

Fitting mechanistic models



A general approach to model fitting

Simulate data given a model and some parameters

Evaluate how likely the observed data are, given the simulated data

Tweak the parameters/model to make the observations more likely

Is this model superior to other models?

Maximum likelihood

Parameter values that make the observed data most likely

Notation

Observed variable D depends on unknown parameter θ

Probability density function of D is f_θ

$$L_d(\theta) = f_\theta(d)$$

What's D ?

Case counts at different times

Sequences

Titers

or some composite of observations

Maximizing the likelihood

means maximizing the log-likelihood

or minimizing the negative log-likelihood

Finding the maximum likelihood

Can be analytically tractable

For our models, it's not

Likelihood optimization

Brute force

Derivative-based methods

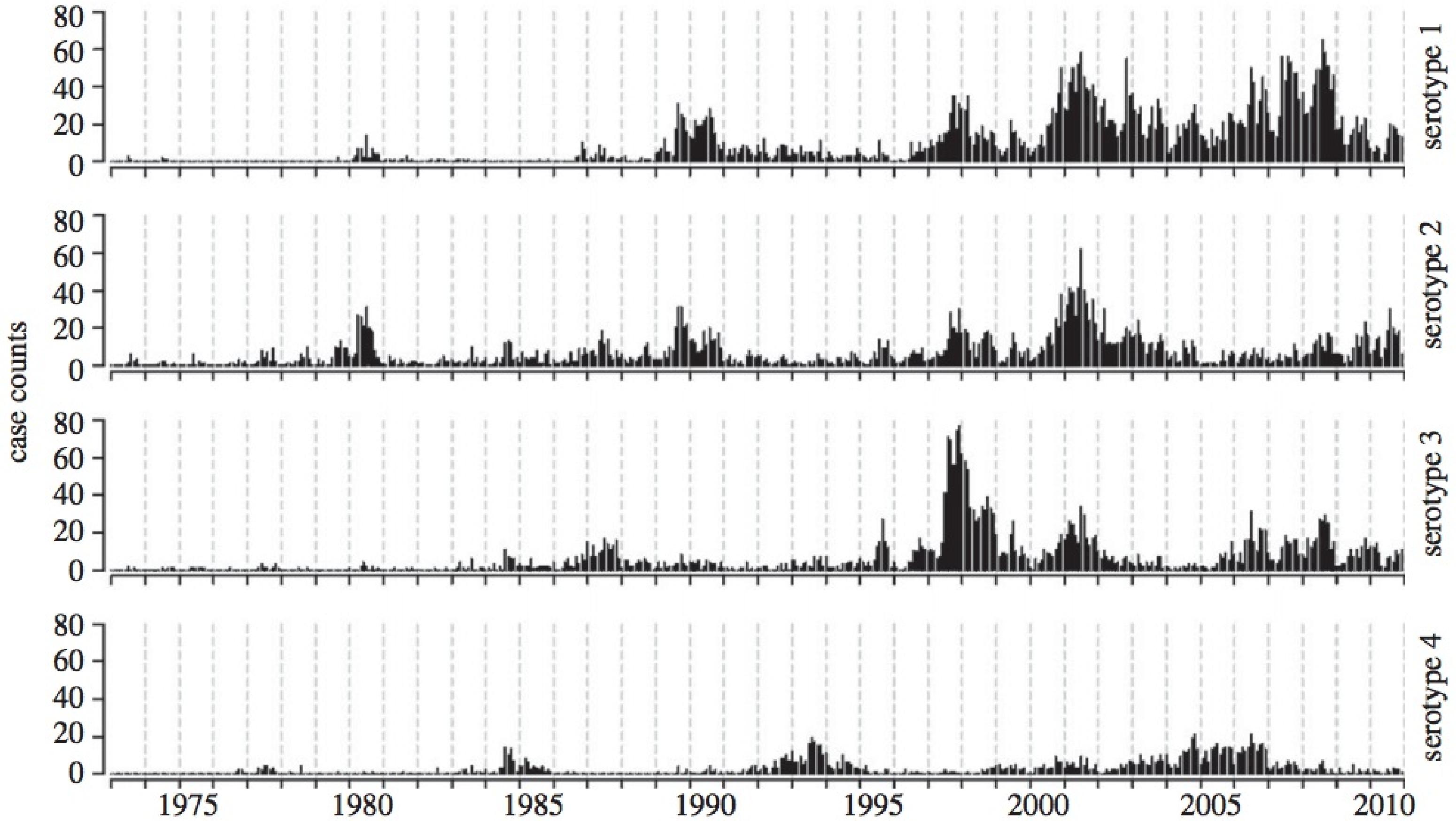
Simplex

Simulated annealing

Sequential Monte Carlo

Many others... but for us, few tried and true

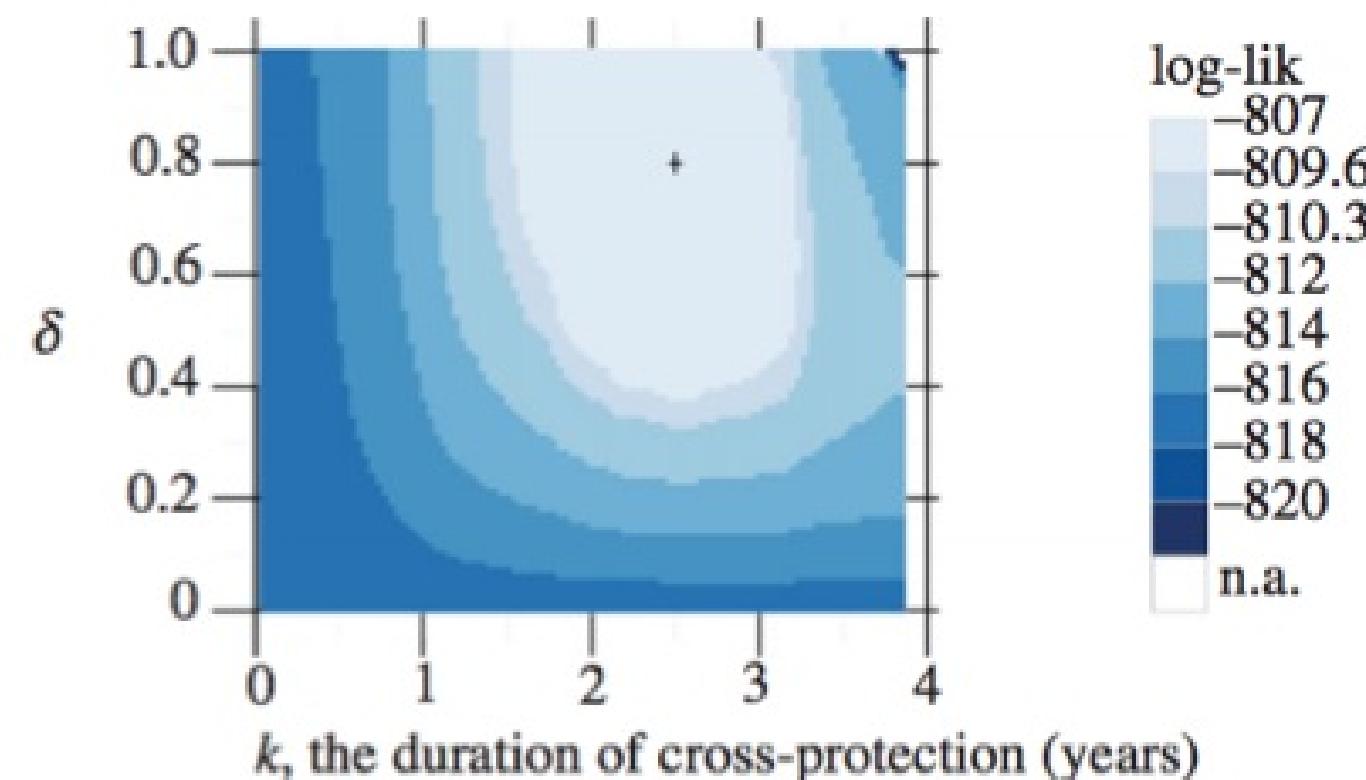
Interacting dengue serotypes



Reich et al. 2013

Likelihood profiles

Hold parameter(s) constant, fit the rest



Reich et al. 2013

Bayesian approaches

Assume prior expectations about our parameters, $p(\theta)$

$$p(\theta|D) \propto p(D|\theta)p(\theta)$$

where $p(D|\theta)$ is the likelihood
and $p(\theta|D)$ is the posterior density

Bayesian optimization

Markov Chain Monte Carlo

Metropolis-Hastings MCMC

Particle MCMC

Hybrid/Hamiltonian Monte Carlo

Many others

Introduction to MCMC

Likelihood with longitudinal data

Data: Infection and clearance times (t_i, t_c) for many people

Model: Force of infection $\lambda_i = \beta I_i$

Gamma-distributed duration of infection (shape k , scale θ)

$$L(\lambda, k, \theta | t_i, t_c) \propto \prod_{\text{time to infection}}^n e^{-\lambda t_i} \underbrace{((t_c - t_i)^{(k-1)} e^{(t_c - t_i)/\theta})}_{\text{infection duration}}$$

This model is biased. Can you see why?

