportion of equilibrium
μ Mutation rate
α Sampling proportion
p Detection probability

Evolutionary processes

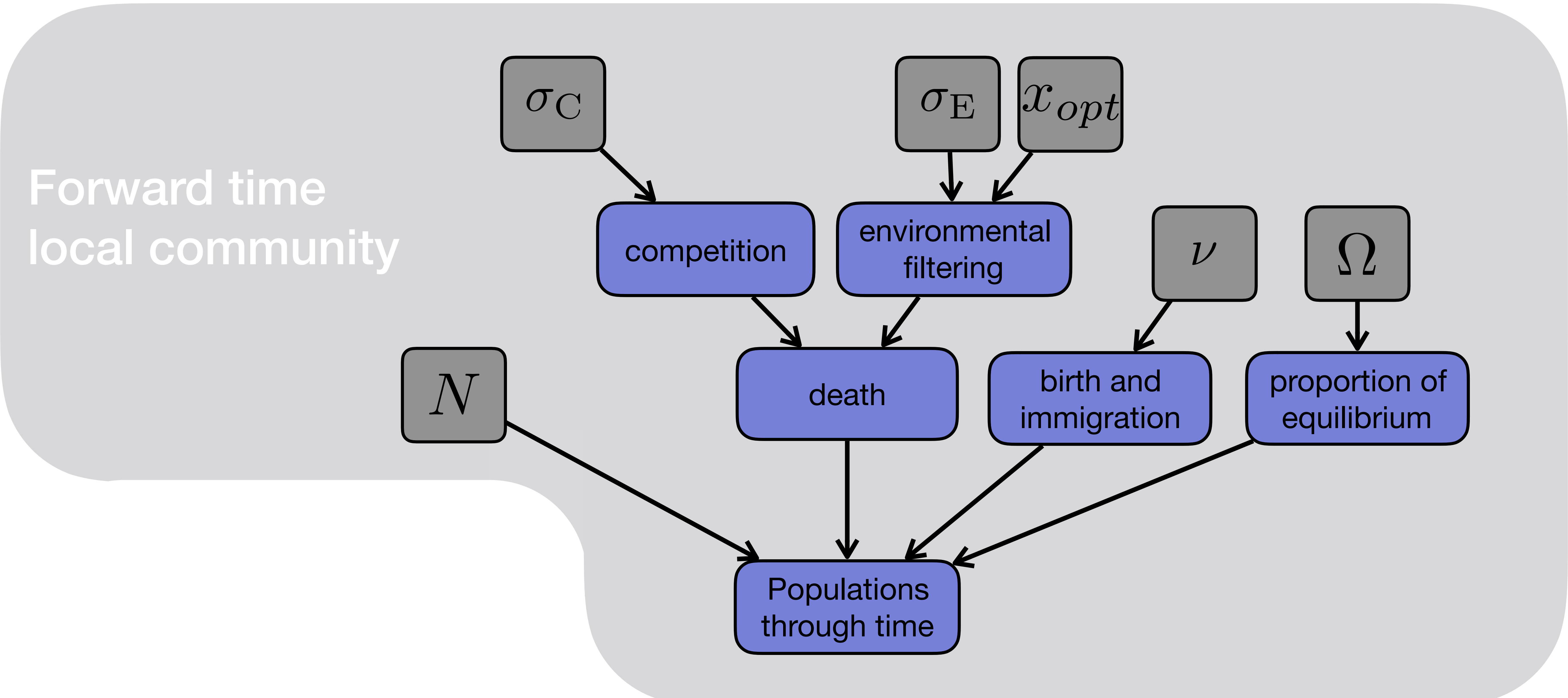


Ecological processes

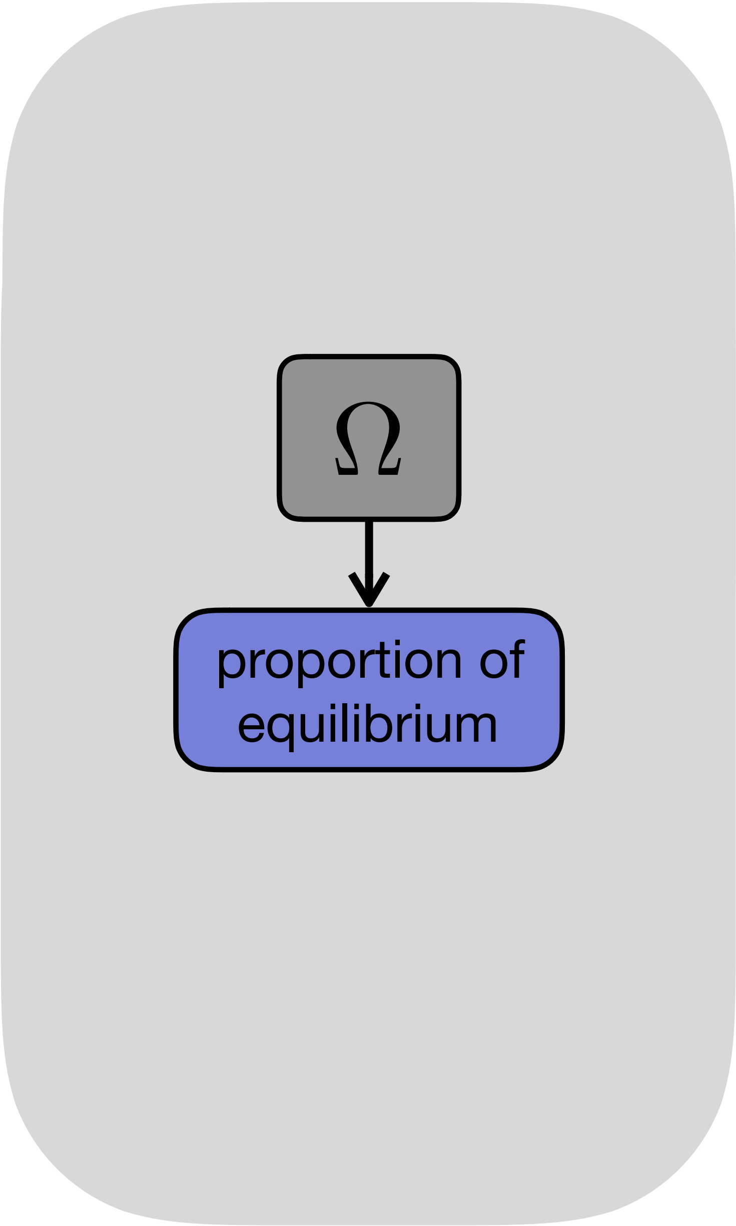


Forward time
local community

Ecological processes

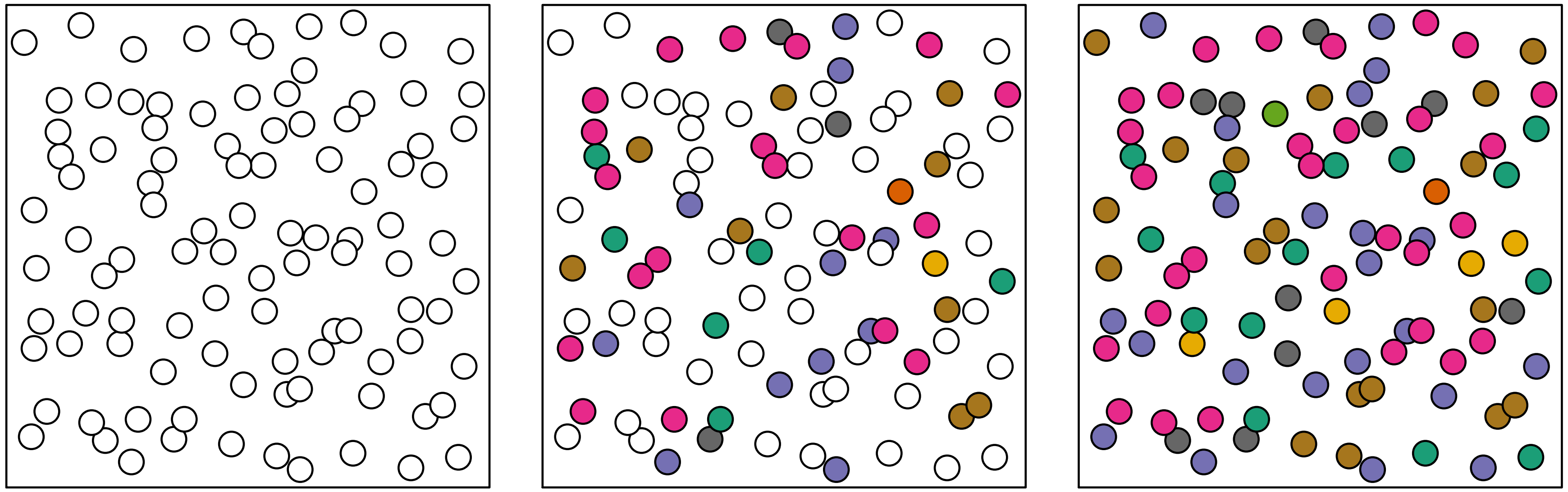


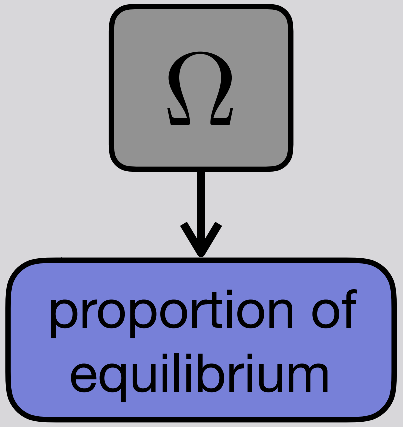
Ecological processes



Ecological processes

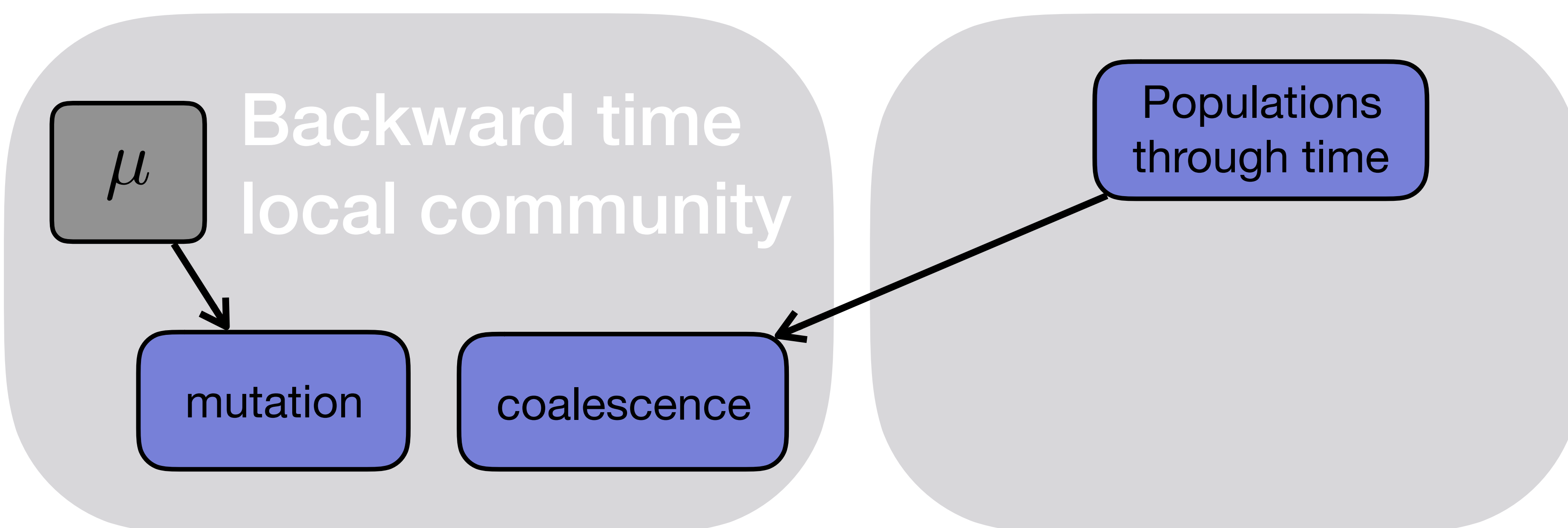
Equilibrium = escape from initial conditions

