NOAA FISHERIES

Northwest Fisheries Science Center

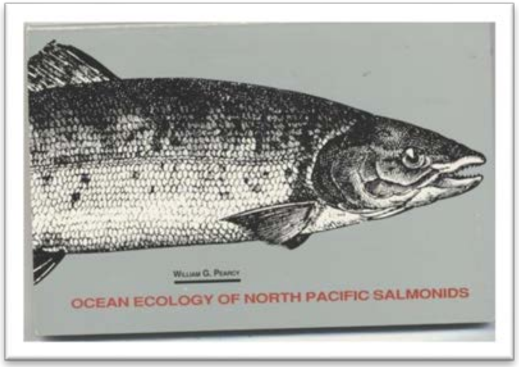
Inferring and forecasting community responses to climate drivers *using multivariate time-series analysis*

Elizabeth Holmes

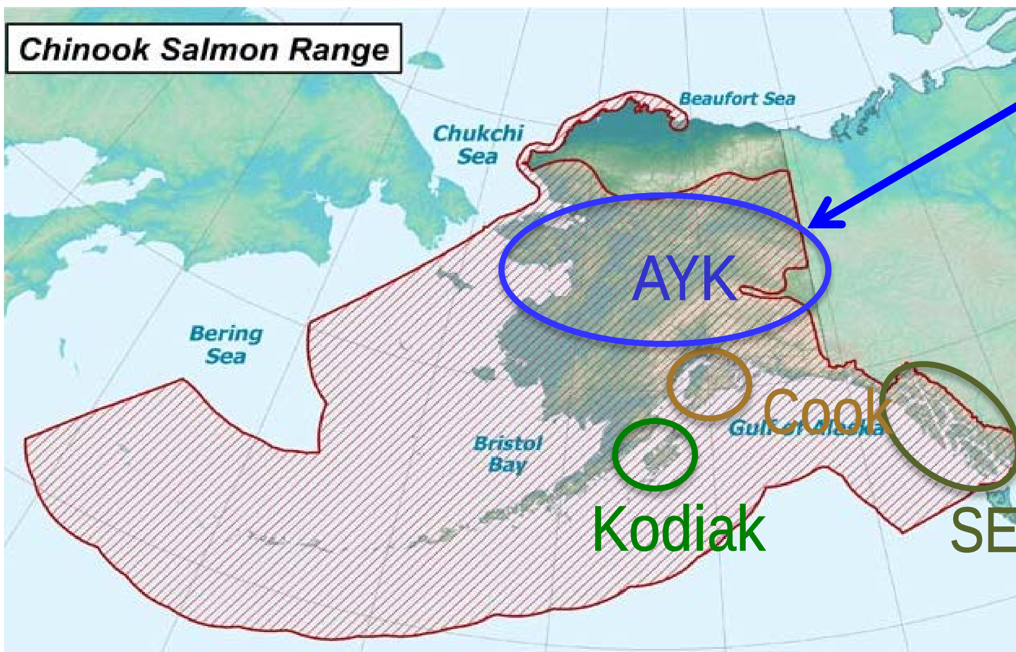
NOAA Fisheries/University of Washington

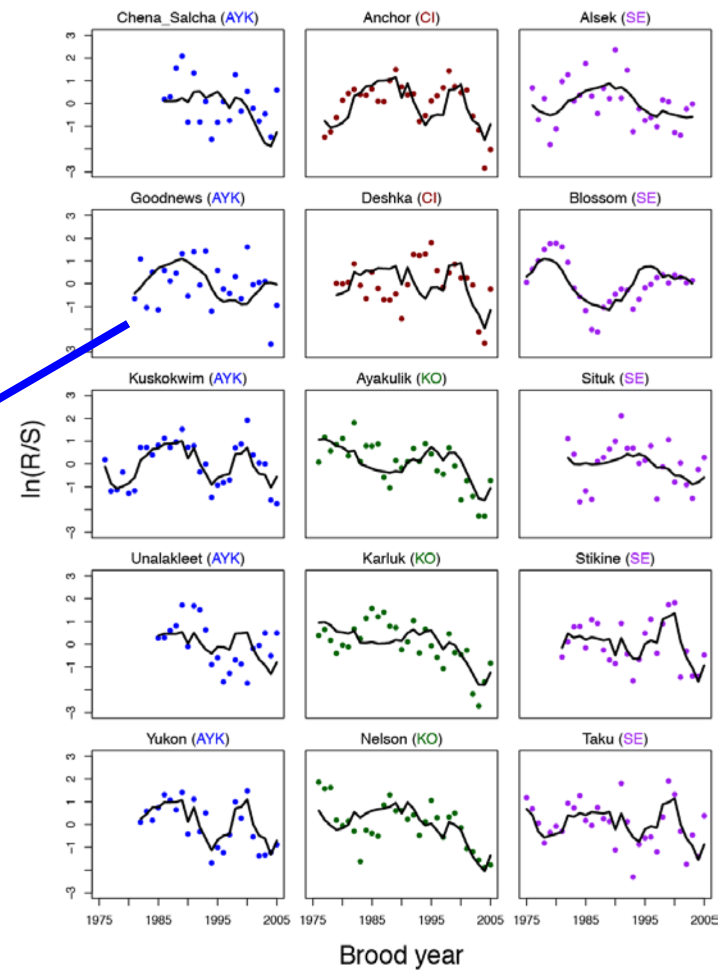
faculty.washington.edu/eeholmes

Multivariate ecological time series



example from work by Mark Scheuerell, NWFSC, NOAA





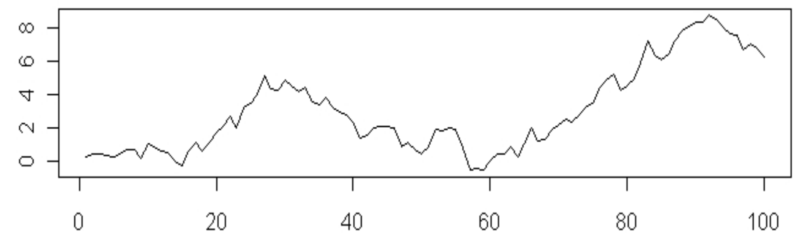
Some problems we have with our data

Multivariate time-series data with

- Lots of gaps (missing data)
- (Unknown) observation error
- Complex (unknown) relationships between observation and underlying process trajectory
- Non-ideal covariate data --- instrumentation changed, multiple time series

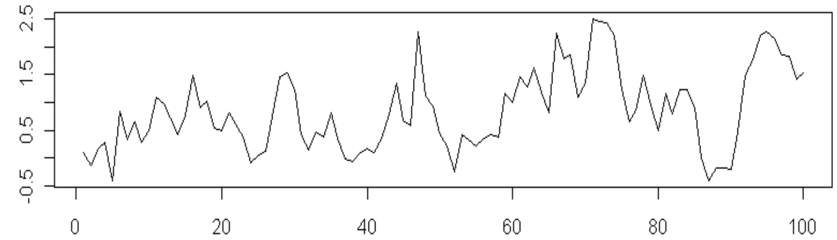
The questions we ask often involve random walks

SOME UNDERLYING "HIDDEN" AUTOREGRESSIVE PROCESS



Random walk

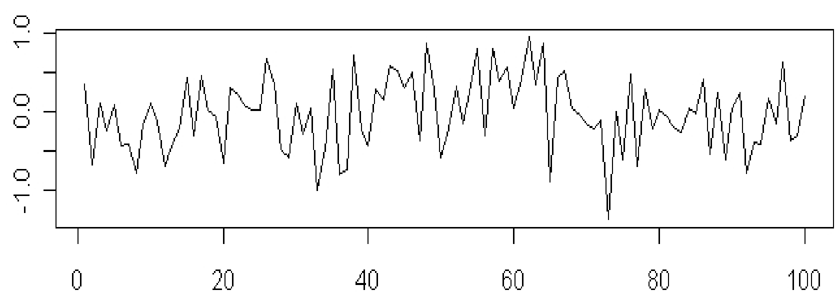
$x(t)=x(t-1)+e(t)$ +u



Mean-reverting random walk

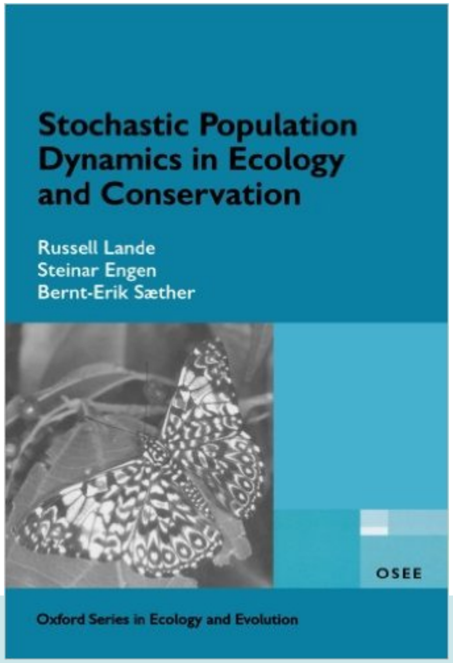
$x(t)=bx(t-1)+e(t)$ +u

+ OBSERVATION PROCESS

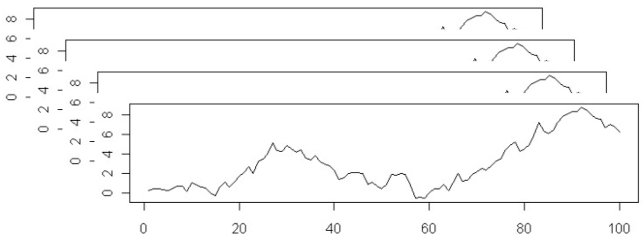


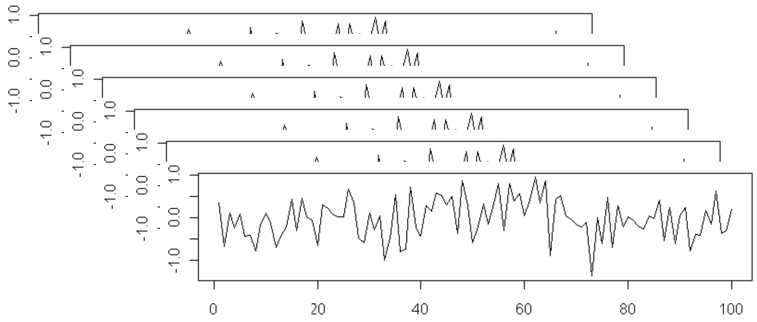
White or autocorrelated noise

$x(t)=e(t)$ +u



Our problems are multivariate





hidden random walks that we want to estimate or 'understand'

observations

Multivariate autoregressive "random walk"