Simulated-based inference

e.g., R-package **pomp**

Fit models to time series or (new) longitudinal data

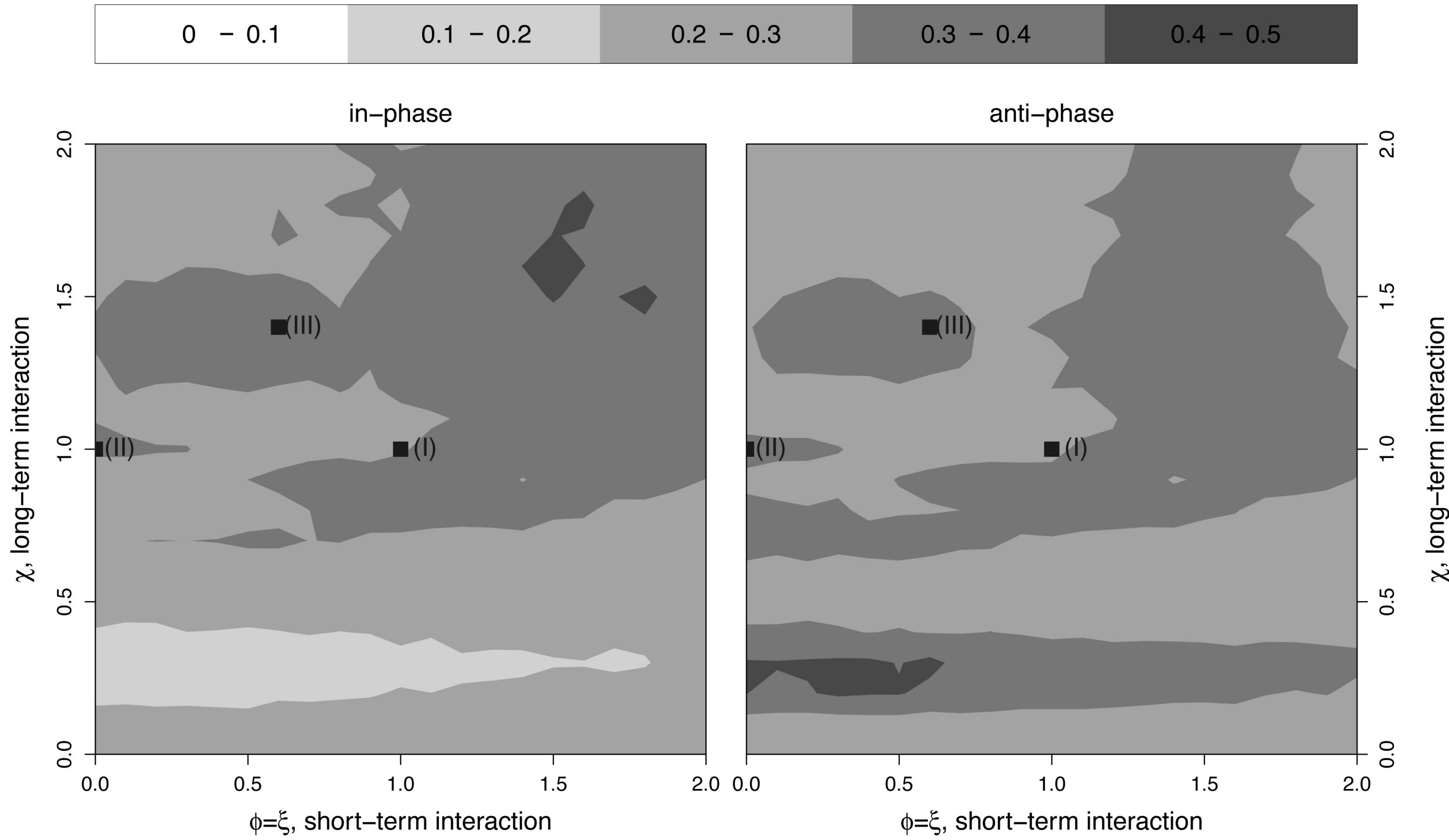
King and Ionides (SISMID module)

Probes and arbitrary metrics

Approximate Bayesian Computation

Time-series probes

Phases insufficient to infer interaction



Challenges fitting multistrain models

Many parameters

Multiple minima

Noise, nonstationarity, etc.

When should we trust a model?

Model "validation"

Confirm convergence

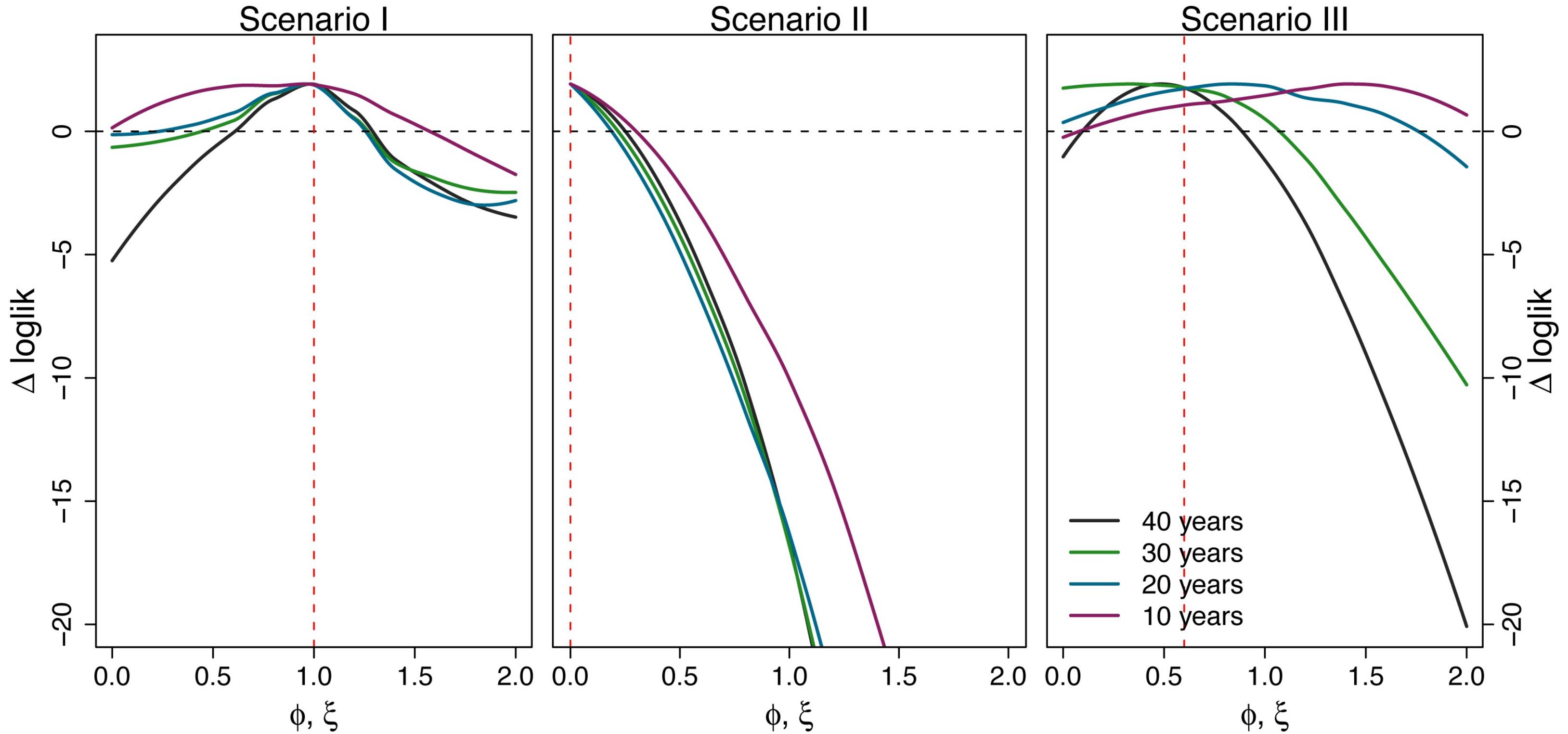
AIC and WAIC

Leave-one-out cross-validation

Out-of-sample prediction

Replicate on simulated data

Gauge the power of your data



Shrestha et al. 2011

(but this sounds so hard)