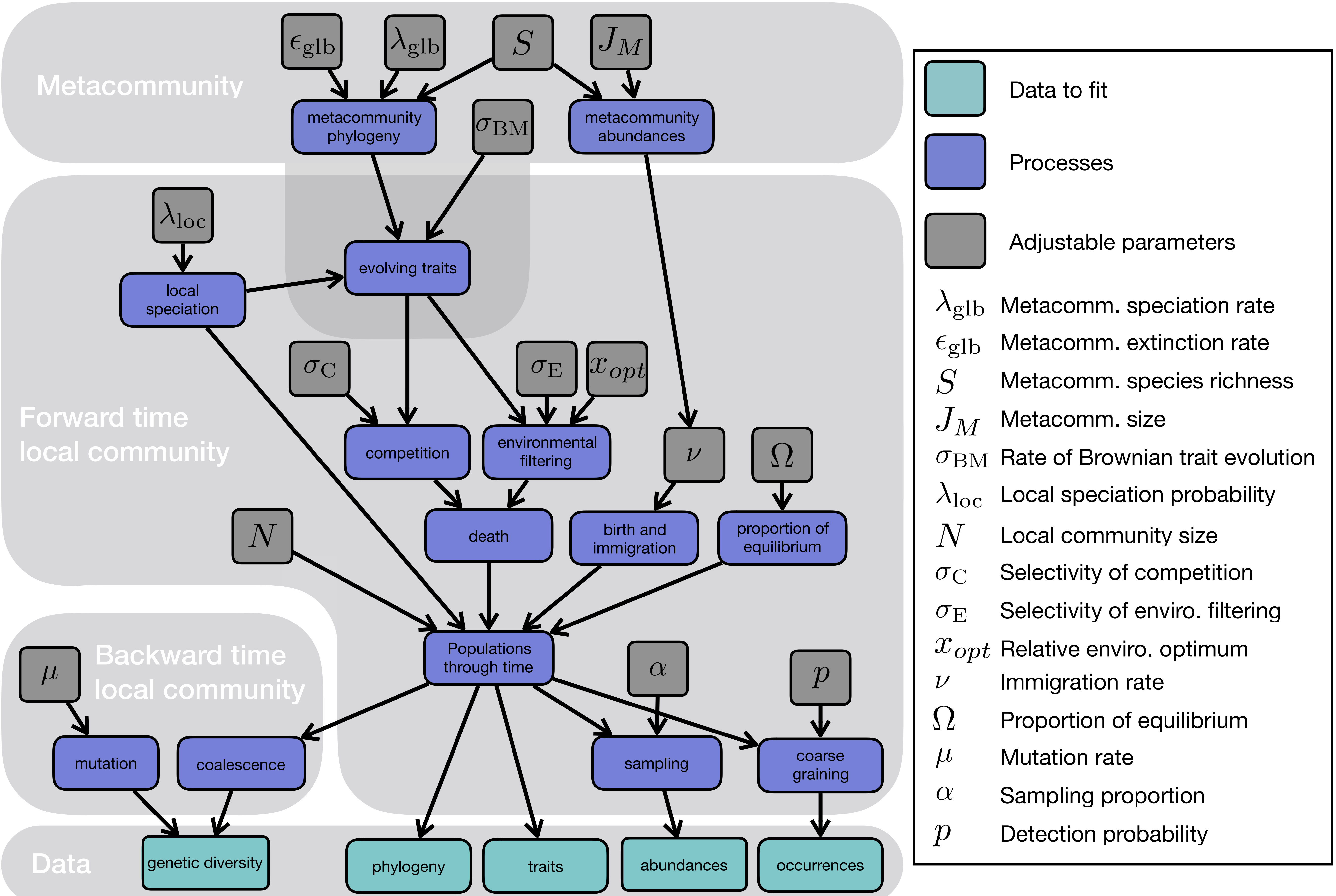




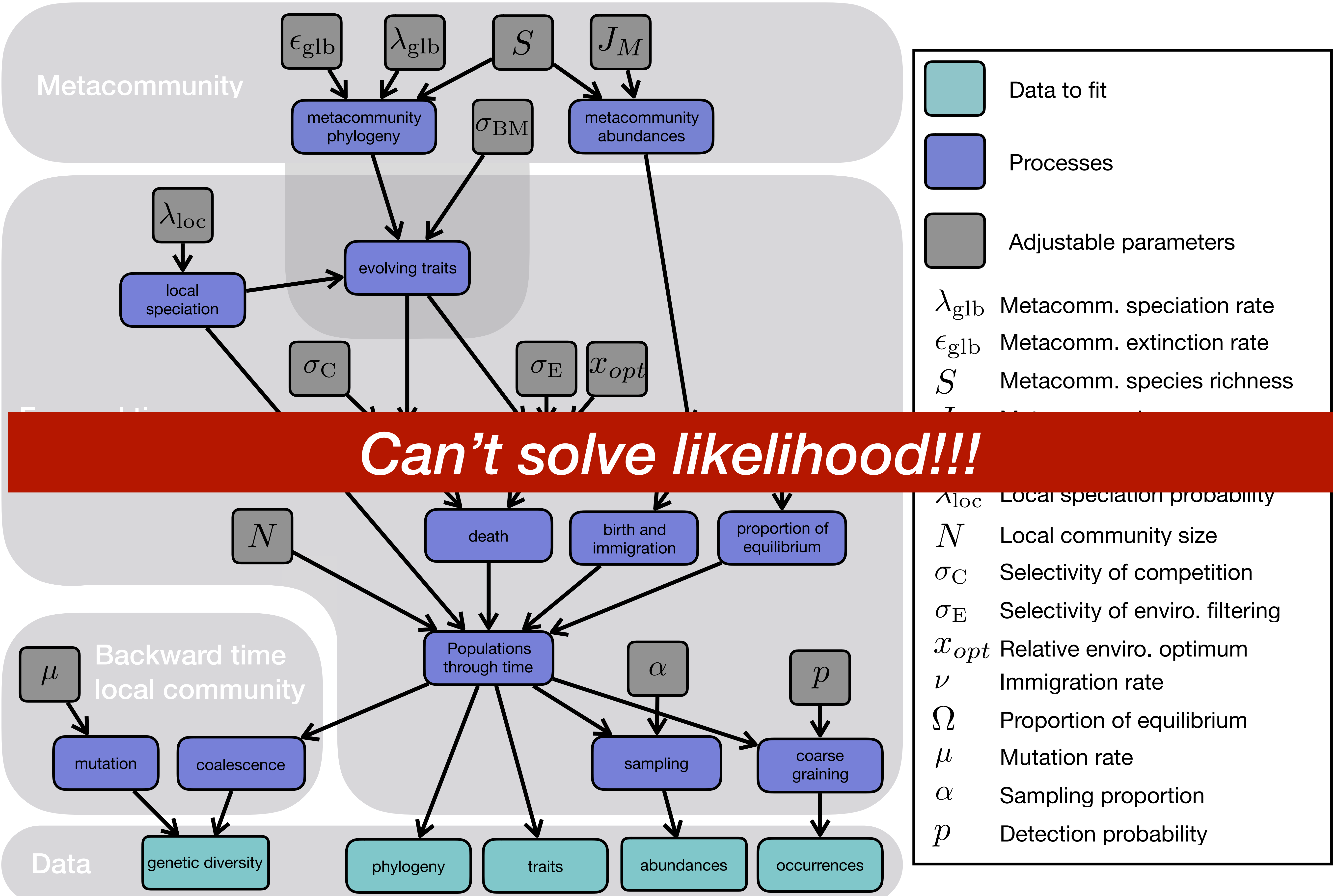
Rosindell, Harmon (2013) *Journal of Biogeography*

Population genetic processes





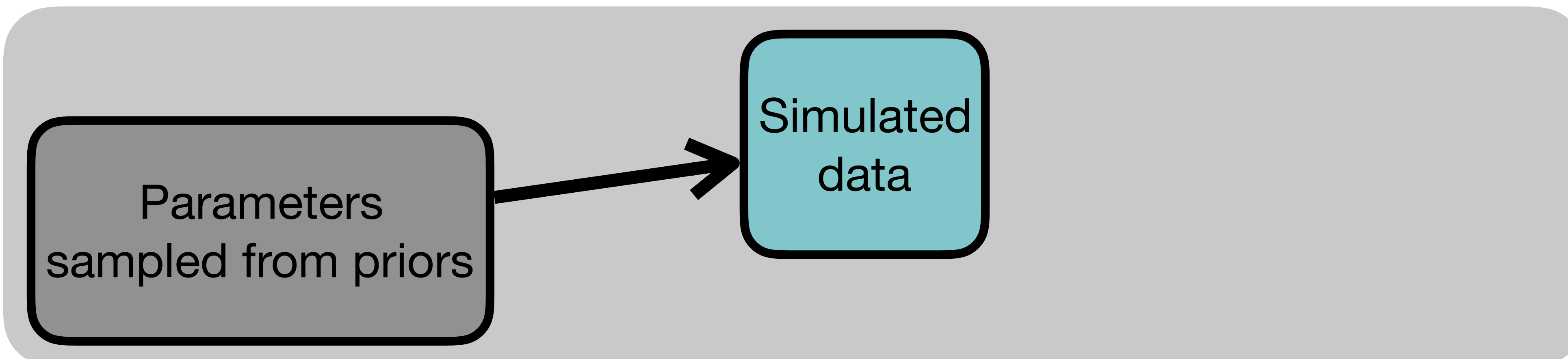
Metacommunity
$\epsilon_{\rm glb}$
$\lambda_{\rm glb}$
S
J_M
metacommunity phylogeny
$\sigma_{\rm BM}$
metacommunity abundances
$\lambda_{\rm loc}$
local speciation
evolving traits
$\sigma_{\rm C}$
$\sigma_{\rm E}$
x_{opt}
Forward time local community
competition
environmental filtering
ν
Ω
N
death
birth and immigration
proportion of equilibrium
Populations through time
α
p
μ
Backward time local community
mutation
coalescence
sampling
coarse graining
Data
genetic diversity
phylogeny
traits
abundances
occurrences
Data to fit
Processes
Adjustable parameters
$\lambda_{\rm glb}$ Metacomm. speciation rate
$\epsilon_{\rm glb}$ Metacomm. extinction rate
S Metacomm. species richness
J_M Metacomm. size
$\sigma_{\rm BM}$ Rate of Brownian trait evolution
$\lambda_{\rm loc}$ Local speciation probability
N Local community size
$\sigma_{\rm C}$ Selectivity of competition
$\sigma_{\rm E}$ Selectivity of enviro. filtering
x_{opt} Relative enviro. optimum
ν Immigration rate
Ω Proportion of equilibrium
μ Mutation rate
α Sampling proportion
p Detection probability