Multivariate with noise

Multivariate auto-regressive state-space (MARSS) models

Process

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}_t = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}_{t-1} + \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}_t, \quad \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}_t \sim MVN\left(\begin{bmatrix} u_1 \\ u_2 \end{bmatrix}, \begin{bmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{bmatrix} \right)$$

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}_t = \begin{bmatrix} z_{11} & z_{12} \\ z_{21} & z_{22} \\ z_{31} & z_{32} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}_t + \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}_t, \quad \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}_t \sim MVN\left(\begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}, \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \right)$$

Observation

$$\mathbf{x}_t = \mathbf{B}_t\mathbf{x}_{t-1} + \mathbf{u}_t + \mathbf{C}_t\mathbf{c}_t + \mathbf{w}_t, \mathbf{w}_t \sim \mathrm{MVN}(0, \mathbf{Q}_t)$$

$$\mathbf{y}_t = \mathbf{Z}_t\mathbf{x}_t + \mathbf{a}_t + \mathbf{D}_t\mathbf{d}_t + \mathbf{v}_t, \mathbf{v}_t \sim \mathrm{MVN}(0, \mathbf{R}_t)$$

A talk in two parts

Part I: Development of a new expectation-maximization algorithm for fitting *constrained* MARSS models

$$
\begin{aligned}
\log \mathbf{L}(\boldsymbol{y}, \boldsymbol{x}; \Theta) = &-\sum_{1}^{T} \frac{1}{2}(\boldsymbol{y}_t - \mathbf{Z}\boldsymbol{x}_t - \mathbf{a})^\top \mathbf{R}^{-1}(\boldsymbol{y}_t - \mathbf{Z}\boldsymbol{x}_t - \mathbf{a}) - \sum_{1}^{T} \frac{1}{2} \log |\mathbf{R}| \\
&-\sum_{1}^{T} \frac{1}{2}(\boldsymbol{x}_t - \mathbf{B}\boldsymbol{x}_{t-1} - \mathbf{u})^\top \mathbf{Q}^{-1}(\boldsymbol{x}_t - \mathbf{B}\boldsymbol{x}_{t-1} - \mathbf{u}) - \sum_{1}^{T} \frac{1}{2} \log |\mathbf{Q}| \\
&-\frac{1}{2}(\boldsymbol{x}_0 - \xi)^\top \Lambda^{-1}(\boldsymbol{x}_0 - \xi) - \frac{1}{2} \log |\Lambda| - \frac{n}{2} \log 2\pi
\end{aligned}
$$

Part II: Using MARSS models to study the effect of climate variation versus species interactions on native fishes in a desert river

Parameter estimation for constrained MARSS models

x are hidden states

c and d are 'inputs'

$$\mathbf{x}_t = \boxed{\mathbf{B}_t}\mathbf{x}_{t-1} + \boxed{\mathbf{u}_t} + \boxed{\mathbf{C}_t}\mathbf{c}_t + \mathbf{w}_t, \mathbf{w}_t \sim \text{MVN}(0, \boxed{\mathbf{Q}_t})$$

$$\mathbf{y}_t = \boxed{\mathbf{Z}_t}\mathbf{x}_t + \boxed{\mathbf{a}_t} + \boxed{\mathbf{D}_t}\mathbf{d}_t + \mathbf{v}_t, \mathbf{v}_t \sim \text{MVN}(0, \boxed{\mathbf{R}_t})$$

y are data (w or w/o missing values)

$$\boxed{\begin{bmatrix} \psi & 0 & 3+\chi \\ 0 & \psi & \psi+2\chi \\ 2 & \theta & 0 \end{bmatrix}} \quad \boxed{\begin{bmatrix} \psi \\ \psi \\ 2 \end{bmatrix}} \quad \boxed{\begin{bmatrix} \psi & \chi & 0 \\ \chi & \psi & 0 \\ 0 & 0 & \theta \end{bmatrix}}$$

Linear constraints on matrix elements

$$\begin{bmatrix} \psi & 0 & 3+\chi \\ 0 & \psi & \psi+2\chi \\ 2 & \theta & 0 \end{bmatrix} \qquad \begin{bmatrix} \psi & \chi & 0 \\ \chi & \psi & 0 \\ 0 & 0 & \theta \end{bmatrix}$$

$$c + \alpha_1\beta_1 + \alpha_2\beta_2 + \dots$$

Models that can be written with a constrained MARSS form

- **Multivariate auto-regressive process**
  - stochastic exponential growth, random walk movement
  - Gompertz density-dependent growth
- **Multi-site data**
  - Combining multiple time series with collection methods that have changed over time
- **Estimation of species interaction matrices ("MAR" modeling)**
- **Linear dynamical models**

  a multivariate regression model where the coefs are random walks
- **Dynamic factor analysis**
  - similar to PCA but for time-series data
  - describe a large number (100s) of time series with a smaller set (1-5) of 'trend' time series
- **AR-p models**
- **Multivariate-response linear regression**

Related to 'shrinkage estimation'

Statistical Applications in Genetics and Molecular Bio

Volume 4, Issue 1 2005

A Shrinkage Approach to L
Covariance Matrix Estim
Implications for Functional

Ju

Biometrika (2010), **0**, 0, *pp.* 1–24

© 2008 Biometrika Trust Advance Access publication on DAY MONTH YEAR

Printed in Great Britain

Sparse Estimation of a Covariance Matrix

BY JACOB BIEN AND ROBERT TIBSHIRANI

Departments of Statistics and Health Research & Policy, Stanford University, Sequoia Hall,

4305-4065, U.S.A.

Multi-Target Shrinkage Estimation for Covariance Matrices

Tomer Lancewicki, *Member, IEEE*, and Mayer Aladjem

***Abstract*—Covariance matrix estimation is problematic when the number of samples is relatively small c number of variables. One way to tackle this p the use of shrinkage estimators that offer a co the sample covariance matrix and a well-c (also known as the "target") with the aim mean-squared error (MSE). The use of only o shrinkage estimators' flexibility when minim this paper, we propose a multi-target shrinkage for covariance matrices that exploits the l**

density with a mean $\mathbf{m}$ and covariance matrix $\boldsymbol{\Sigma} = \{\sigma_{ij}\}$.

It can be hard to impose these constraints---in a general way

MARSS models are widely used in economics, finance and engineering

though perhaps with different names, e.g. VARSS, DLM, or simply state-space

If MARSS models have such a history, why not just use existing algorithms?

Existing methods could deal with

- Some types of MARSS models with missing data
- observation error
- time-varying parameters

B[illegible]

- Parameter estimation based on Newton methods can struggle with some MARSS models and some data. → EM methods
- Existing EM methods did not [illegible] ints
- Existing EM methods did not partially deterministic systems (variances of 0)

Goal

- Develop a general robust algorithm for constrained MARSS models based on Expectation-Maximization algorithms

Deriving algorithms is fun and captivating. I didn't really need any more justification than that for putting years of effort into this!

Holmes, E. E. 2010, 2012. Derivation of the EM algorithm for constrained and unconstrained multivariate autoregressive state-space (MARSS) models.

Finding MLE parameters for *constrained* MARSS models:

Joint likelihood of y(data) and x(hidden states) given a set of parameters

$$
\begin{aligned}
&\mathrm{E}_{\mathbf{XY}}[\log \mathbf{L}(\boldsymbol{Y},\boldsymbol{X};\Theta);\boldsymbol{Y}(1)=\boldsymbol{y}(1),\Theta_j] = \Psi = \\
&\quad -\frac{1}{2}\sum_{1}^{T}\Big(\mathrm{E}[\boldsymbol{Y}_t^\top\mathbf{R}^{-1}\boldsymbol{Y}_t] - \mathrm{E}[\boldsymbol{Y}_t^\top\mathbf{R}^{-1}\mathbf{Z}\boldsymbol{X}_t] - \mathrm{E}[(\mathbf{Z}\boldsymbol{X}_t)^\top\mathbf{R}^{-1}\boldsymbol{Y}_t] - \mathrm{E}[\mathbf{a}^\top\mathbf{R}^{-1}\boldsymbol{Y}_t] - \mathrm{E}[\boldsymbol{Y}_t^\top\mathbf{R}^{-1}\mathbf{a}] \\
&\quad + \mathrm{E}[(\mathbf{Z}\boldsymbol{X}_t)^\top\mathbf{R}^{-1}\mathbf{Z}\boldsymbol{X}_t] + \mathrm{E}[\mathbf{a}^\top\mathbf{R}^{-1}\mathbf{Z}\boldsymbol{X}_t] + \mathrm{E}[(\mathbf{Z}\boldsymbol{X}_t)^\top\mathbf{R}^{-1}\mathbf{a}] + \mathrm{E}[\mathbf{a}^\top\mathbf{R}^{-1}\mathbf{a}]\Big) - \frac{T}{2}\log|\mathbf{R}| \\
&\quad -\frac{1}{2}\sum_{1}^{T}\Big(\mathrm{E}[\boldsymbol{X}_t^\top\mathbf{Q}^{-1}\boldsymbol{X}_t] - \mathrm{E}[\boldsymbol{X}_t^\top\mathbf{Q}^{-1}\mathbf{B}\boldsymbol{X}_{t-1}] - \mathrm{E}[(\mathbf{B}\boldsymbol{X}_{t-1})^\top\mathbf{Q}^{-1}\boldsymbol{X}_t] \\
&\quad - \mathrm{E}[\mathbf{u}^\top\mathbf{Q}^{-1}\boldsymbol{X}_t] - \mathrm{E}[\boldsymbol{X}_t^\top\mathbf{Q}^{-1}\mathbf{u}] + \mathrm{E}[(\mathbf{B}\boldsymbol{X}_{t-1})^\top\mathbf{Q}^{-1}\mathbf{B}\boldsymbol{X}_{t-1}] \\
&\quad + \mathrm{E}[\mathbf{u}^\top\mathbf{Q}^{-1}\mathbf{B}\boldsymbol{X}_{t-1}] + \mathrm{E}[(\mathbf{B}\boldsymbol{X}_{t-1})^\top\mathbf{Q}^{-1}\mathbf{u}] + \mathbf{u}^\top\mathbf{Q}^{-1}\mathbf{u}\Big) - \frac{T}{2}\log|\mathbf{Q}| \\
&\quad -\frac{1}{2}\Big(\mathrm{E}[\boldsymbol{X}_0^\top\mathbf{V}_0^{-1}\boldsymbol{X}_0] - \mathrm{E}[\boldsymbol{\xi}^\top\boldsymbol{\Lambda}^{-1}\boldsymbol{X}_0] - \mathrm{E}[\boldsymbol{X}_0^\top\boldsymbol{\Lambda}^{-1}\boldsymbol{\xi}] + \boldsymbol{\xi}^\top\boldsymbol{\Lambda}^{-1}\boldsymbol{\xi}\Big) - \frac{1}{2}\log|\boldsymbol{\Lambda}| - \frac{n}{2}\log\pi
\end{aligned}
$$