State-space reconstruction

A system with two state variables

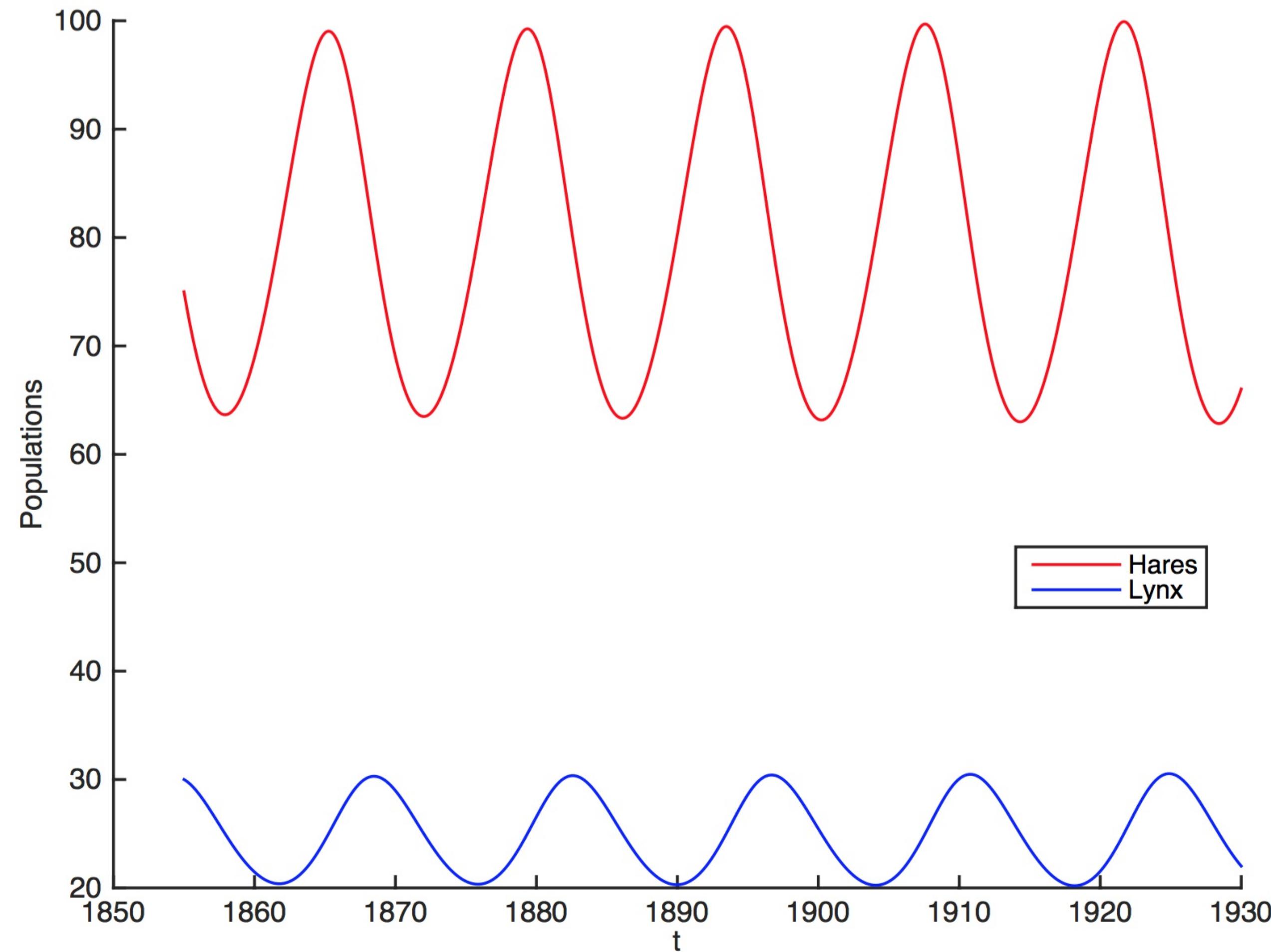
$$H' = aH - bHL$$

$$L' = cHL - dL$$

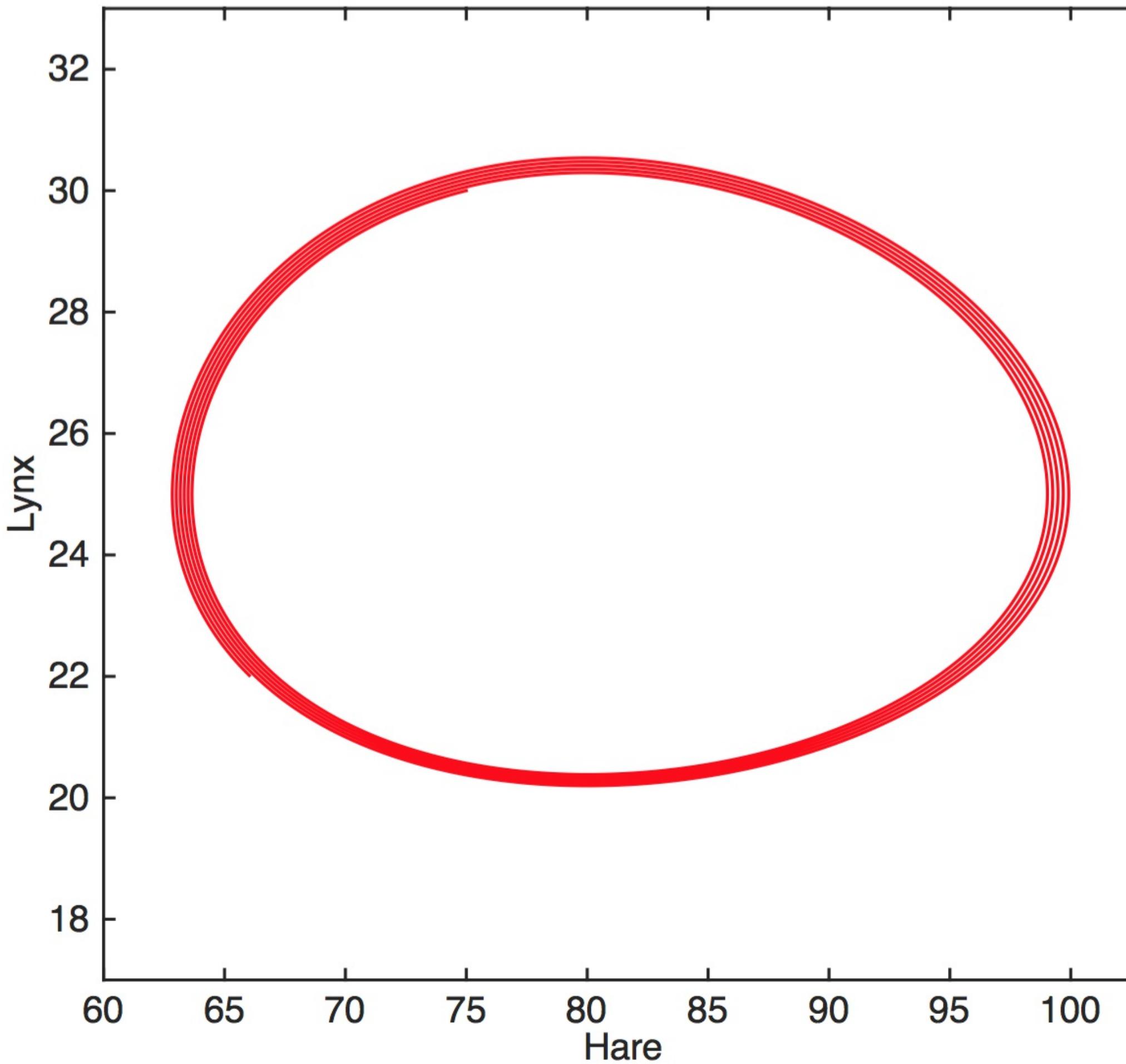
H hares, L lynxes

hare birth rate a , predation rate b ,
consumption rate c , death rate d

Solve for $H(t)$, $L(t)$ by numerical integration



Attractor is a limit cycle



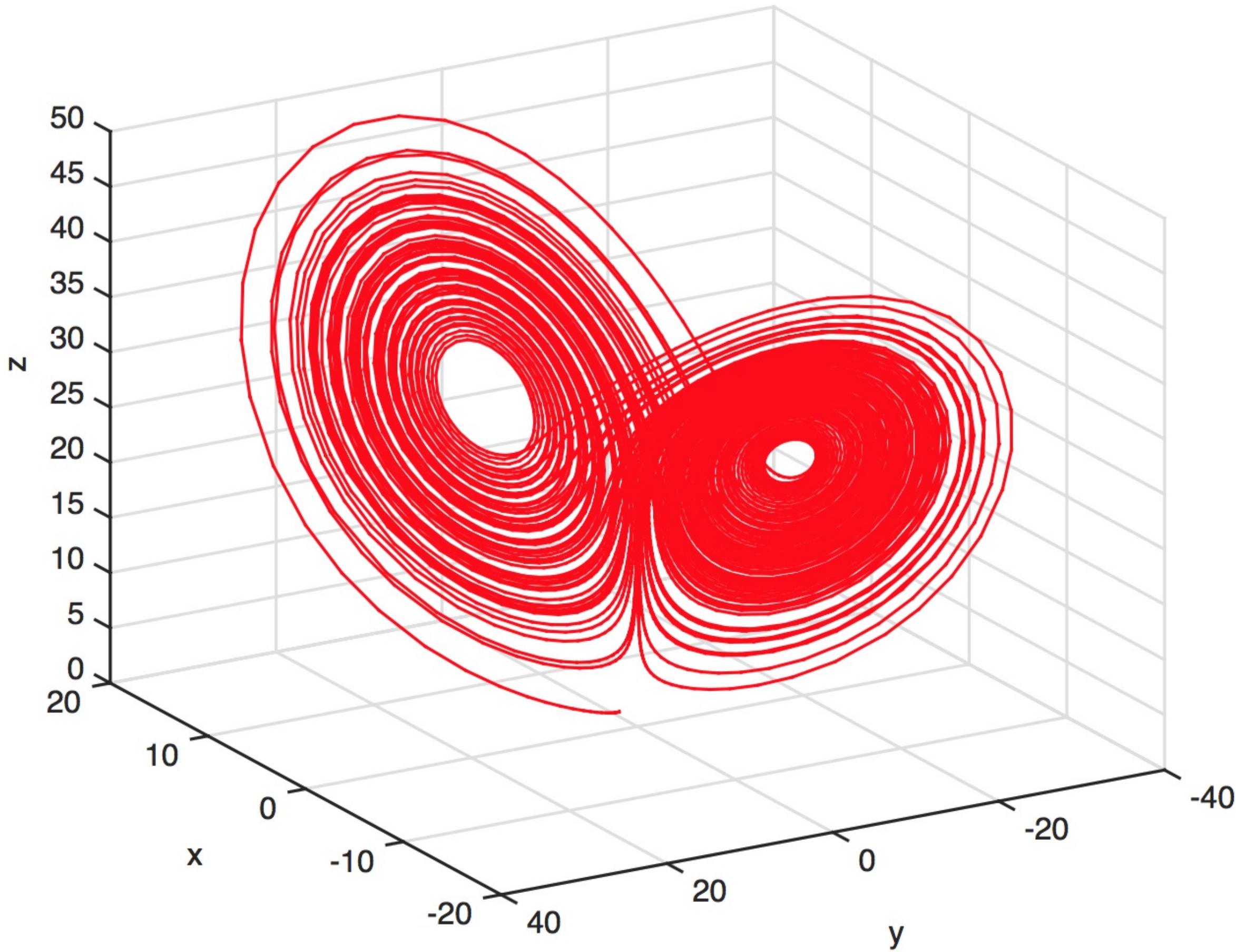
A more complex system

$$x' = \sigma(y - x)$$

$$y' = x(\rho - z) - y$$

$$z' = xy - \beta z$$

The Lorenz attractor



Implications of state-space reconstruction

We can detect underlying structure

We can detect and predict without understanding

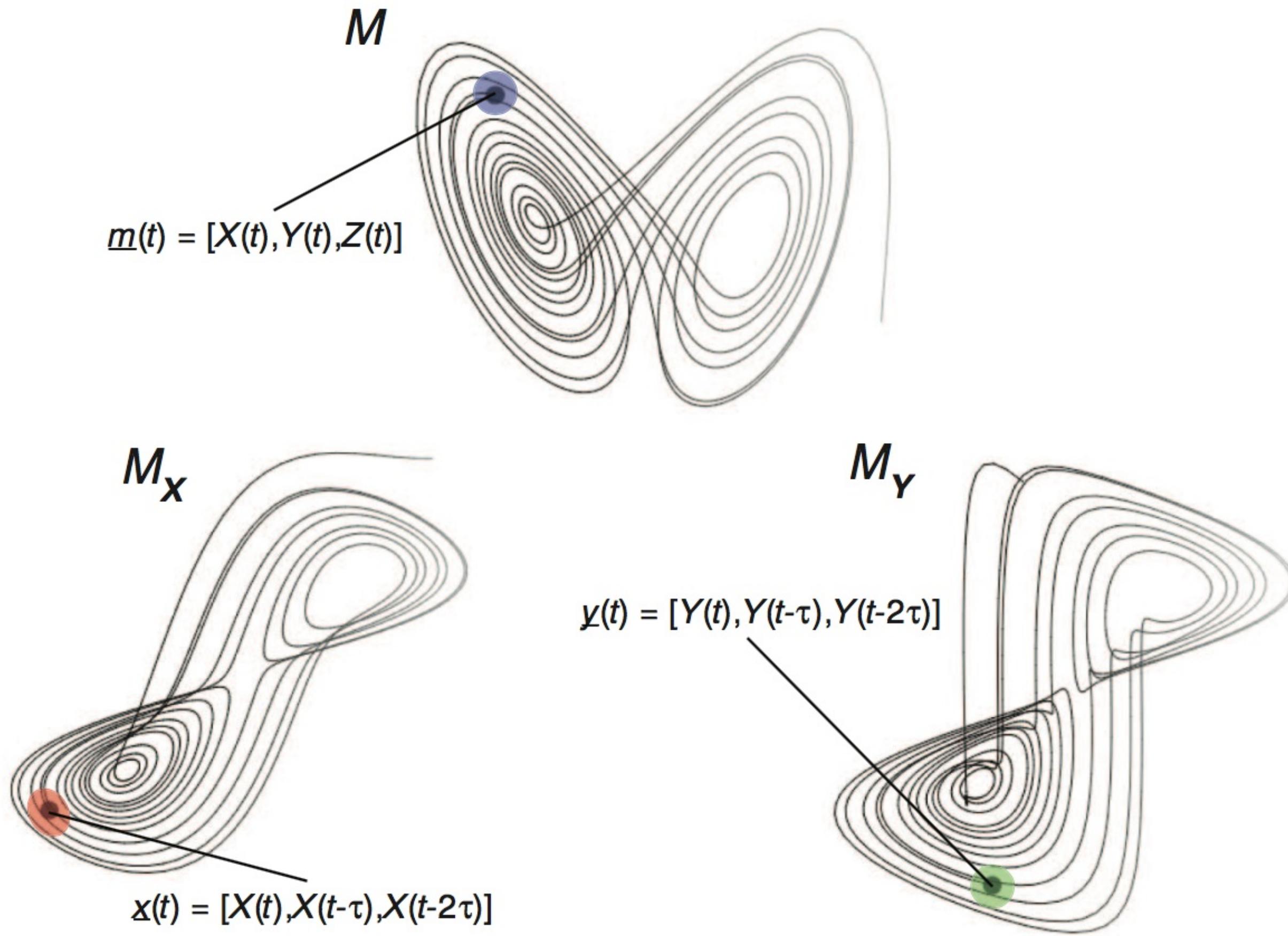
New claim: We can infer causal interactions

Takens' theorem

Very roughly, the system's attractor is **diffeomorphic to** (can be mapped without loss of information) to the individual attractors of the state variables in some delay-embedding space.

Takens 1981

Manifolds and shadow manifolds



Introduction to Takens' Theorem

Causal inference

Through their shadow manifolds, variables in the same dynamical system can predict each other.