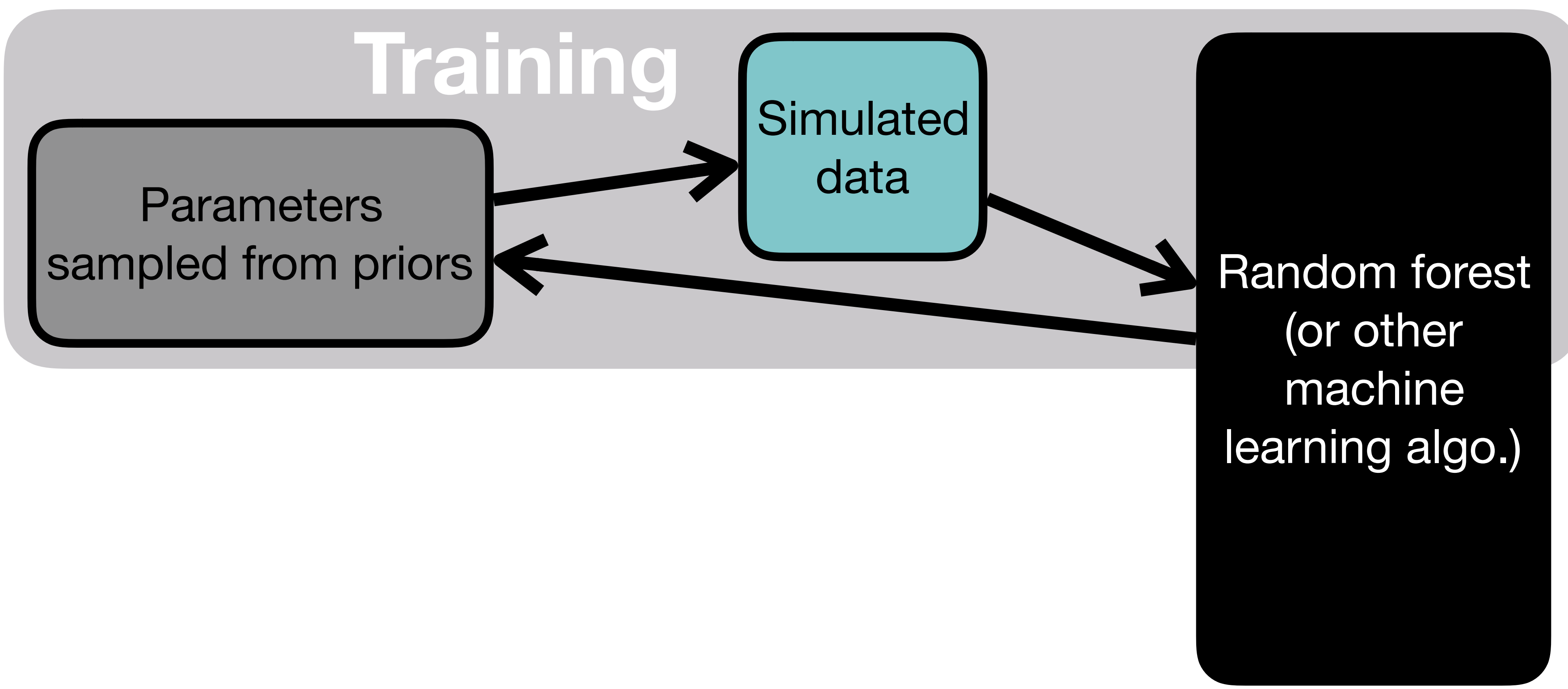
Metacommunity
ϵ_{glb}
λ_{glb}
S
J_M
metacommunity phylogeny
σ_{BM}
metacommunity abundances
λ_{loc}
local speciation
evolving traits
σ_{C}
σ_{E}
x_{opt}
Can't solve likelihood!!!
N
death
birth and immigration
proportion of equilibrium
Populations through time
Backward time local community
μ
mutation
coalescence
α
sampling
p
coarse graining
Data
genetic diversity
phylogeny
traits
abundances
occurrences
Data to fit
Processes
Adjustable parameters
λ_{glb} Metacomm. speciation rate
ϵ_{glb} Metacomm. extinction rate
S Metacomm. species richness
λ_{loc} Local speciation probability
N Local community size
σ_{C} Selectivity of competition
σ_{E} Selectivity of enviro. filtering
x_{opt} Relative enviro. optimum
ν Immigration rate
Ω Proportion of equilibrium
μ Mutation rate
α Sampling proportion
p Detection probability