log L(y,x;Θ) = f(y,x,Θ)

Idea is to find the Θ that maximizes the **log L(y;Θ)**

Newton-based method

- The Kalman filter will give you the marginal likelihood

- Write a function to compute the marginal likelihood **L(y; Q)**, or there are a number of R packages that do this. I use KFAS.

- Maximize that using a Newton-type algorithm. You could use optim().

- You can do this in a few lines of code.

Works great for lots of problems. But for many big multivariate problems it doesn't work so great.

How do you deal with constrained variance-covariance matrices?

Expectation-Maximization algorithm

Joint likelihood of y and x is log $L(y,x;\Theta) = f(y,x,\Theta)$

The EM algorithm maximizes the expected value of the joint likelihood

$$\mathrm{E}_{\mathbf{XY}}[\log \mathbf{L}(\boldsymbol{Y}, \boldsymbol{X}; \Theta); \boldsymbol{Y}(1) = \boldsymbol{y}(1), \Theta_j]$$

1. Start with Θ_1
2. Compute the expectations involving X and Y conditioned on Θ_1 and the data
3. Put those in $\mathrm{E}_{\mathbf{XY}}[\log \mathbf{L}(\boldsymbol{Y}, \boldsymbol{X}; \Theta); \boldsymbol{Y}(1) = \boldsymbol{y}(1), \Theta_j]$
4. Maximize that (3) with respect to Θ to get Θ_2
   - *This is the hard part.*
   - Take the partial derivative of (3) wrt Θ and solve for Θ
5. Repeat until convergence

NOAA FISHERIES

We need to take the partial derivative of the expectation of the LL wrt parameters: u, R, Q, B, a, Z

$$
\begin{aligned}
&\mathrm{E}_{\mathbf{XY}}[\log \mathbf{L}(\boldsymbol{Y},\boldsymbol{X};\Theta);\boldsymbol{Y}(1)=\boldsymbol{y}(1),\Theta_j] = \Psi = \\
&-\frac{1}{2}\sum_1^T\Bigg(\mathrm{E}[\boldsymbol{Y}_t^\top\mathbf{R}^{-1}\boldsymbol{Y}_t] - \mathrm{E}[\boldsymbol{Y}_t^\top\mathbf{R}^{-1}\mathbf{Z}\boldsymbol{X}_t] - \mathrm{E}[(\mathbf{Z}\boldsymbol{X}_t)^\top\mathbf{R}^{-1}\boldsymbol{Y}_t] - \mathrm{E}[\mathbf{a}^\top\mathbf{R}^{-1}\boldsymbol{Y}_t] - \mathrm{E}[\boldsymbol{Y}_t^\top\mathbf{R}^{-1}\mathbf{a}] \\
&+ \mathrm{E}[(\mathbf{Z}\boldsymbol{X}_t)^\top\mathbf{R}^{-1}\mathbf{Z}\boldsymbol{X}_t] + \mathrm{E}[\mathbf{a}^\top\mathbf{R}^{-1}\mathbf{Z}\boldsymbol{X}_t] + \mathrm{E}[(\mathbf{Z}\boldsymbol{X}_t)^\top\mathbf{R}^{-1}\mathbf{a}] + \mathrm{E}[\mathbf{a}^\top\mathbf{R}^{-1}\mathbf{a}]\Bigg) - \frac{T}{2}\log|\mathbf{R}| \\
&-\frac{1}{2}\sum_1^T\Bigg(\mathrm{E}[\boldsymbol{X}_t^\top\mathbf{Q}^{-1}\boldsymbol{X}_t] - \mathrm{E}[\boldsymbol{X}_t^\top\mathbf{Q}^{-1}\mathbf{B}\boldsymbol{X}_{t-1}] - \mathrm{E}[(\mathbf{B}\boldsymbol{X}_{t-1})^\top\mathbf{Q}^{-1}\boldsymbol{X}_t] \\
&- \mathrm{E}[\mathbf{u}^\top\mathbf{Q}^{-1}\boldsymbol{X}_t] - \mathrm{E}[\boldsymbol{X}_t^\top\mathbf{Q}^{-1}\mathbf{u}] + \mathrm{E}[(\mathbf{B}\boldsymbol{X}_{t-1})^\top\mathbf{Q}^{-1}\mathbf{B}\boldsymbol{X}_{t-1}] \\
&+ \mathrm{E}[\mathbf{u}^\top\mathbf{Q}^{-1}\mathbf{B}\boldsymbol{X}_{t-1}] + \mathrm{E}[(\mathbf{B}\boldsymbol{X}_{t-1})^\top\mathbf{Q}^{-1}\mathbf{u}] + \mathbf{u}^\top\mathbf{Q}^{-1}\mathbf{u}\Bigg) - \frac{T}{2}\log|\mathbf{Q}| \\
&-\frac{1}{2}\Bigg(\mathrm{E}[\boldsymbol{X}_0^\top\mathbf{V}_0^{-1}\boldsymbol{X}_0] - \mathrm{E}[\boldsymbol{\xi}^\top\boldsymbol{\Lambda}^{-1}\boldsymbol{X}_0] - \mathrm{E}[\boldsymbol{X}_0^\top\boldsymbol{\Lambda}^{-1}\boldsymbol{\xi}] + \boldsymbol{\xi}^\top\boldsymbol{\Lambda}^{-1}\boldsymbol{\xi}\Bigg) - \frac{1}{2}\log|\boldsymbol{\Lambda}| - \frac{n}{2}\log\pi
\end{aligned}
$$

$$\mathbf{x}_t = \mathbf{B}_t\mathbf{x}_{t-1} + \mathbf{u}_t + \mathbf{C}_t\mathbf{c}_t + \mathbf{w}_t, \mathbf{w}_t \sim \mathrm{MVN}(0, \mathbf{Q}_t)$$

Solve for the u that maximizes the E[LL] ψ

$$\Psi = -\mathbf{a}^\top\mathbf{u} - \mathbf{u}^\top\mathbf{a} + \mathbf{u}^\top\mathbf{u}$$

$$\frac{\partial\Psi}{\partial\mathbf{u}} = -\frac{\partial\mathbf{a}^\top\mathbf{u}}{\partial\mathbf{u}} - \frac{\partial\mathbf{u}^\top\mathbf{a}}{\partial\mathbf{u}} + \frac{\partial\mathbf{u}^\top\mathbf{u}}{\partial\mathbf{u}}$$

$$\frac{\partial\Psi}{\partial\mathbf{u}} = -2\mathbf{a}^\top + 2\mathbf{u}^\top$$

$$\mathbf{u} = \mathbf{a}$$

$$\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \\ u_5 \end{bmatrix}$$

$$\mathbf{u} = \begin{bmatrix} u_1 \\ u_1 \\ u_2 \\ u_2 \\ u_2 \end{bmatrix}$$

$$\mathbf{u} = \begin{bmatrix} 1 \\ 0 \\ u_1 \\ u_1 \\ u_2 \end{bmatrix}$$

$$\mathbf{u} = \begin{bmatrix} u_1 \\ u_1 + u_2 \\ u_2 \\ 3 + u_2 + u_3 \\ 4 \end{bmatrix}$$

$$\mathbf{u} = \begin{bmatrix} u_1 \\ u_1 + u_2 \\ u_2 \\ 3 + u_2 + u_3 \\ 4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 3 \\ 4 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \mathbf{f} + \mathbf{D}\mathfrak{u}$$

$$\Psi = -\text{constant} - \mathbf{a}^\top \mathbf{D}\mathfrak{u} - \mathfrak{u}^\top \mathbf{D}^\top \mathbf{a} + \mathfrak{u}^\top \mathbf{D}^\top \mathbf{D}\mathfrak{u}$$

$$\frac{\partial \Psi}{\partial \mathfrak{u}} = -2\mathbf{a}^\top \mathbf{D} + 2\mathfrak{u}^\top \mathbf{D}^\top \mathbf{D}$$

$$\mathfrak{u}^\top = \mathbf{a}^\top \mathbf{D}(\mathbf{D}^\top \mathbf{D})^{-1}$$

$$\mathbf{x}_t = \mathbf{B}_t\mathbf{x}_{t-1} + \mathbf{u}_t + \mathbf{C}_t\mathbf{c}_t + \mathbf{w}_t, \mathbf{w}_t \sim \mathrm{MVN}(0, \mathbf{Q}_t)$$