If X drives Y , increasing the number of observations of Y should improve predictions of states of X .

Sugihara et al. 2012

Convergent cross-mapping

To infer if X drives Y :

- Construct the shadow manifold of Y , \mathbf{M}_Y (for some E, τ).
(Each point in \mathbf{M}_Y is given by
$$\vec{y}(t) = \{y_t, y_{t-\tau}, y_{t-2\tau}, \dots, y_{t-(E-1)\tau}\}.$$
)
- For each $X(t)$, identify its analogues $\vec{x}(t)$ and $\vec{y}(t)$.
- Find the $E + 1$ nearest neighbors of $\vec{y}(t)$ and weight them by their Euclidean distances to $\vec{y}(t)$.
- To make a prediction $\hat{X}(t)$, multiply these weights by the respective points in \mathbf{M}_X . Let ρ be the correlation between $\vec{x}(t)$ and $\hat{X}(t)$.
- First make predictions from \mathbf{M}_Y constructed with only a few points in the time series, L_{\min} , and then with many, L_{\max} .
- If ρ increases with more information on \mathbf{M}_Y , X drives Y .

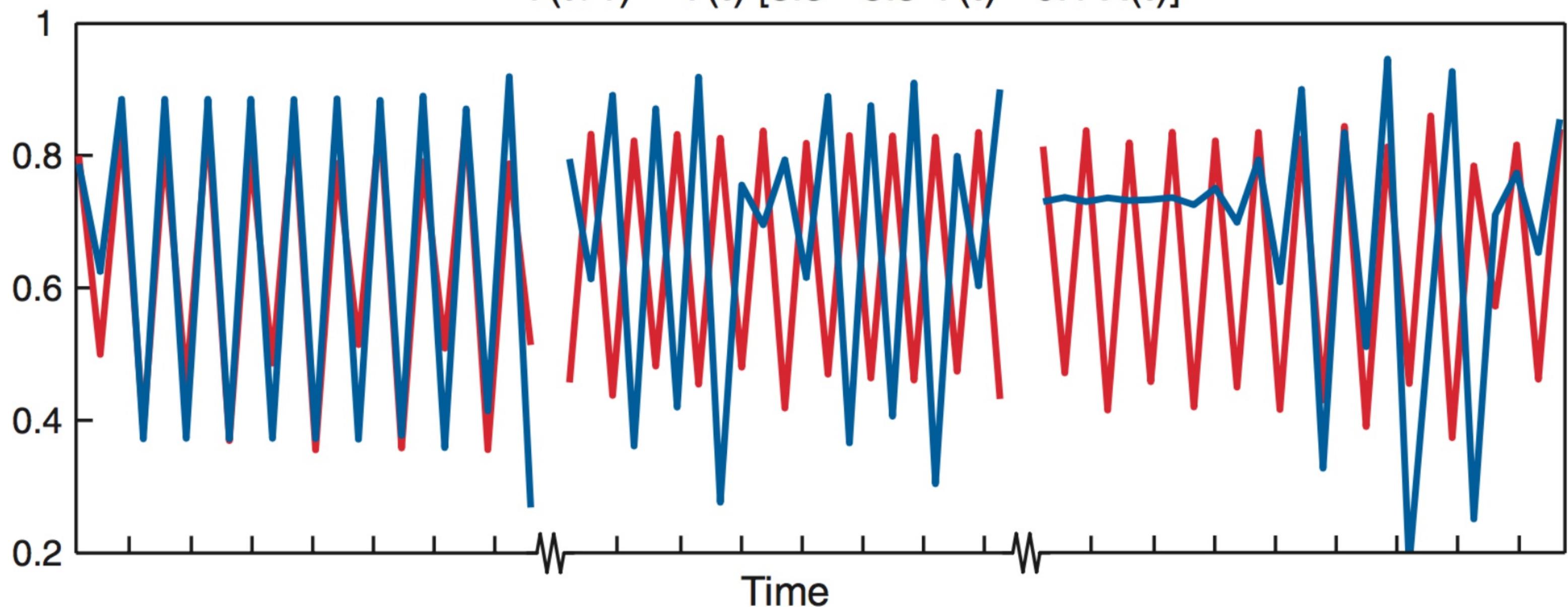
Introduction to convergent cross-mapping

What do you expect ρ to converge to?

Deterministic toy model

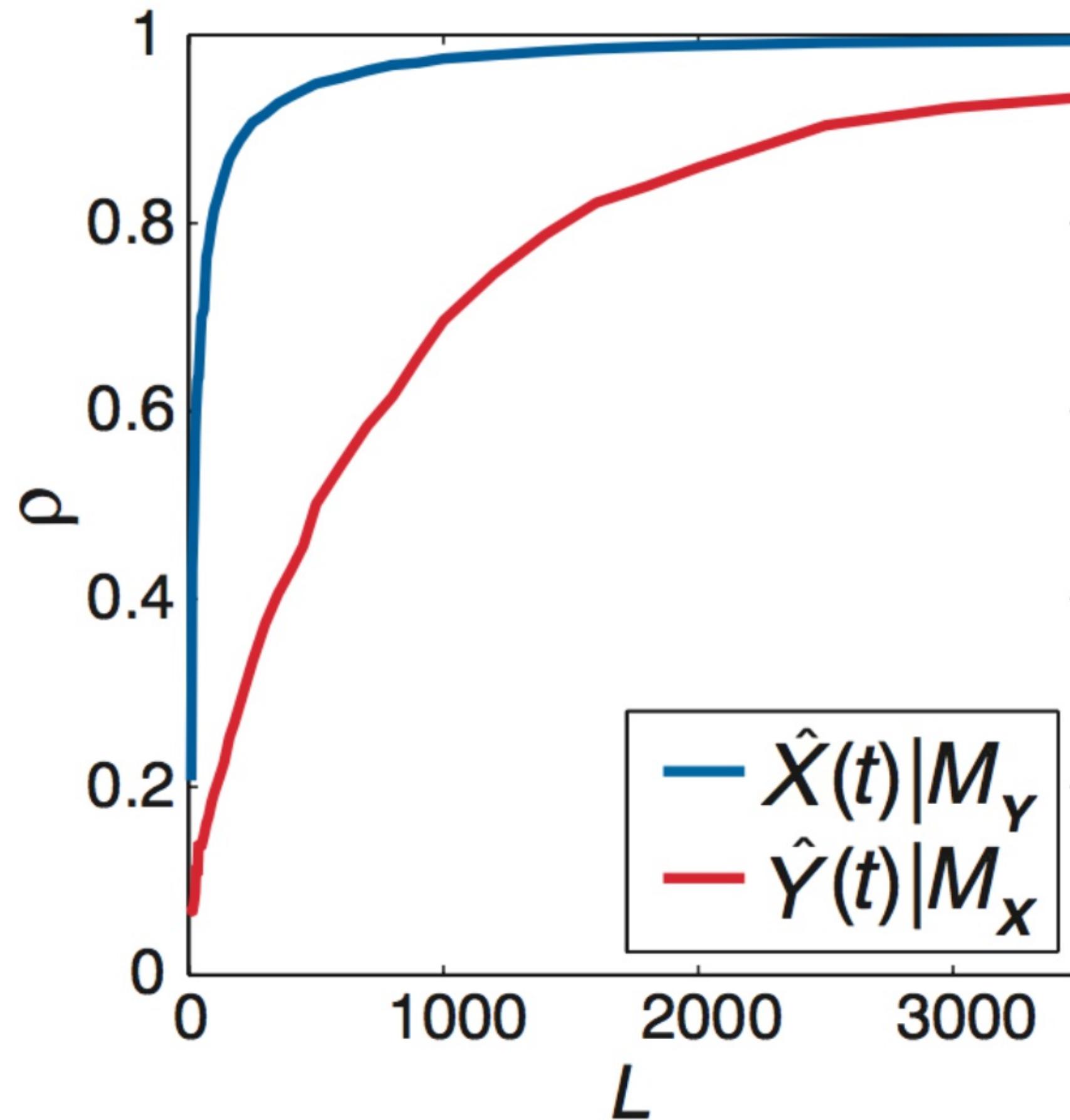
$$X(t+1) = X(t) [3.8 - 3.8 X(t) - 0.02 Y(t)]$$