Likelihood free inference

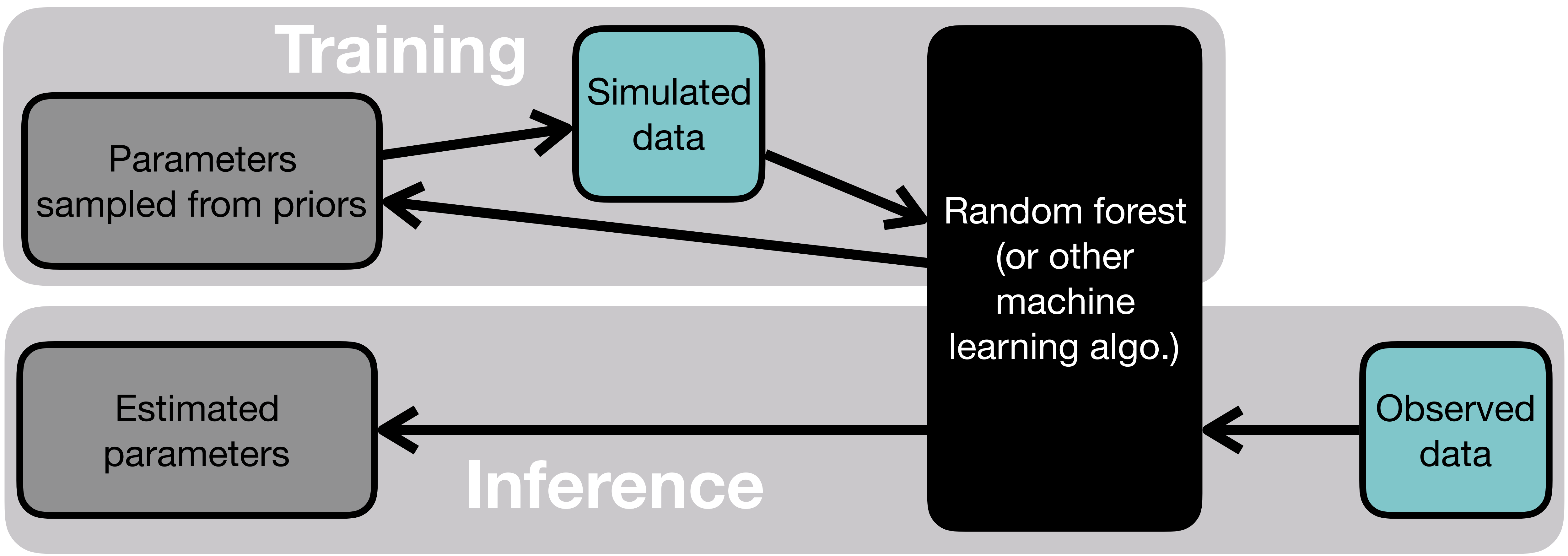
Likelihood free inference



Likelihood free inference



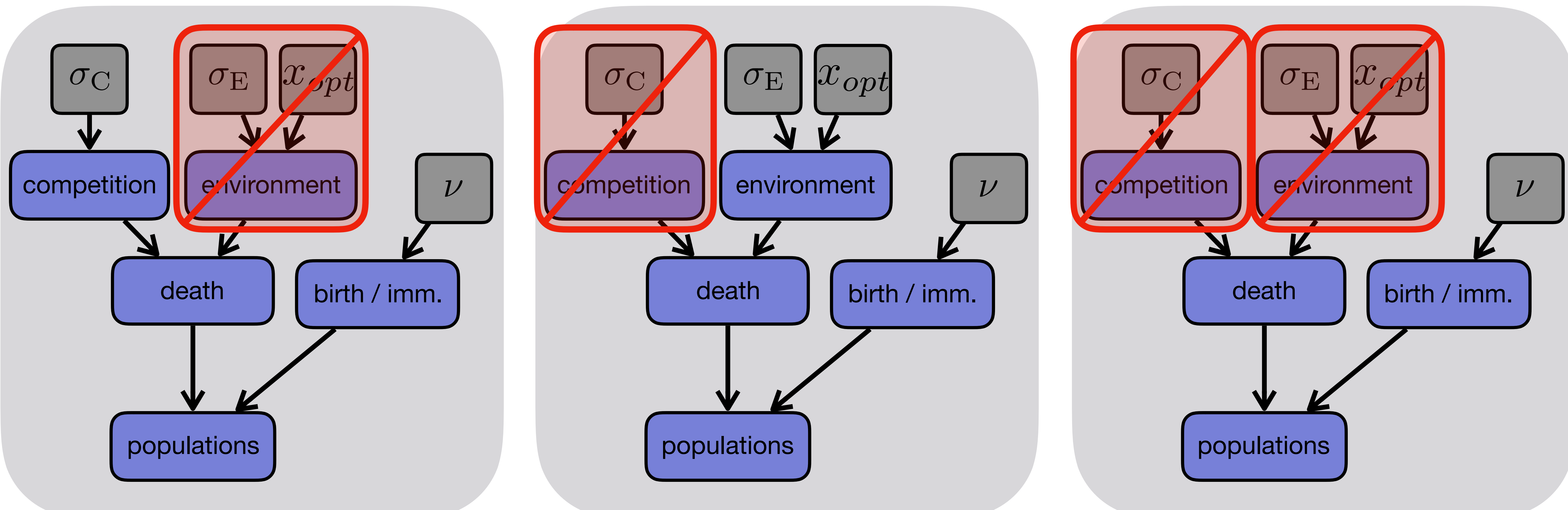
Likelihood free inference



A few MESS-y results

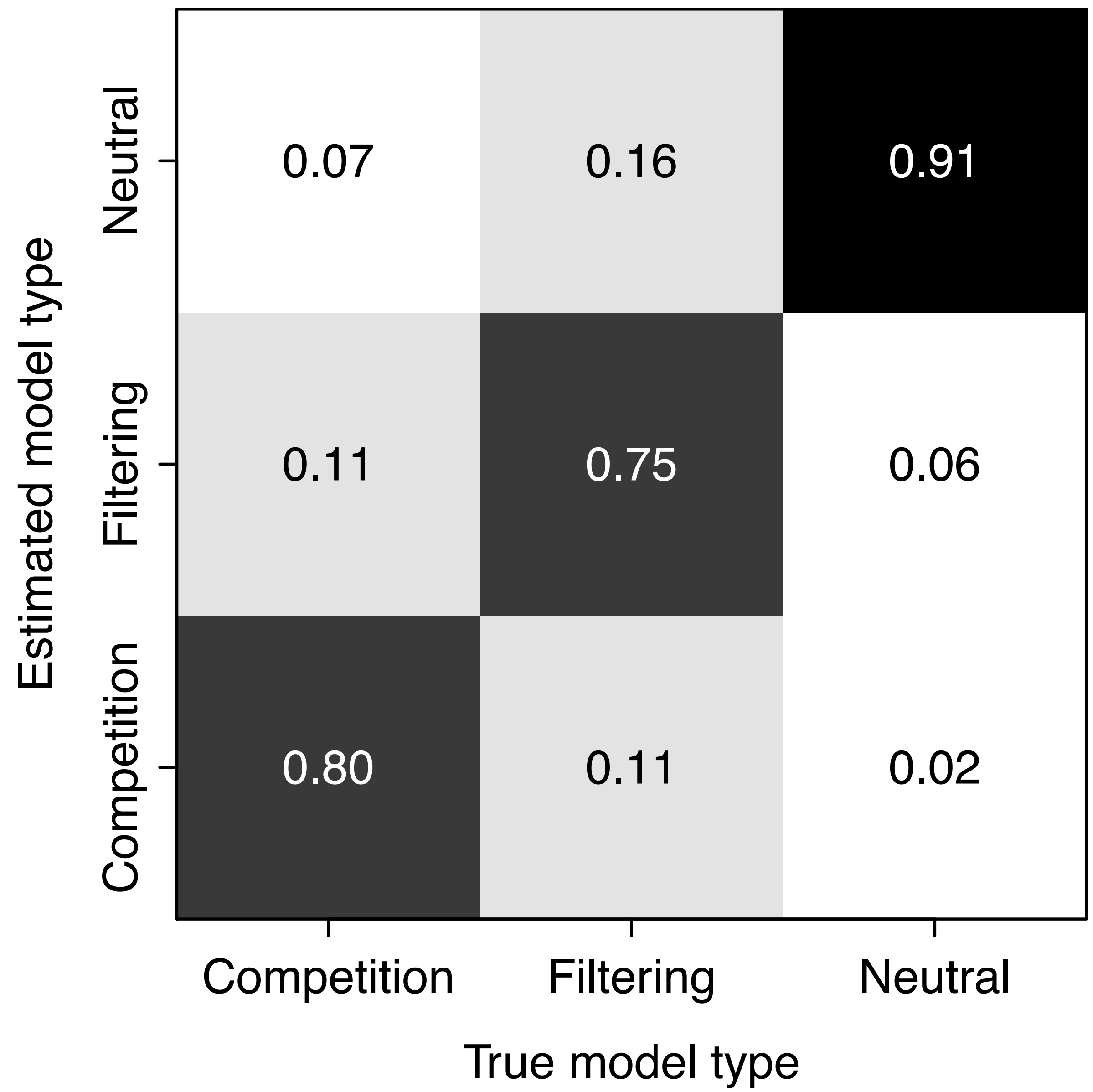
A few MESS-y results

Competition | Environmental filtering | Neutral



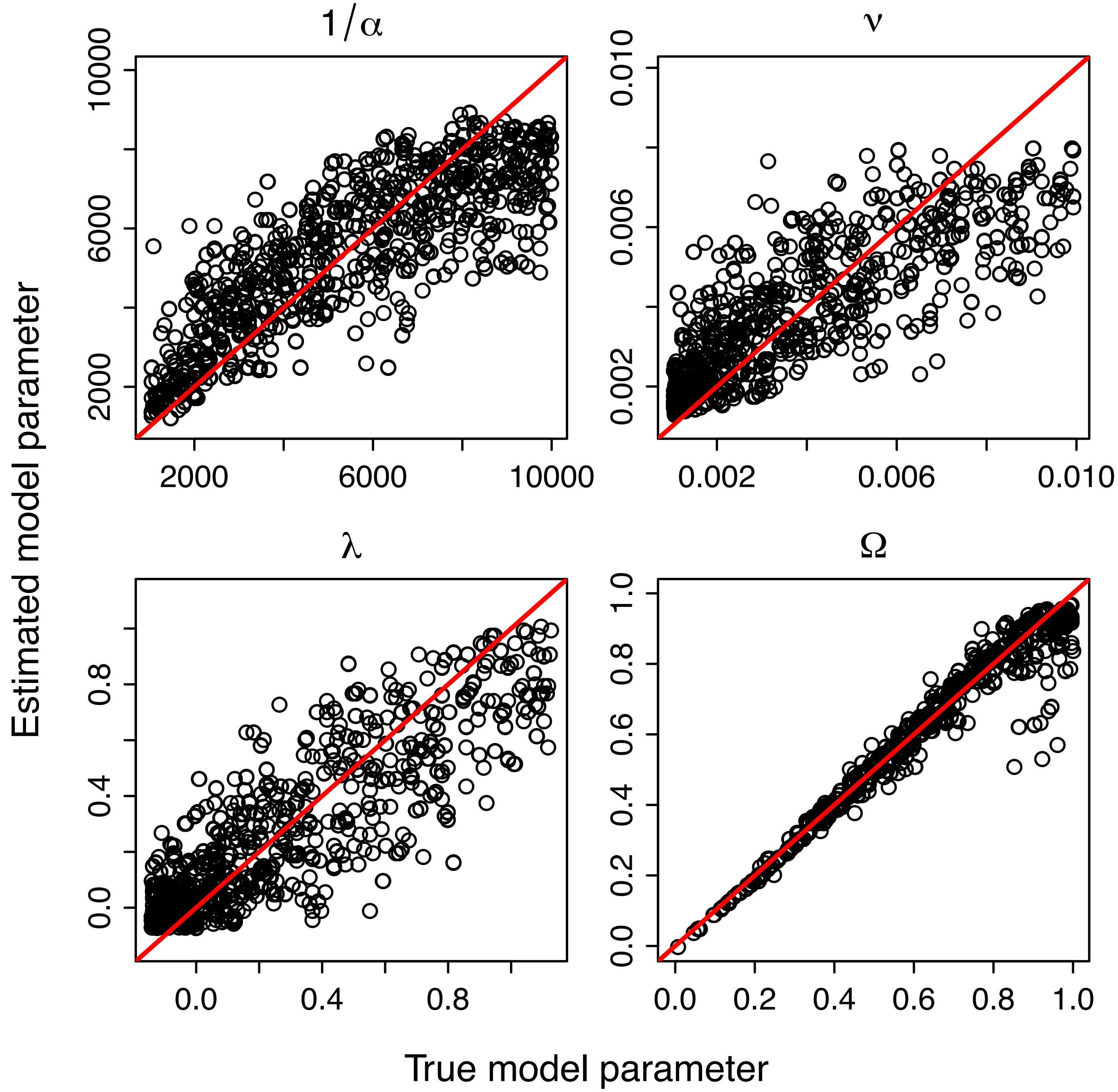
Overcast, …, Rominger. pre-print: biorxiv 10.1101/2020.01.30.927236