Solve for the B that maximizes the E[LL] ψ

$$\Psi = -\mathbf{a}^\top\mathbf{B}\mathbf{b} - \mathbf{b}^\top\mathbf{B}^\top\mathbf{a} + \mathbf{c}^\top\mathbf{B}^\top\mathbf{B}\mathbf{c}$$

$$\frac{\partial\Psi}{\partial\mathbf{B}} = -\frac{\partial\mathbf{a}^\top\mathbf{B}\mathbf{b}}{\partial\mathbf{B}} - \frac{\partial\mathbf{b}^\top\mathbf{B}^\top\mathbf{a}}{\partial\mathbf{B}} + \frac{\partial\mathbf{c}^\top\mathbf{B}^\top\mathbf{B}\mathbf{c}}{\partial\mathbf{B}}$$

$$\frac{\partial\Psi}{\partial\mathbf{B}} = -2\mathbf{b}\mathbf{a}^\top + 2\mathbf{c}\mathbf{c}^\top\mathbf{B}^\top$$

$$\mathbf{B}^\top = (\mathbf{c}\mathbf{c}^\top)^{-1}\mathbf{b}\mathbf{a}^\top$$

$$\mathbf{B} = \begin{bmatrix} b_1 & b_2 \\ b_3 & b_4 \end{bmatrix}$$

$$\mathbf{B} = \begin{bmatrix} b_1 & 0 \\ 0 & b_2 \end{bmatrix}$$

$$\mathbf{B} = \begin{bmatrix} b_1 & b_2 \\ b_3 & b_1 \end{bmatrix}$$

$$\mathbf{B} = \begin{bmatrix} 1 & b_1 \\ 0 & b_1 + b_2 \end{bmatrix}$$

Write the log likelihood in vec form

$$\Psi = -\mathbf{a}^\top \mathbf{B}\mathbf{b} - \mathbf{b}^\top \mathbf{B}^\top \mathbf{a} + \mathbf{c}^\top \mathbf{B}^\top \mathbf{B}\mathbf{c}$$

$$\vec{\mathbf{b}} = \mathrm{vec}(\mathbf{B})$$

$$\Psi = \mathrm{vec}(\Psi) = \mathrm{vec}(-\mathbf{a}^\top \mathbf{B}\mathbf{b} - \mathbf{b}^\top \mathbf{B}^\top \mathbf{a} + \mathbf{c}^\top \mathbf{B}^\top \mathbf{B}\mathbf{c})$$

$$\Psi = -\mathbf{d}^\top \vec{\mathbf{b}} - \vec{\mathbf{b}}^\top \mathbf{d} + \vec{\mathbf{b}}^\top \mathbf{C}\vec{\mathbf{b}}$$

$$\mathbf{d} = (\mathbf{b}^\top \otimes \mathbf{a}^\top)$$

$$\mathbf{C} = (\mathbf{c}^\top \otimes \mathbf{I})^\top (\mathbf{c}^\top \otimes \mathbf{I})$$

$$\mathbf{B} = \begin{bmatrix} 1 & b_1 \\ 0 & b_1 + b_2 \end{bmatrix}$$

$$\vec{\mathbf{b}} = \begin{bmatrix} 1 \\ b_1 \\ 0 \\ b_1 + b_2 \end{bmatrix}$$

This leads to a simple solution for linear constraints within parameter matrices

$$\mathbf{B} = \begin{bmatrix} 1 & b_1 \\ 0 & b_1 + b_2 \end{bmatrix}$$

$$\vec{\mathbf{b}} = \begin{bmatrix} 1 \\ b_1 \\ 0 \\ b_1 + b_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \mathbf{f} + \mathbf{D}\mathfrak{b}$$

$$\partial\Psi\partial\mathfrak{b} = -\text{constant} - \mathbf{d}^\top\mathbf{D}\mathfrak{b} - \mathfrak{b}^\top\mathbf{D}^\top\mathbf{d} + \mathfrak{b}^\top\mathbf{D}^\top\mathbf{C}\mathbf{D}\mathfrak{b}$$

Solve for $\mathfrak{b}$ that maximizes Ψ.

$$\mathfrak{b}^\top = \mathbf{d}^\top\mathbf{D}(\mathbf{D}^\top\mathbf{C}\mathbf{D})^{-1}$$

$$\mathbf{x}_t = \mathbf{B}_t\mathbf{x}_{t-1} + \mathbf{u}_t + \mathbf{C}_t\mathbf{c}_t + \mathbf{w}_t, \mathbf{w}_t \sim \mathrm{MVN}(0, \mathbf{Q}_t)$$

Solving for constrained variances Q (and R) follow the same steps as for B

We don't need the 'trace' trick to pull the Q-1 out, because the vec() does that

$$\mathrm{vec}(\mathbf{a}^\top\mathbf{Q}^{-1}\mathbf{b}) = (\mathbf{b}^\top \otimes \mathbf{a}^\top)\mathrm{vec}(\mathbf{Q}^{-1}) = \mathbf{c}^\top\mathbf{D}\mathbf{q}'$$

The estimate has the familiar form

$$\mathbf{q}^\top = \mathbf{d}^\top\mathbf{D}(\mathbf{D}^\top\mathbf{D})^{-1}$$

Works for all legal forms (that I've come up) of variance-covariance matrices

$$\mathbf{Q} = \begin{bmatrix} q_1 & 0 & 0 & 0 & 0 \\ 0 & f_1 & 0 & 0 & 0 \\ 0 & 0 & q_2 & 0 & 0 \\ 0 & 0 & 0 & f_2 & 0 \\ 0 & 0 & 0 & 0 & q_2 \end{bmatrix}$$

$$\mathbf{Q} = \begin{bmatrix} \alpha & \beta & \beta & \beta \\ \beta & \alpha & \beta & \beta \\ \beta & \beta & \alpha & \beta \\ \beta & \beta & \beta & \alpha \end{bmatrix}$$

$$\mathbf{Q} = \begin{bmatrix} \mathbb{B}_1 & 0 & 0 \\ 0 & \mathbb{B}_2 & 0 \\ 0 & 0 & \mathbb{B}_3 \end{bmatrix}$$

$$\mathbf{Q} = \begin{bmatrix} \mathbb{E}_1 & \mathbb{C}_{1,2} & \mathbb{C}_{1,3} \\ \mathbb{C}_{1,2} & \mathbb{E}_2 & \mathbb{C}_{2,3} \\ \mathbb{C}_{1,3} & \mathbb{C}_{2,3} & \mathbb{E}_3 \end{bmatrix}$$

$$\mathbf{q}^\top = \mathbf{d}^\top \mathbf{D}(\mathbf{D}^\top \mathbf{D})^{-1}$$

$$\mathbf{Q} = \begin{bmatrix} \mathbb{F} & 0 & 0 & 0 \\ 0 & \mathbb{G}_1 & 0 & 0 \\ 0 & 0 & \mathbb{G}_2 & 0 \\ 0 & 0 & 0 & \mathbb{G}_3 \end{bmatrix}$$

Expectation-Maximization algorithm

Advantage?

1) It can make certain types of model fitting problems tractable by being considerably faster and more stable
2) Very insensitive to initial conditions. EM algorithms are very good at getting close to the maximum.
3) Very flexible---within the MARSS structure
4) One algorithm for many different MARSS models.