$$Y(t+1) = Y(t) [3.5 - 3.5 Y(t) - 0.1 X(t)]$$



Sugihara et al. 2012

Under determinism, perfect predictability

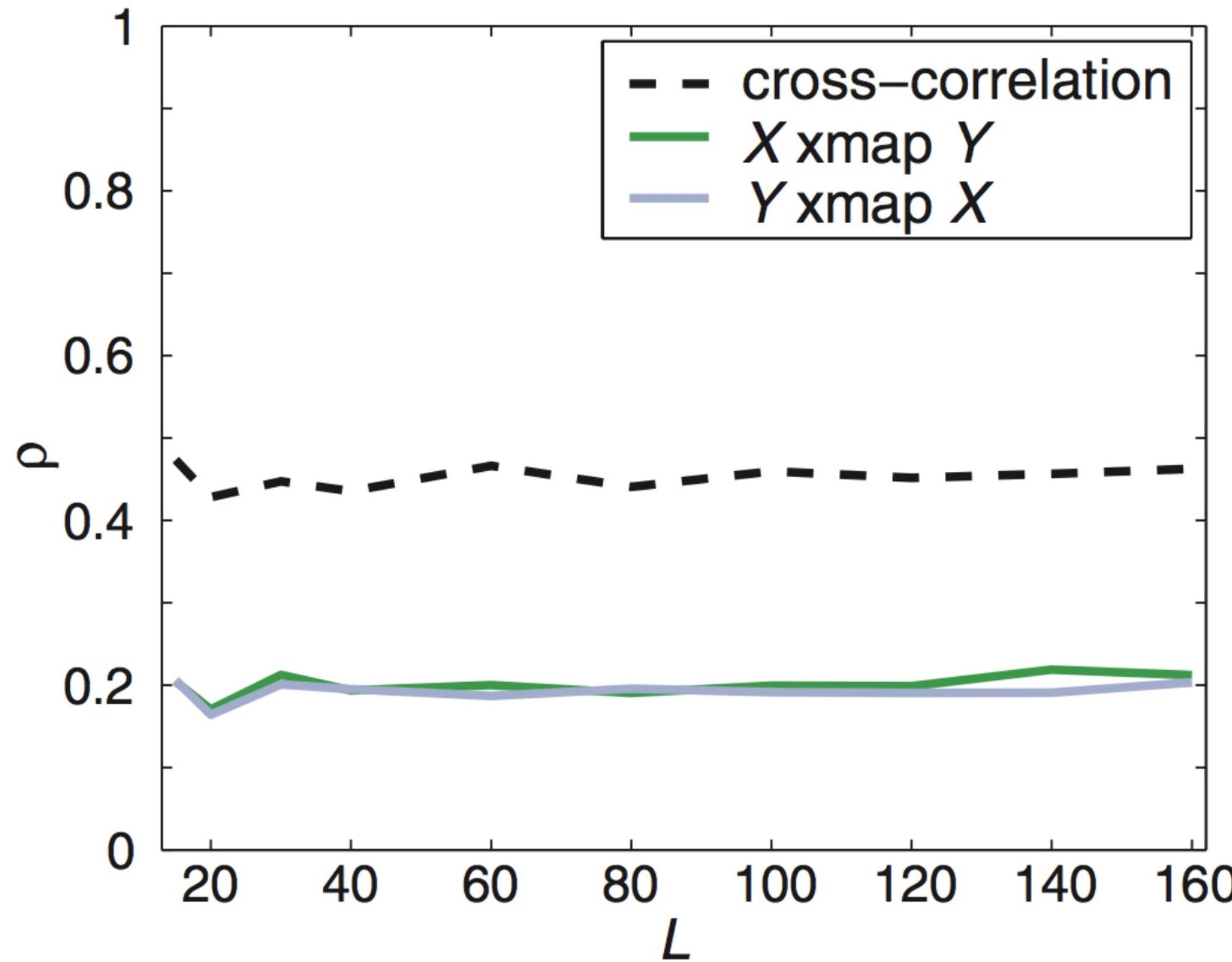


Sugihara et al. 2012

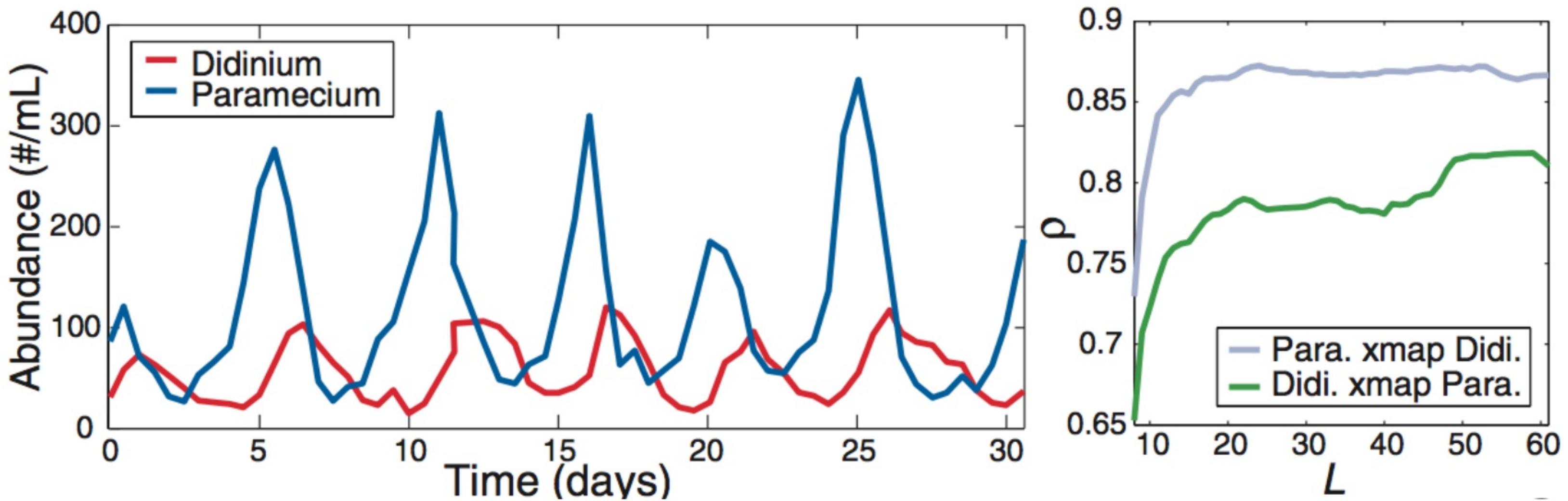
What about non-interacting variables
sharing a driver?