A few MESS-y results



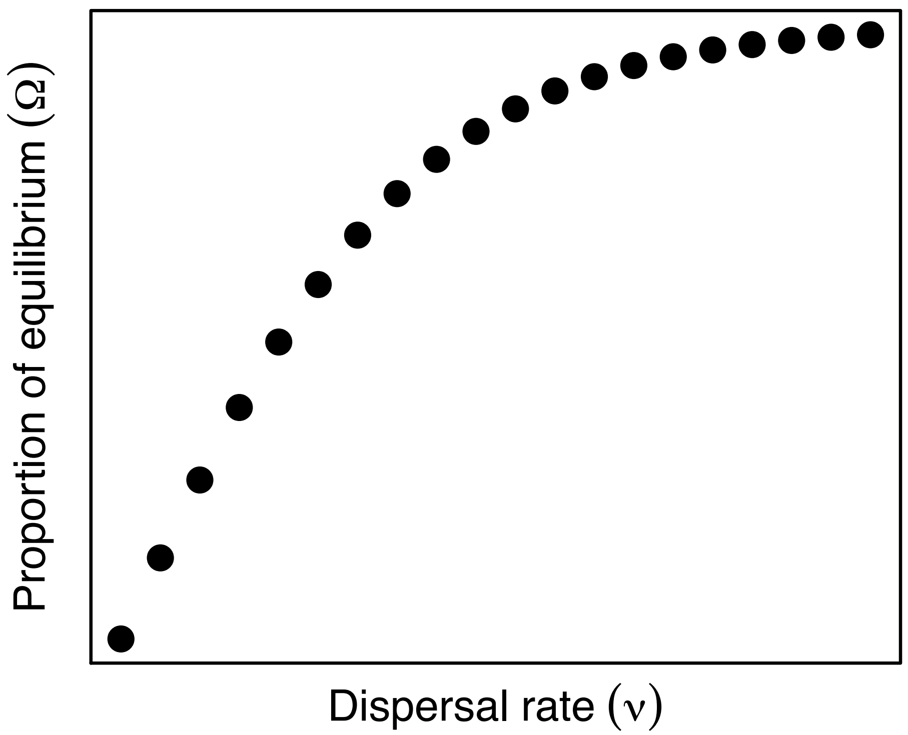
Overcast, …, Rominger. pre-print: biorxiv 10.1101/2020.01.30.927236