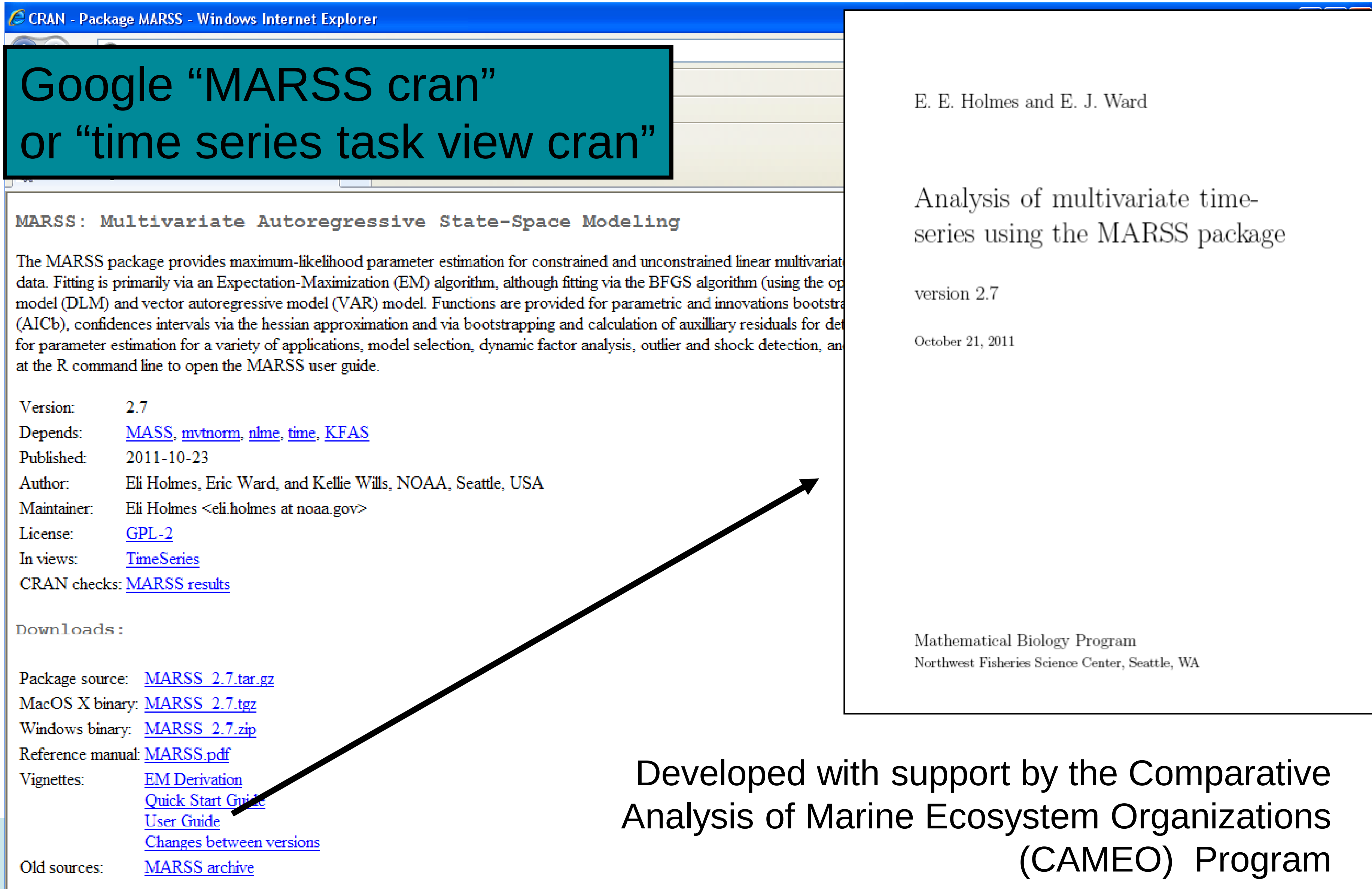
Expectation-Maximization algorithm

Disdvantages?

1) Locked into the MARSS structure
    1) Especially bad when your data have zeros and you want ‘zeros’ model
2) MVN errors
3) Getting a profile of the likelihood is painful (relative to Bayesian approach)
4) E-M algorithms are really slow. If a Newton-esque method works (I use BFGS), it is much faster.

MARSS R package on CRAN

Google "MARSS cran"
or "time series task view cran"


CRAN - Package MARSS - Windows Internet Explorer

MARSS: Multivariate Autoregressive State-Space Modeling

The MARSS package provides maximum-likelihood parameter estimation for constrained and unconstrained linear multivariat
data. Fitting is primarily via an Expectation-Maximization (EM) algorithm, although fitting via the BFGS algorithm (using the op
model (DLM) and vector autoregressive model (VAR) model. Functions are provided for parametric and innovations bootstra
(AICb), confidences intervals via the hessian approximation and via bootstrapping and calculation of auxilliary residuals for det
for parameter estimation for a variety of applications, model selection, dynamic factor analysis, outlier and shock detection, an
at the R command line to open the MARSS user guide.

Version:	2.7
Depends:	MASS, mvtnorm, nlme, time, KFAS
Published:	2011-10-23
Author:	Eli Holmes, Eric Ward, and Kellie Wills, NOAA, Seattle, USA
Maintainer:	Eli Holmes <eli.holmes at noaa.gov>
License:	GPL-2
In views:	TimeSeries
CRAN checks:	MARSS results

Downloads:

Package source:	MARSS_2.7.tar.gz
MacOS X binary:	MARSS_2.7.tgz
Windows binary:	MARSS_2.7.zip
Reference manual:	MARSS.pdf
Vignettes:	EM Derivation Quick Start Guide User Guide Changes between versions
Old sources:	MARSS archive

E. E. Holmes and E. J. Ward

Analysis of multivariate time-series using the MARSS package

version 2.7

October 21, 2011

Mathematical Biology Program
Northwest Fisheries Science Center, Seattle, WA

Developed with support by the Comparative Analysis of Marine Ecosystem Organizations (CAMEO) Program

LIMNOLOGY
and
OCEANOGRAPHY: METHODS

Limnol. Oceanogr.: Methods 10, 2012, 54–64

Assessing marine plankton community structure from long-term monitoring data with multivariate autoregressive (MAR) models: a comparison of fixed station versus spatially distributed sampling data

Lindsay P. Scheef[1], Daniel E. Pendleton[2], Steph*
Scheuerell[2], and David G. Johns[4]
[1]National Center for Ecological Analysis & Synthesi
[2]Northwest Fisheries Science Center, NOAA Fisherie
[3]Channel Islands National Marine Sanctuary, Santa
[4]Sir Alister Hardy Foundation for Ocean Science, Pl

Abstract
We examined how marine plankton int
analysis of time-series, differ based on dat
English Channel) and four similar time-seri
po
th
fo
in
fo
ex
en

Global Change Biology

Global Change Biology (2012) **18**, 2498–2508, doi: 10.1111/j.1365-2486.2012.02702.x

Climate shifts the interaction web of a marine plankton community

TESSA B. FRANCIS*, MARK D. SCHEUERELL*, RICHARD D. BRODEUR†, PHILLIP S. LEVIN*, JAMES J. RUZICKA‡, NICK TOLIMIERI* and WILLIAM T. PETERSON†
**National Oceanic and Atmospheric Administration, Northwest Fisheries Science Center, 2725 Montlake Blvd. E, Seattle, WA 98112, USA, †National Oceanic and Atmospheric Administration, Northwest Fisheries Science Center, Hatfield Marine Science*
ive Institute of Marine Resources Studies, Oregon

ibed, yet accurate predictions about ecosystem
g biotic responses in a food-web context to sup-
ent. Here we conduct time-series analyses with
bundance in the Northern California Current
zooplankton community interactions. Autore-
rm vs. cool ocean climate conditions. Negative
e major warm phase during the time series,
red zooplankton communities. Local environ-
El Niño/Southern Oscillation) were associated
condary environmental correlates of zooplank-
ring the warm phase for upwelling as a covari-
gh simultaneous quantitation of community
oplankton community structure varies with

Global Change Biology (2008) **14**, 1947–1958, doi: 10.1111/j.1365-2486.2008.01616.x

Sixty years of environmental change in the world's largest freshwater lake – Lake Baikal, Siberia

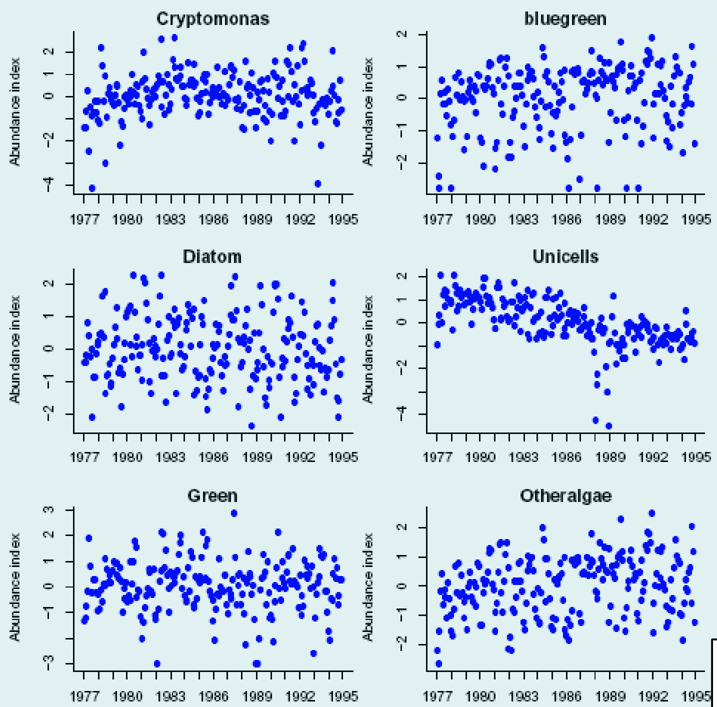
STEPHANIE E. HAMPTON*, LYUBOV R. IZMEST'EVA†, MARIANNE V. MOORE‡, STEPHEN L. KATZ§[1], BRIAN DENNIS¶ and EUGENE A. SILOW†
**National Center for Ecological Analysis & Synthesis, University of California – Santa Barbara, Santa Barbara, CA 93101, USA, †Scientific Research Institute of Biology, Irkutsk State University, Irkutsk 664003, Russia, ‡Department of Biological Sciences, Wellesley College, Wellesley, MA 02481, USA, §Northwest Fisheries Science Center, NOAA Fisheries, Seattle, WA 98112, USA, ¶Departments of Statistics and Fish and Wildlife Resources, University of Idaho, Moscow, ID 83844, USA*

OnlineOpen: This article is available free online at www.blackwell-synergy.com

Abstract

High-resolution data collected over the past 60 years by a single family of Siberian scientists on Lake Baikal reveal significant warming of surface waters and long-term changes in the basal food web of the world's largest, most ancient lake. Attaining depths over 1.6 km, Lake Baikal is the deepest and most voluminous of the world's great lakes. Increases in average water temperature (1.21 °C since 1946), chlorophyll *a* (300% since

Estimating species interaction strengths from time-series data

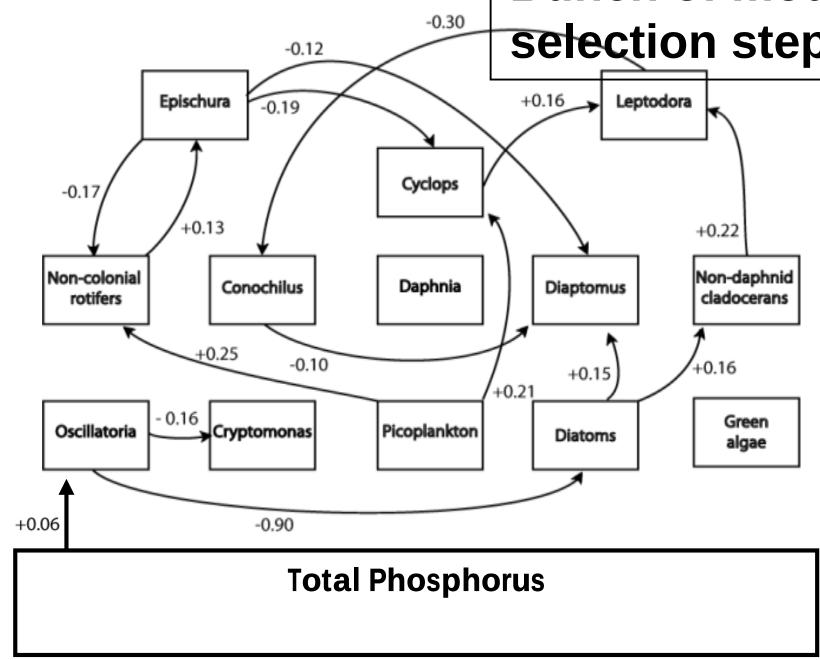


Time series of abundance for the species/groups in the community plus time series of covariates