x

X and Y do not interact but share a driver

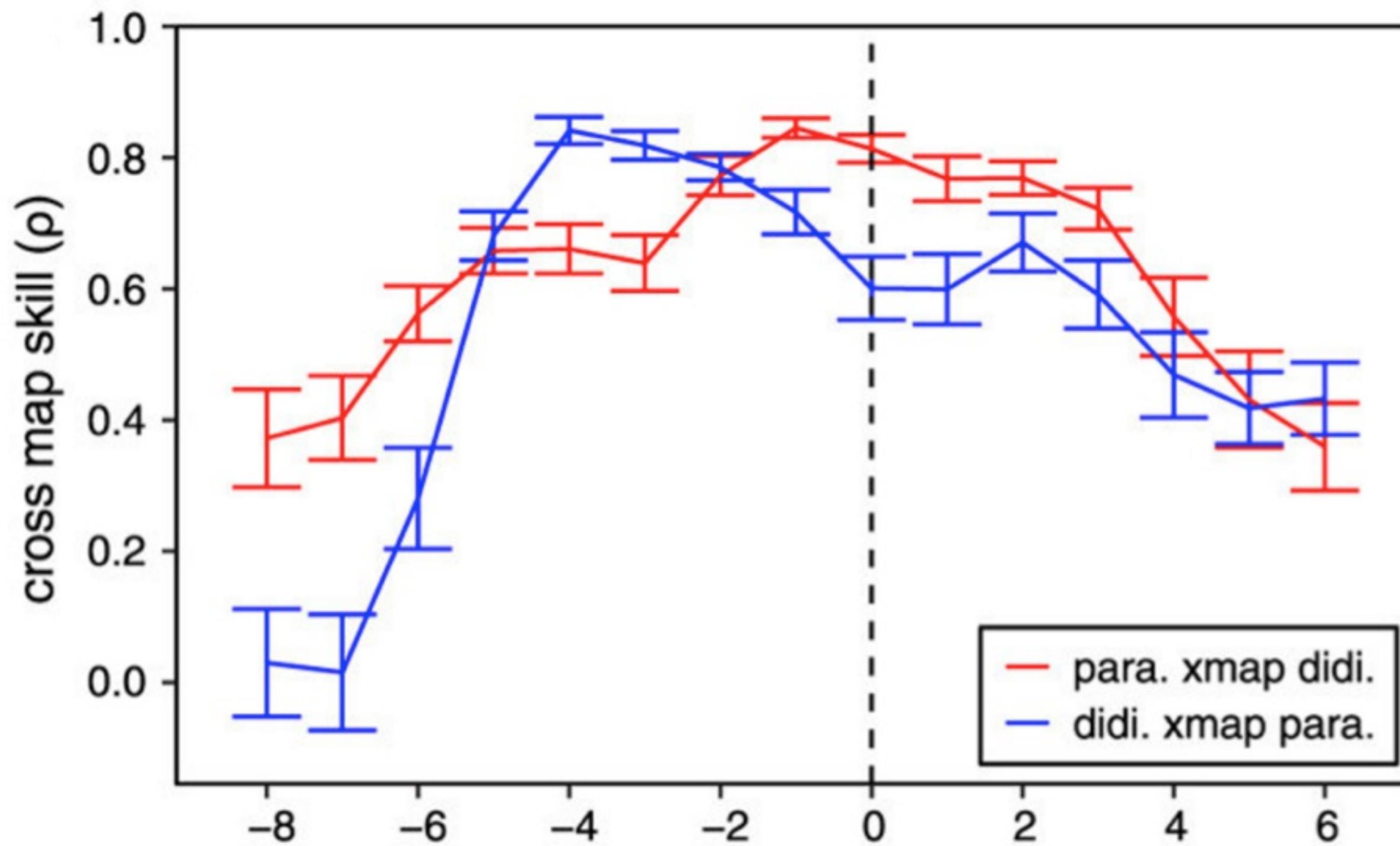


Applied to predator-prey cycles

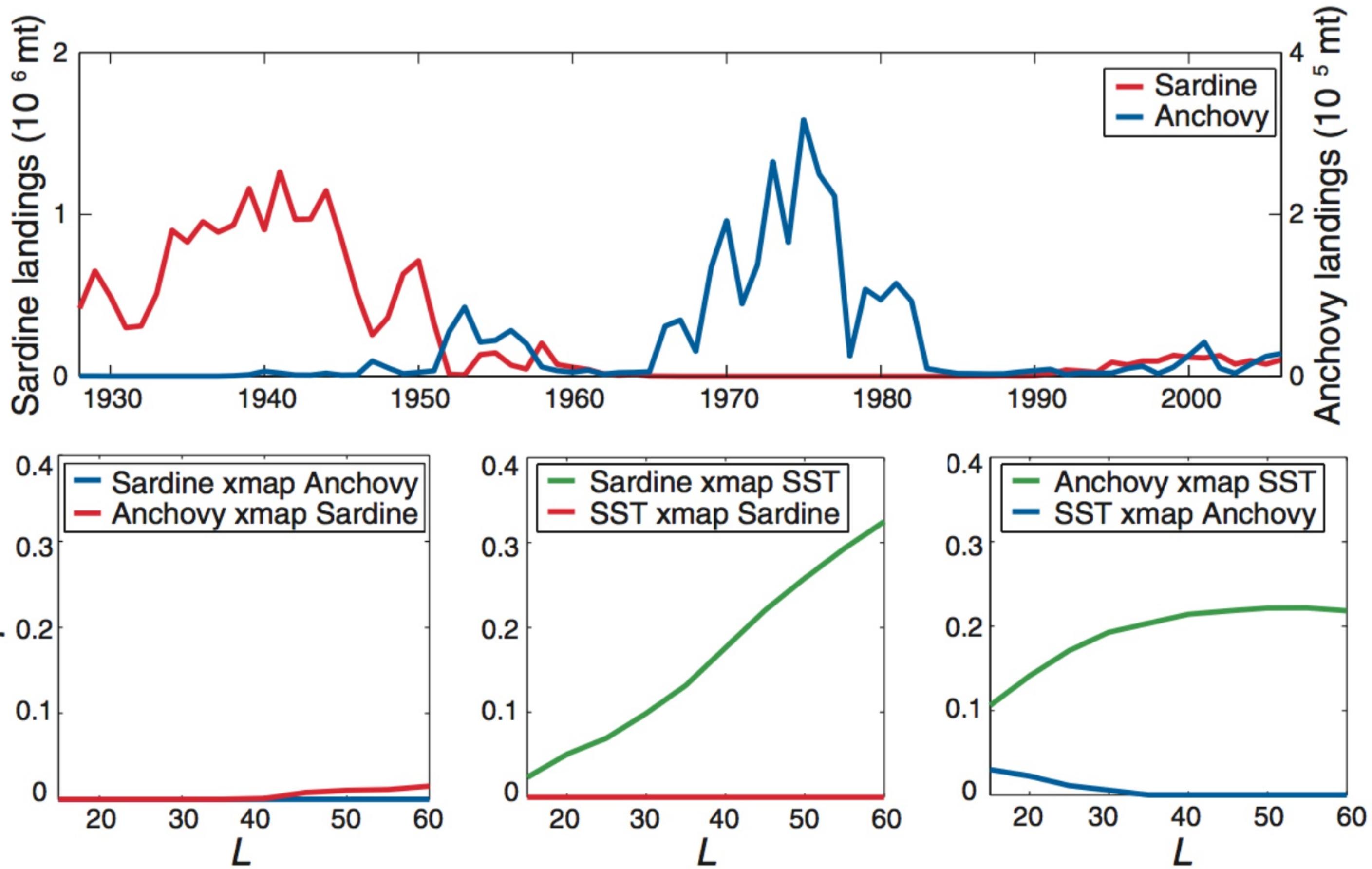


Sugihara et al. 2012

Cross-map lag shows direction

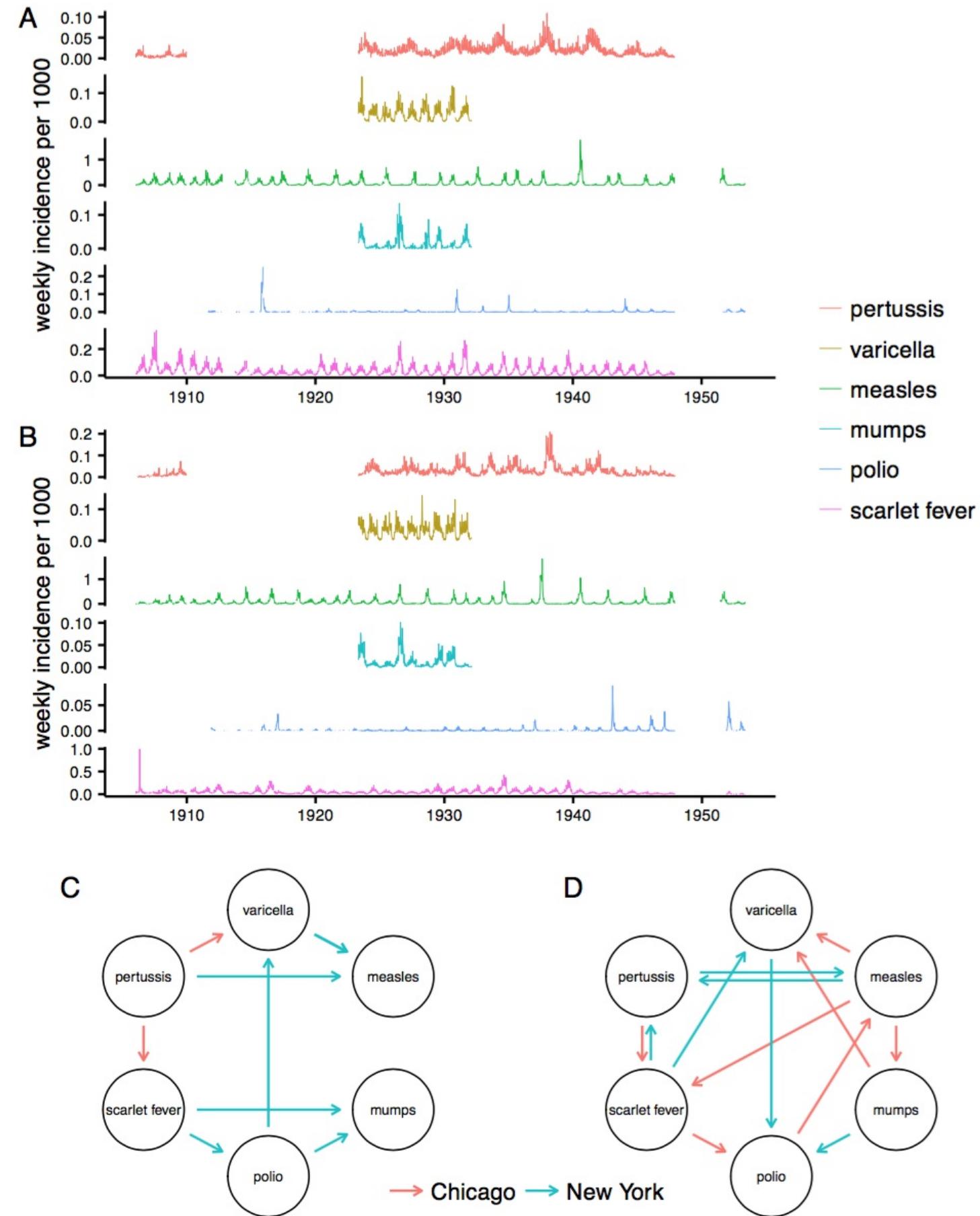


Anchovies, sardines, and SST



What are the assumptions?

Application to real time series



Back to mechanistic models,
then...