A few MESS-y results

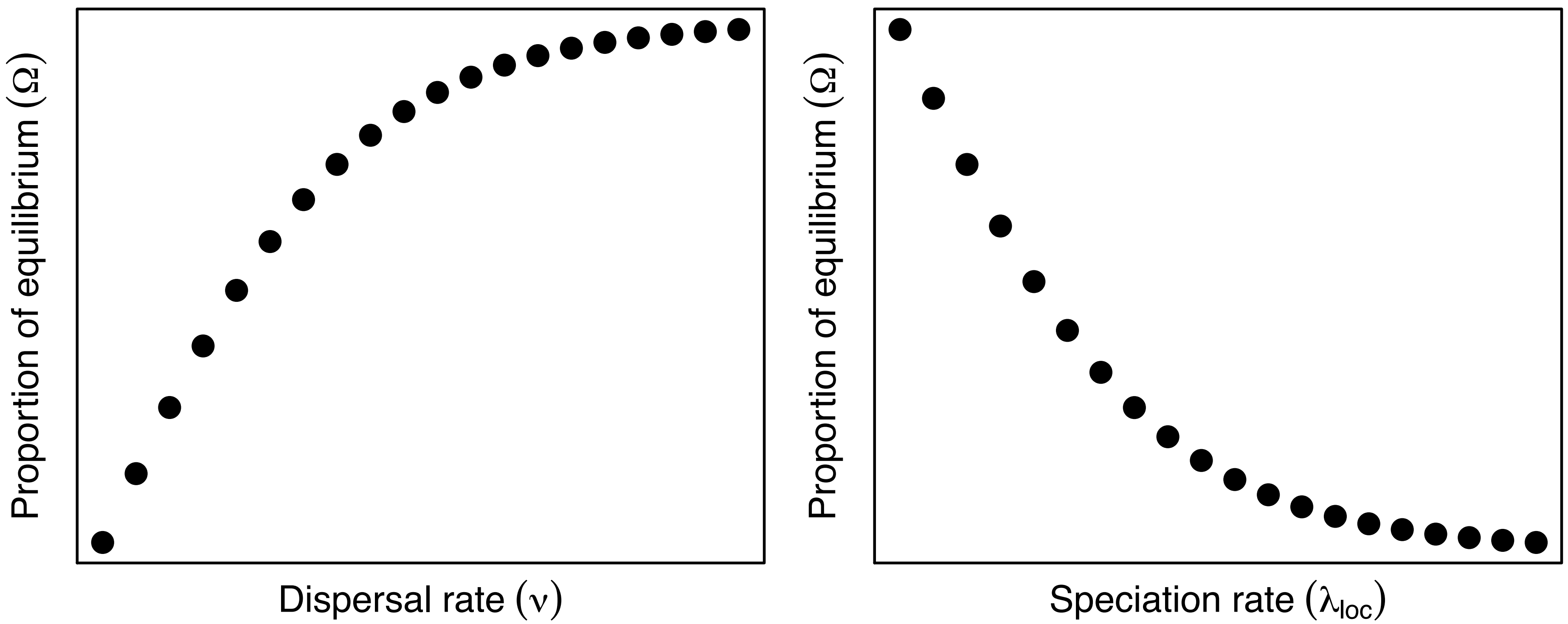


Overcast, …, Rominger. pre-print: biorxiv 10.1101/2020.01.30.927236

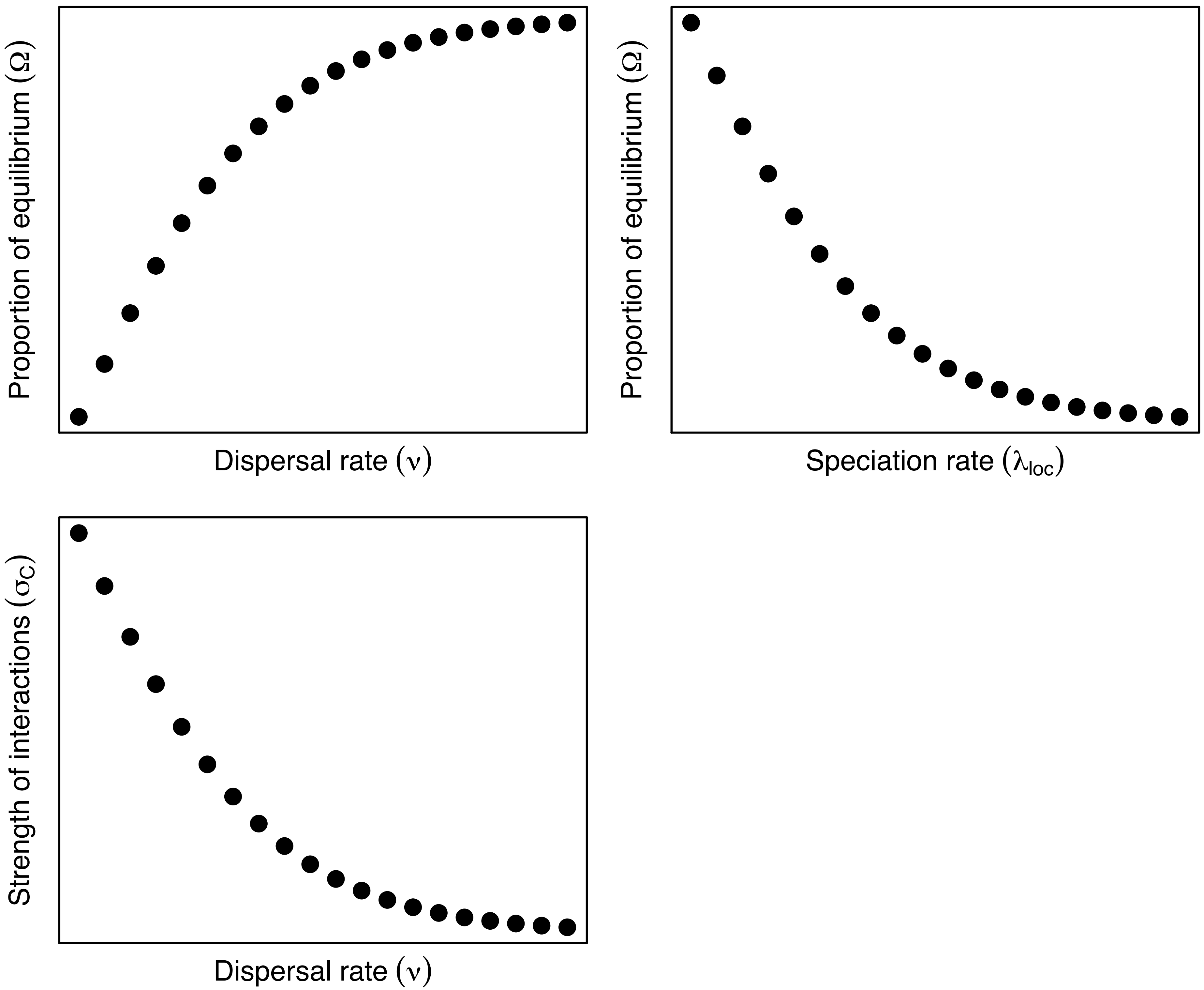
Studying biodiversity with MESS