“MAR”

$$\mathbf{x}_t = \mathbf{B}\mathbf{x}_{t-1} + \mathbf{u} + \mathbf{C}\mathbf{f}_t + \mathbf{w}_t$$

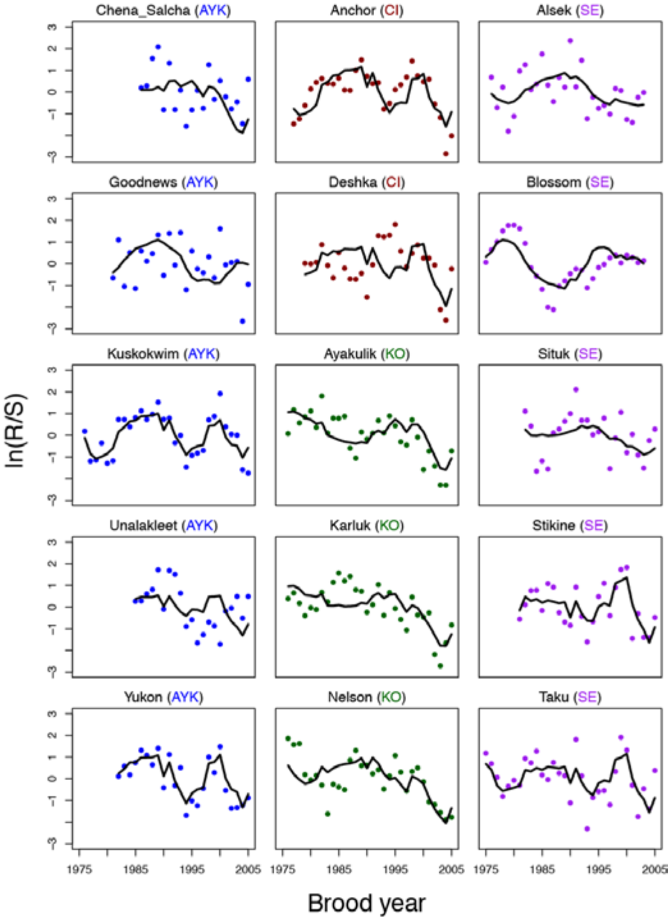
Bunch of model selection steps



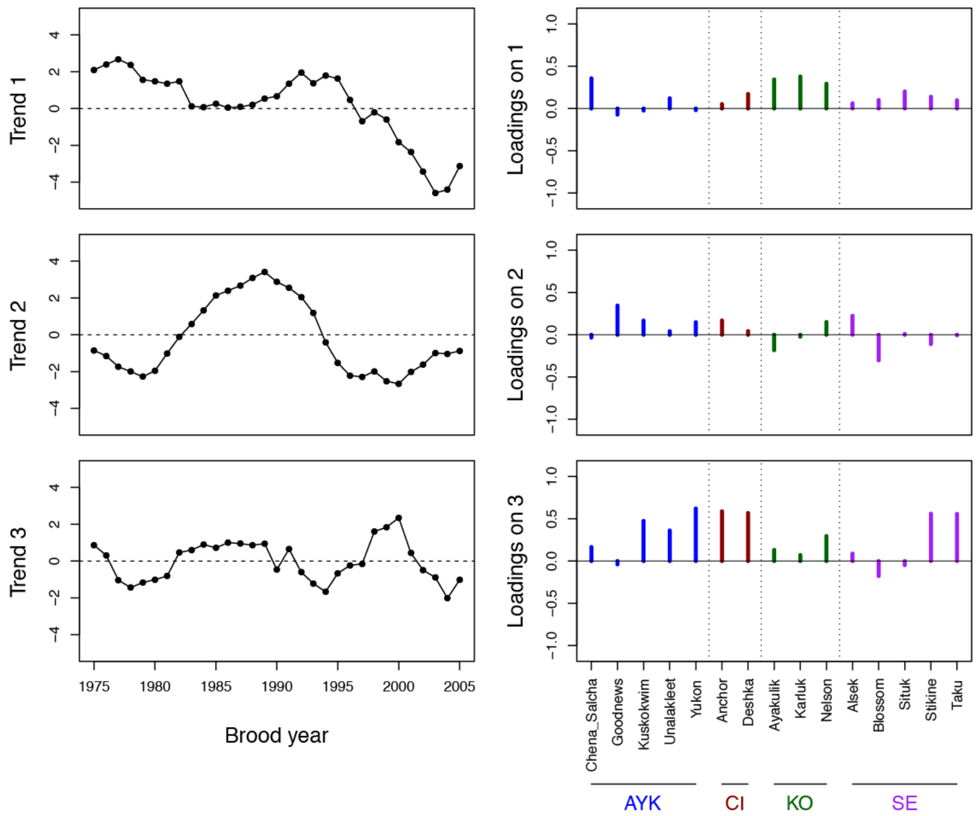
Mark Scheuerell, Eric Ward (NOAA); Stephanie Hampton, Lindsay Scheef (NCEAS)

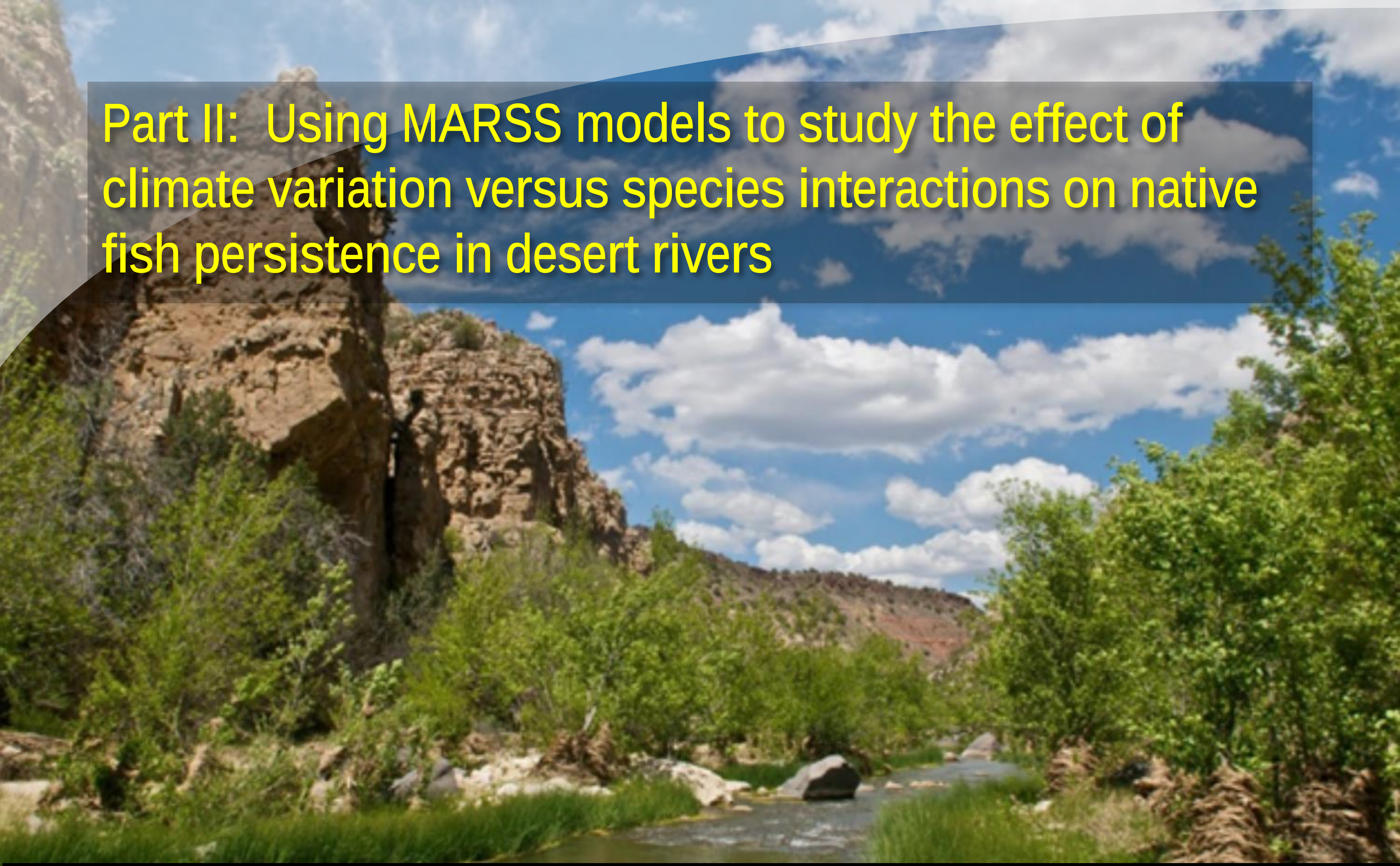
Dynamic Factor Analysis: PCA for time series”