Building confidence in your model

Predict something else

Exploit natural and unnatural disturbances

Streptococcus pneumoniae

Carried by 20-80% of young children

Transmitted mostly between healthy carriers

>90 serotypes

Some serotypes seem better at everything

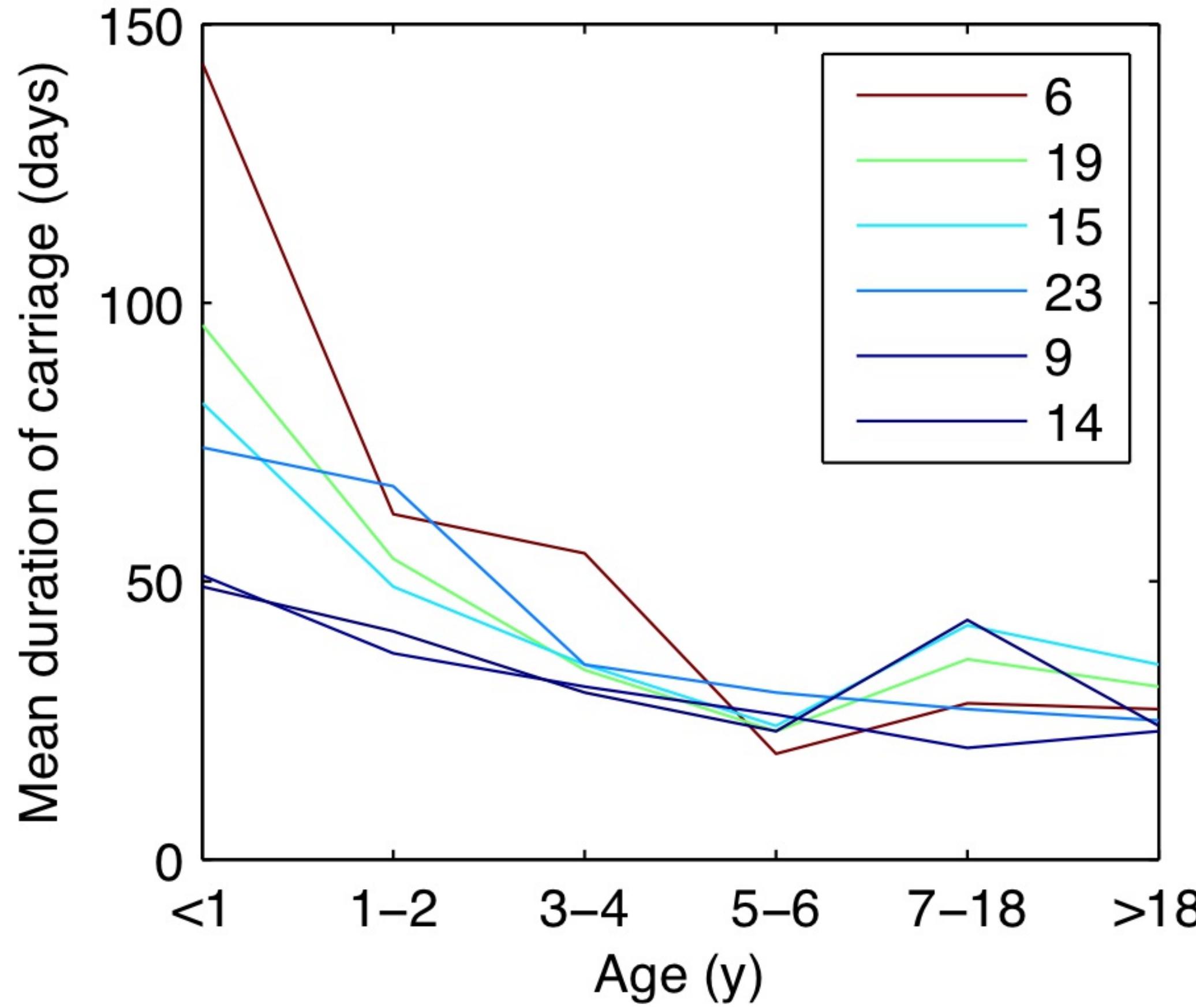
Little evidence for anticapsular immunity

Table 3. Effect of prior colonization with a particular serotype on new acquisition of the same or different serotype in toddlers in Israel. GEE analysis. Odds ratio with 95% confidence interval. Adjusted for prior exposure to other types, age, age at study entry, swab + at prior visit.

		Previous Colonization with type:							
		6A	6B	14	15	19A	19F	23A	23F
New acquisition of type:	6A	0.48*	0.55	1.02	1.27	0.63	0.62	0.87	0.75
	6B	0.76 (0.31-1.90)	0.87 (0.26-2.87)	0.79 (0.24-2.62)	2.16 (0.88-5.31)	2.09 (0.69-6.36)	1.38 (0.58-3.28)	1.46 (0.50-4.28)	1.95 (0.44-2.60)
	14	0.96 (0.46-2.03)	1.02 (0.39-2.67)	0.08* (0.01-0.66)	0.37 (0.43-1.93)	0.72 (0.24-2.17)	1.01 (0.49-2.08)	0.69 (0.23-2.10)	0.82 (0.40-1.68)
	15	1.14 (0.71-1.84)	1.24 (0.70-2.20)	1.15 (0.67-1.99)	1.07 (0.64-1.79)	1.42 (0.80-2.50)	1.09 (0.68-1.73)	2.07* (1.24-3.46)	1.16 (0.72-1.85)
	19A	1.32 (0.58-3.03)	1.95 (0.74-5.15)	2.24 (0.99-5.11)	0.53 (0.22-1.26)	0.58 (0.15-2.21)	1.32 (0.60-2.93)	1.28 (0.49-3.37)	1.15 (0.52-2.56)
	19F	1.43 (0.85-2.41)	0.87 (0.44-1.73)	0.65 (0.34-1.27)	1.02 (0.60-1.72)	0.80 (0.37-1.74)	0.90 (0.50-1.61)	1.16 (0.60-2.26)	0.86 (0.52-1.42)
	23A	1.37 (0.64-2.95)	1.57 (0.69-3.56)	1.22 (0.54-2.75)	0.95 (0.45-2.02)	1.73 (0.74-4.03)	1.65 (0.81-3.36)	0.51 (0.14-1.84)	1.38 (0.67-2.86)
	23F	0.63 (0.36-1.09)	1.57 (0.83-2.95)	1.41 (0.78-2.54)	1.00 (0.56-1.79)	1.17 (0.60-2.30)	0.90 (0.52-1.56)	0.45 (0.17-1.17)	0.47* (0.26-0.86)

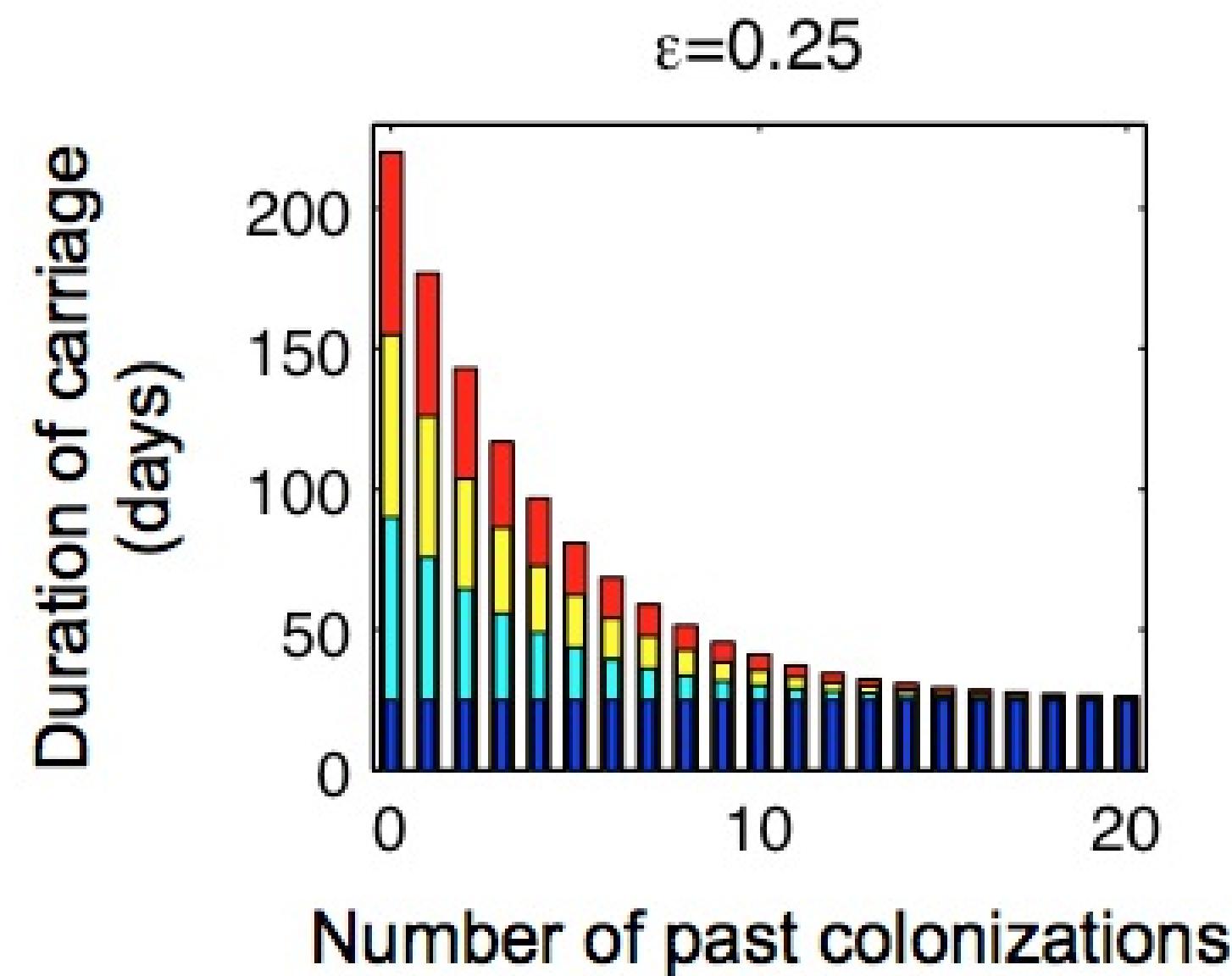
NOTE: * p<0.05.