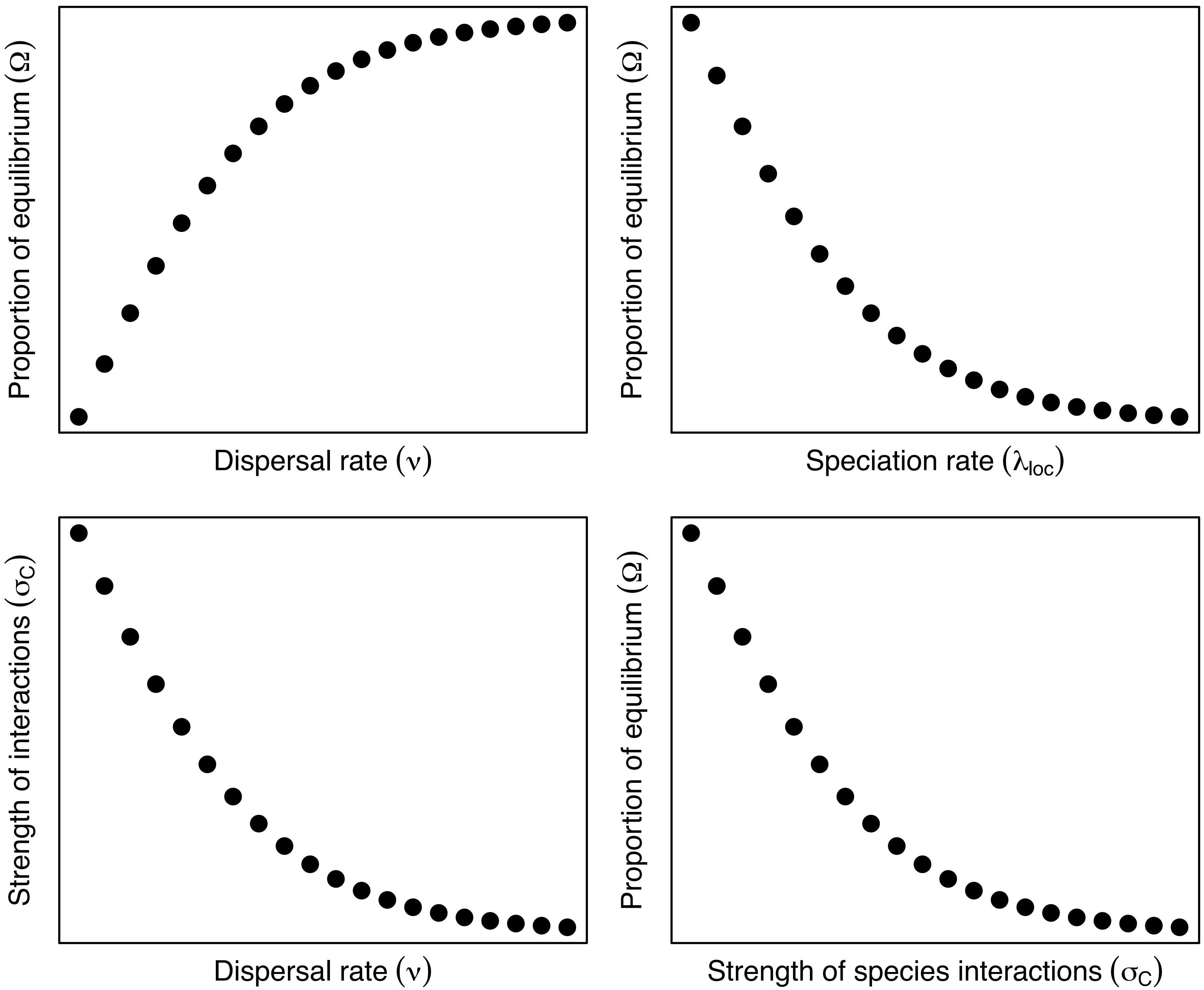
Studying biodiversity with MESS



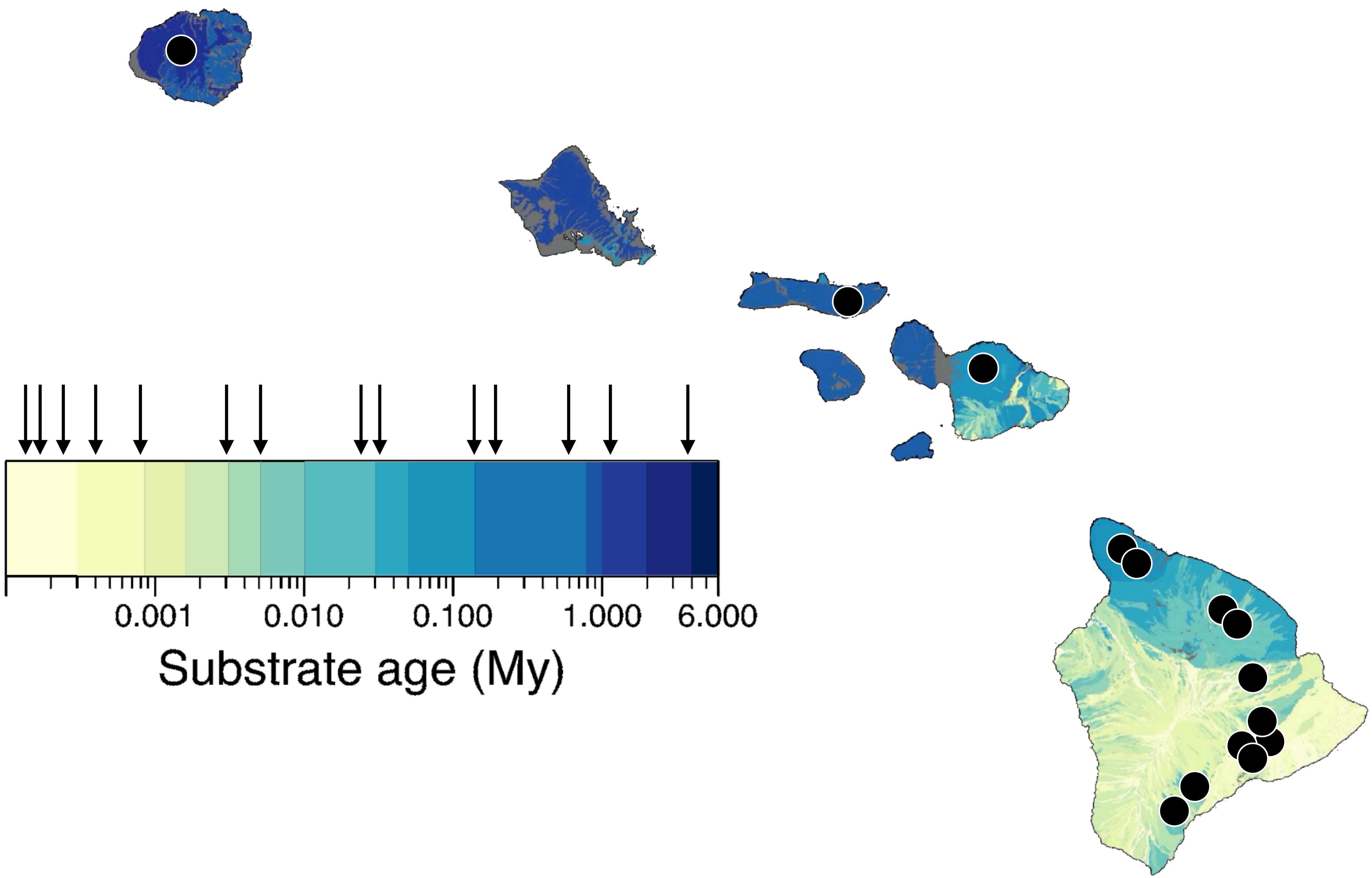
Studying biodiversity with MESS



Studying biodiversity with MESS



New opportunities



New opportunities

Student project ideas:

- Applying MESS to data from Hawaì`i

New opportunities

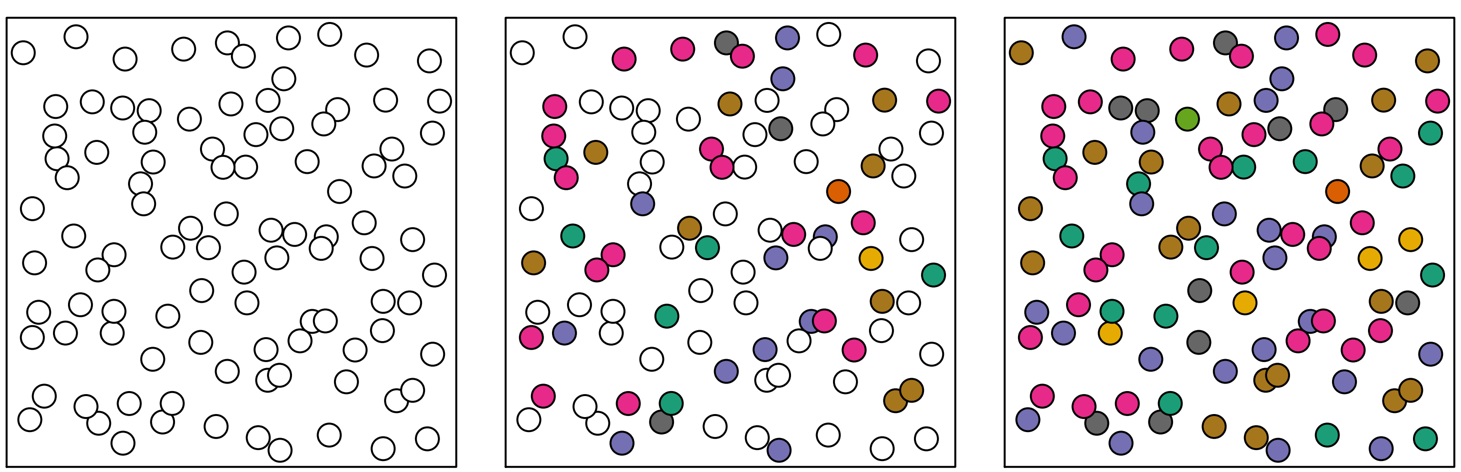
Student project ideas:

- Applying MESS to data from Hawaì`i
- Build eco-evo lesson plans using MESS

New opportunities

Student project ideas:

- Applying MESS to data from Hawai`i
- Build eco-evo lesson plans using MESS
- Creating alternative definitions of equilibrium



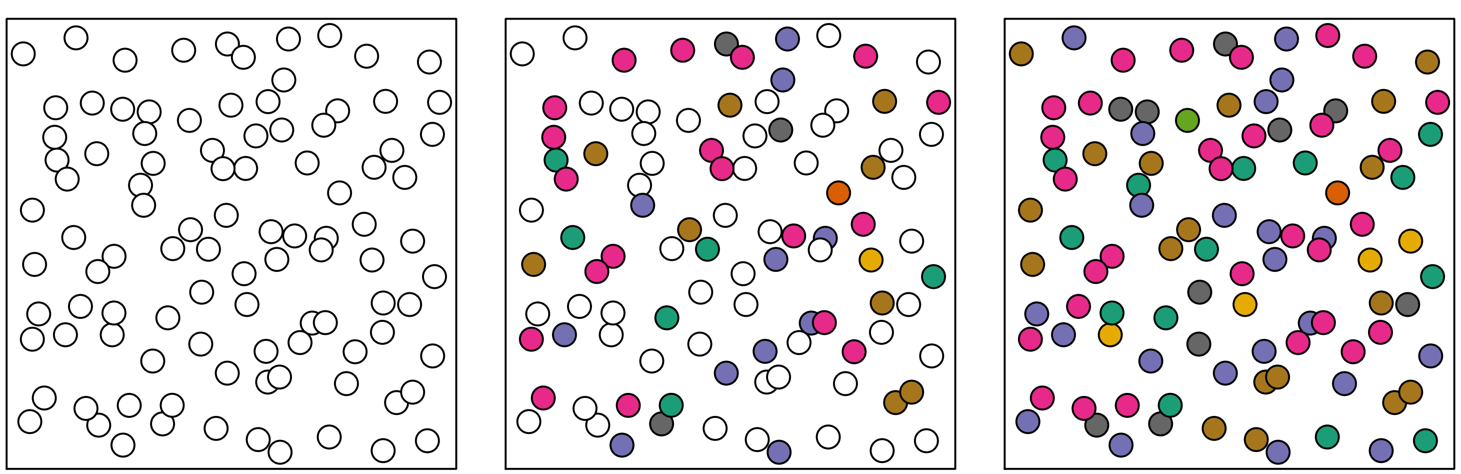
What about…?

CTCTCAGTGC
GTCACAGTGC
GTCACAGTTC
GACACAGTTC
GACAGAGTTC

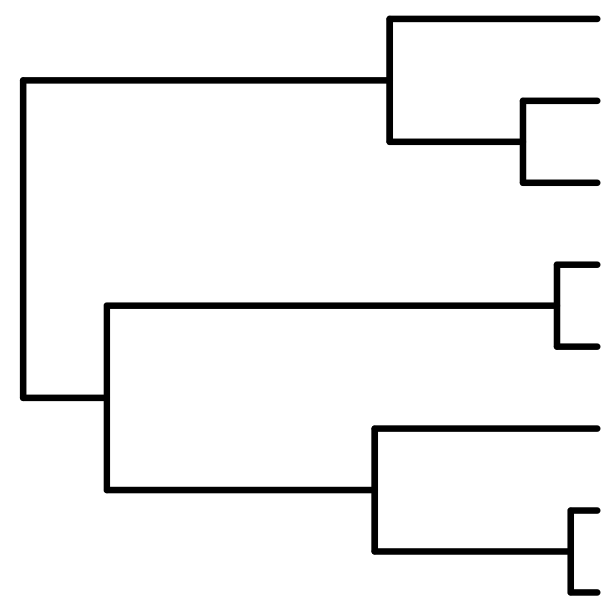
New opportunities

Student project ideas:

- Applying MESS to data from Hawai`i
- Build eco-evo lesson plans using MESS
- Creating alternative definitions of equilibrium



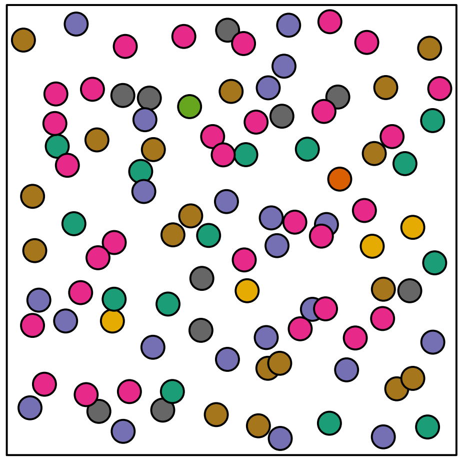
What about…?



New opportunities

Student project ideas:

- Applying MESS to data from Hawaì`i
- Build eco-evo lesson plans using MESS
- Creating alternative definitions of equilibrium
- Does the equilibrium state correspond to the MaxEnt prediction?



$$\Longleftrightarrow \quad R(n, \varepsilon) = \frac{1}{Z} e^{-\lambda_1 n - \lambda_2 n \varepsilon}$$

Mahalo!

J. Harte

R. Gillespie

D. Gruner

L. Schneider

I. Overcast

M. Hickerson

J. Rosindell L. Harmon

A. Carnaval