Lot of time series



Reduce to a few hidden drivers with weighting

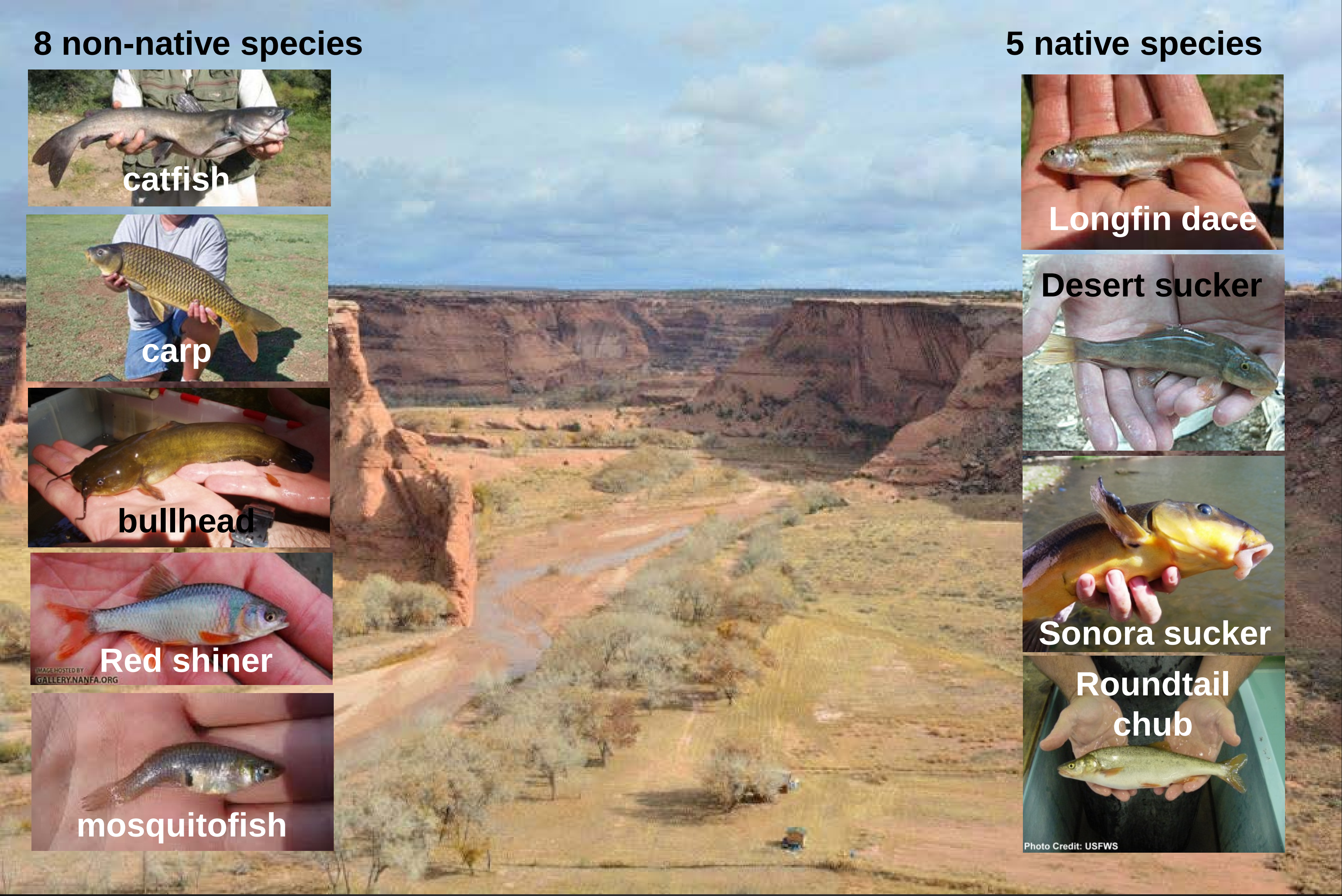




Part II: Using MARSS models to study the effect of climate variation versus species interactions on native fish persistence in desert rivers

Albert Ruhí, Postdoctoral Scholar
John Sabo, PI

ASU GLOBAL INSTITUTE of SUSTAINABILITY

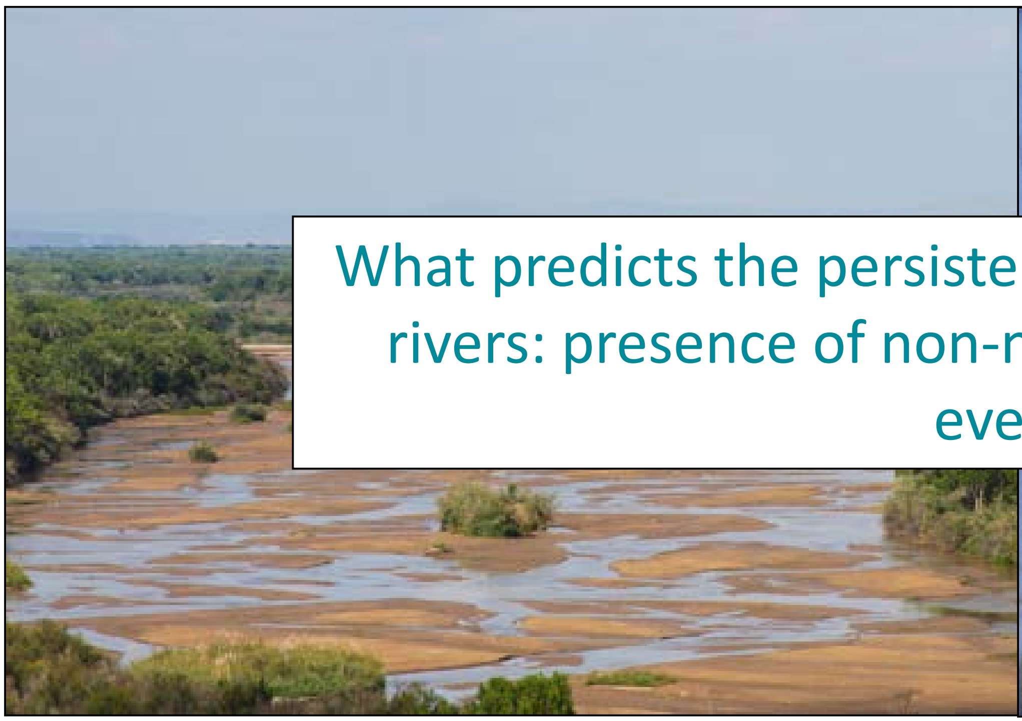


“Poor” but highly endemic, imperiled, and invaded fish communities
(Olden & Poff 2006; Olden et al. 2006; Jaeger et al. 2014; Pool & Olden 2012)

These systems are prone to large discharge variation: flash flood and doughts

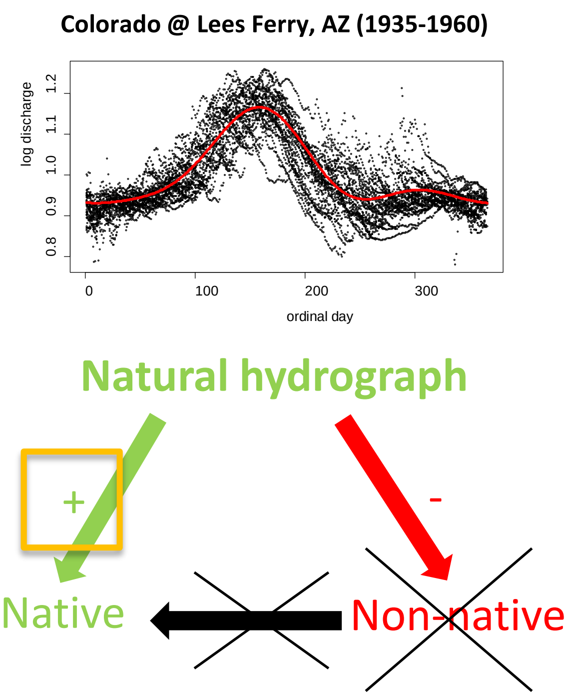
The American Southwest is experiencing increases in extreme low water events

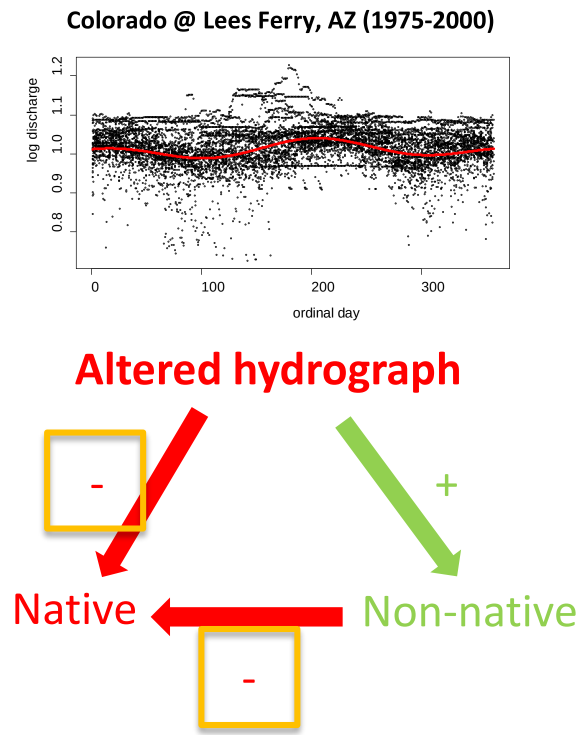
- Increasing human appropriation of freshwater resources and associated low & zero flows
- Expanding climatic droughts caused by global warming



What predicts the persistence of native fish in desert rivers: presence of non-natives or extreme water events

The presence of non-natives confounds detection of the effect of flow





Direct effects of flow or flow-mediated effects of non-natives?

Global Change Biology 2015, 21:1482-1496.