Non-serotype-specific immunity



Hogberg et al. 2007

Fitted duration of carriage



Cobey and Lipsitch 2012

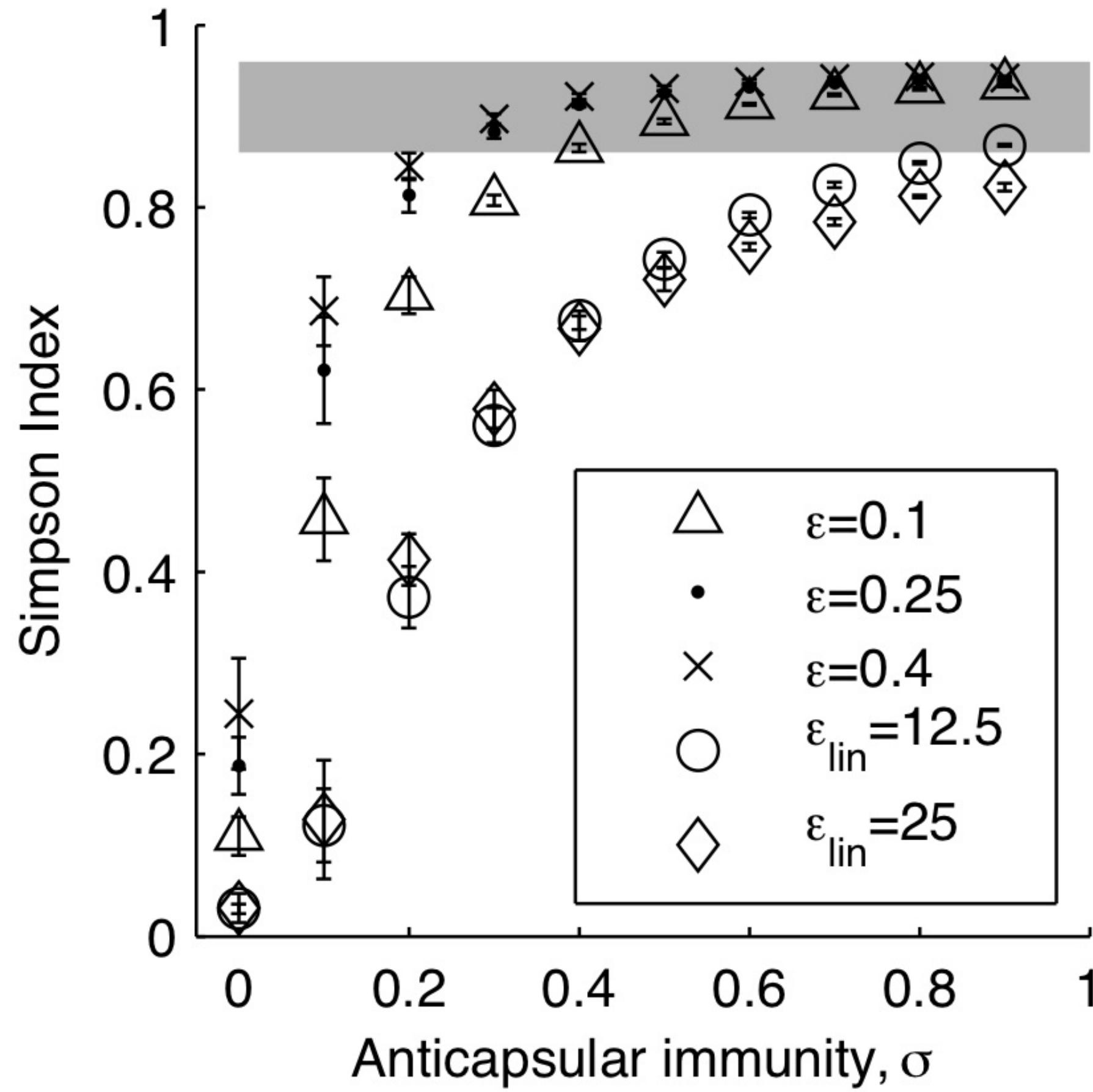
Approach

For each value of serotype-specific immunity, σ

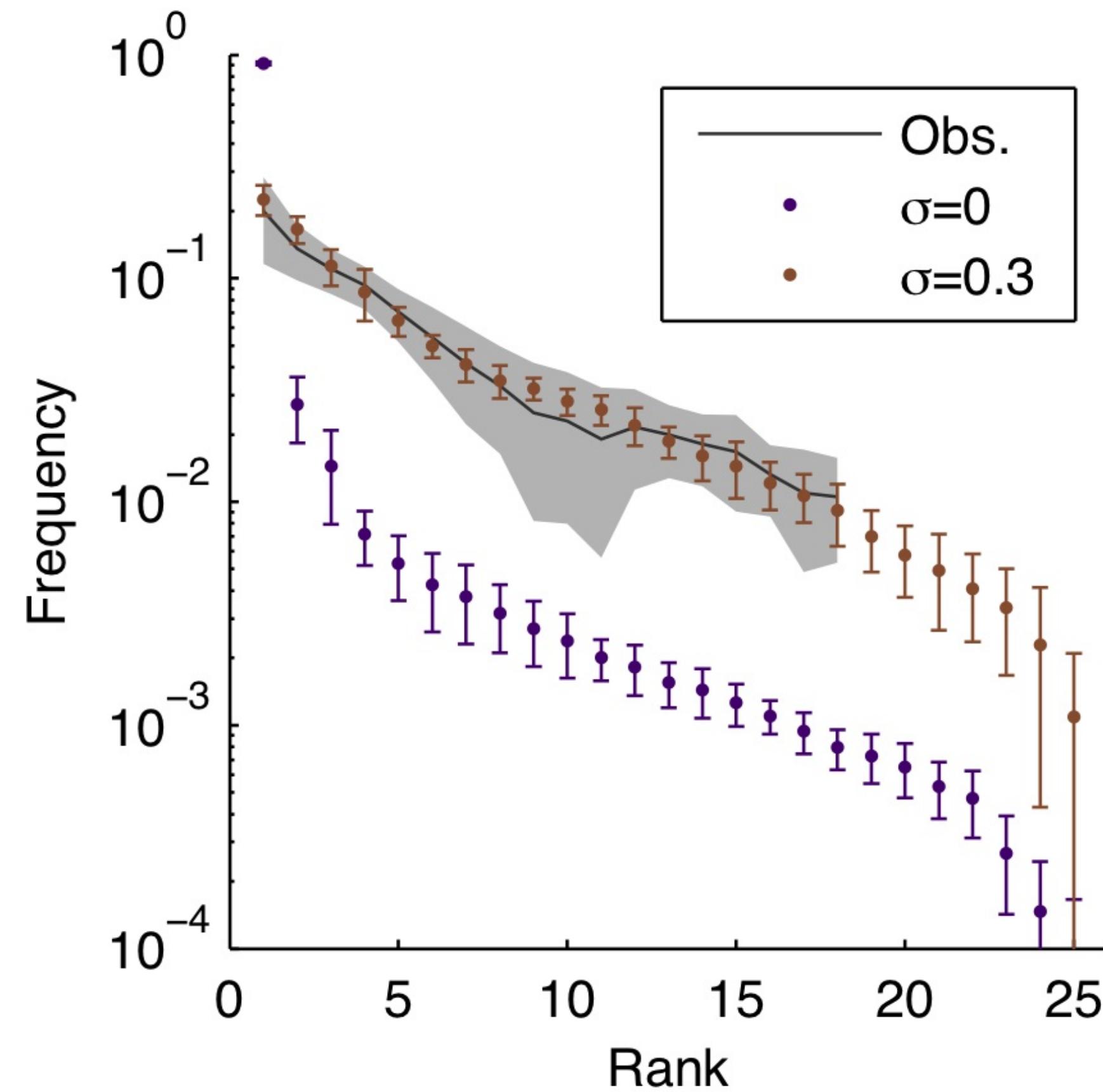
Fit the transmission rate to obtain 40% prevalence in kids

(Later, sensitivity analysis on fixed parameters)

Model reproduces diversity



...including rank-frequency



Other matched patterns

Increase in serotype diversity with age

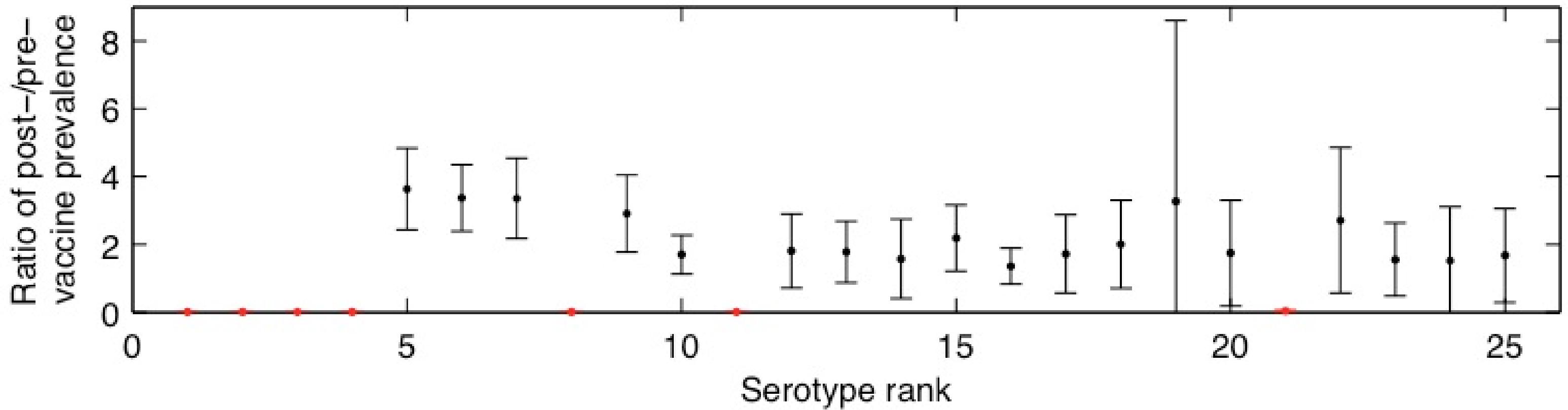
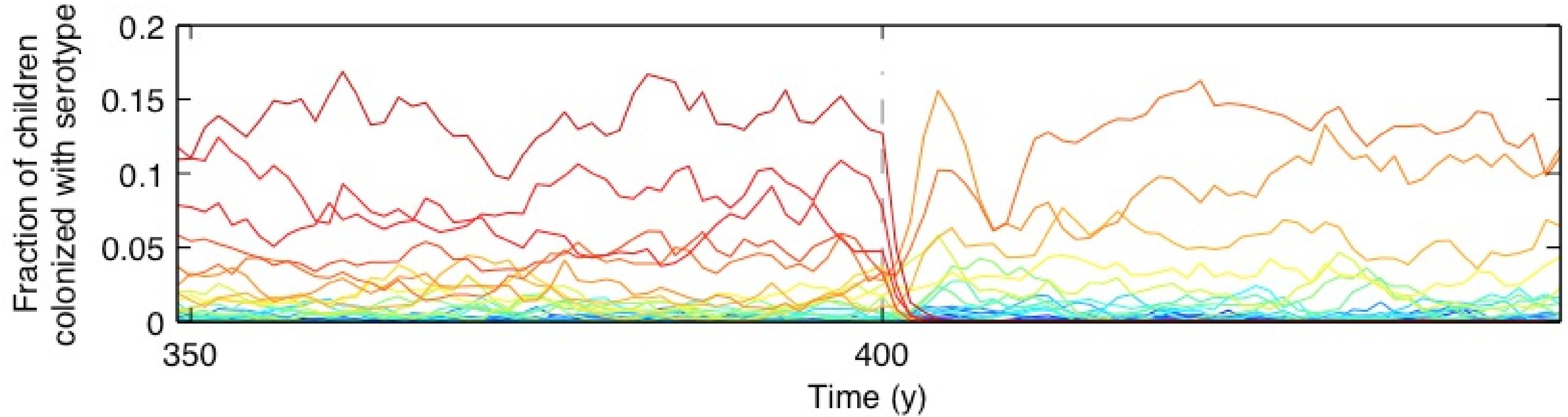
Stable rank order

Decrease in carriage duration with age

Frequency of co-colonizations

Epidemics of rarer serotypes

Vaccinations as natural experiments



Cobey and Lipsitch 2012