Anomalous droughts, not invasion, decrease persistence of native fishes in a desert river

Albert Ruhí[1], Elizabeth E. Holmes[2], John N. Rinne[3] & John L. Sabo[1,4]

① Global Institute of Sustainability, Arizona State University, Tempe AZ;
② Northwest Fisheries Science Center, NOAA Fisheries Service, Seattle WA;
③ Southwest Forest Science Complex, Rocky Mountain Research Station, Flagstaff AZ;
④ School of Life Sciences, Arizona State University, Tempe AZ

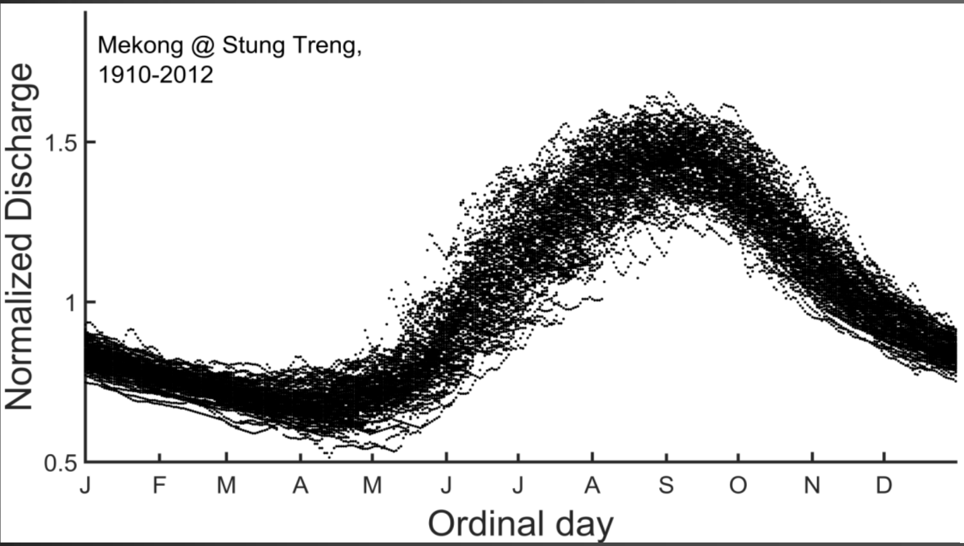
Upper Verde River (AZ)

7 sites
(Rocky Mountain Research Station, USDA Forest Service)
Sullivan Lake to Sycamore Creek
Fish monitored yearly for 15 years (1994-2008)
Daily discharge (USGS station 09504000, “Verde River near Clarkdale, AZ”) for 45 years (1966-2010)

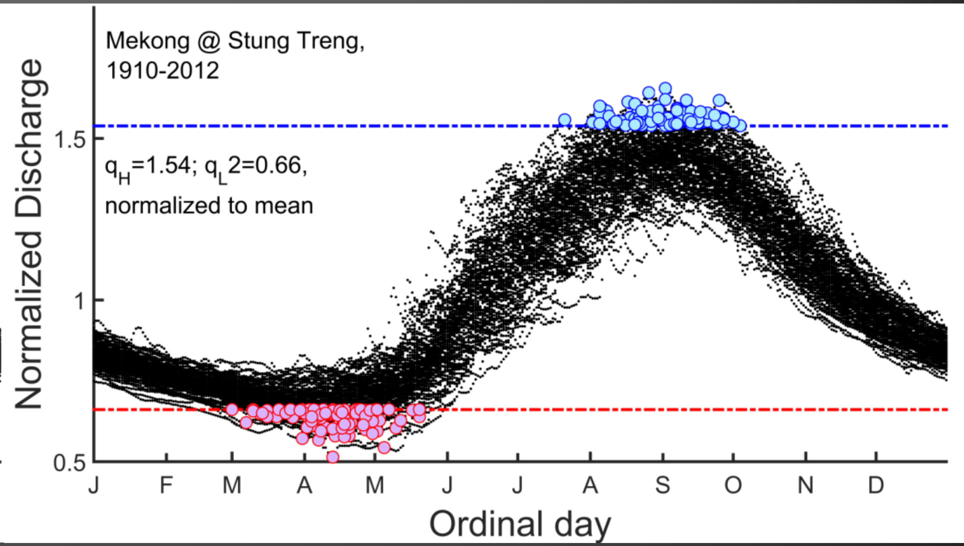
…ideal setting to study the abiotic and biotic factors that influence fish assemblage structure in desert rivers
(e.g. Stefferud & Rinne, 1995; Rinne & Stefferud, 1996)

Methods: What is DFFT?

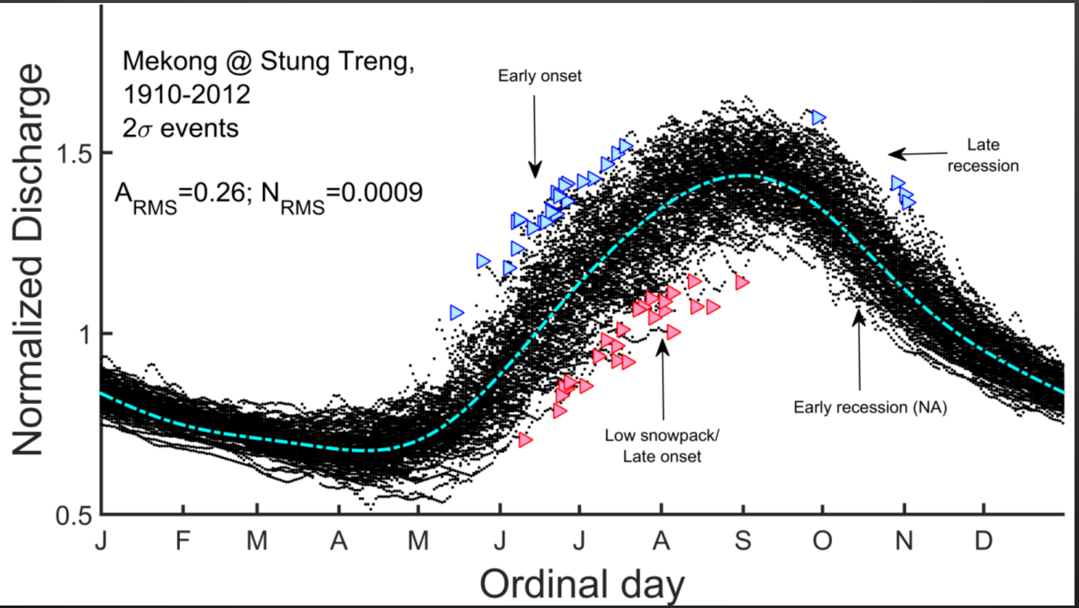
100-year hydrograph



Extreme events (Q_2)

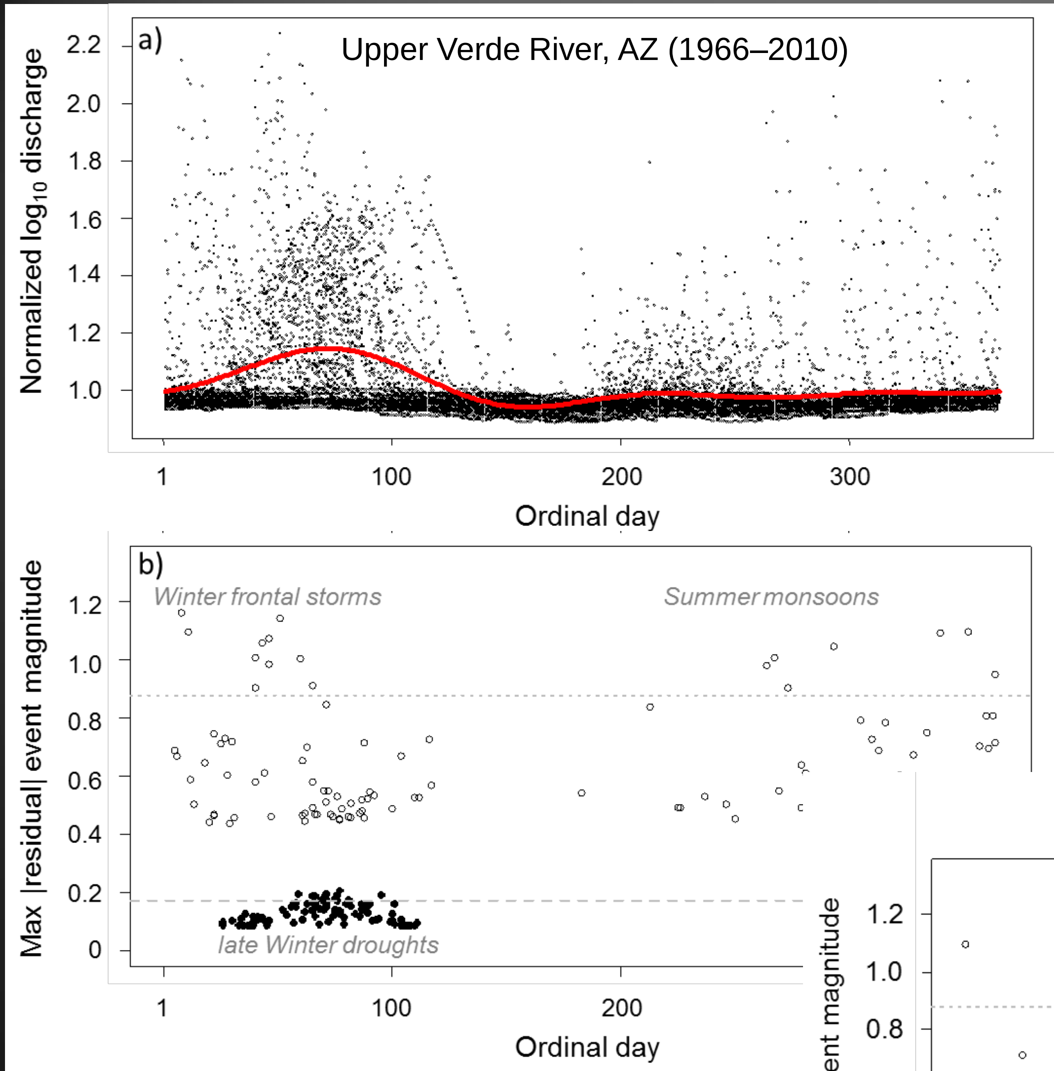


Spectral anomalies (2σ)



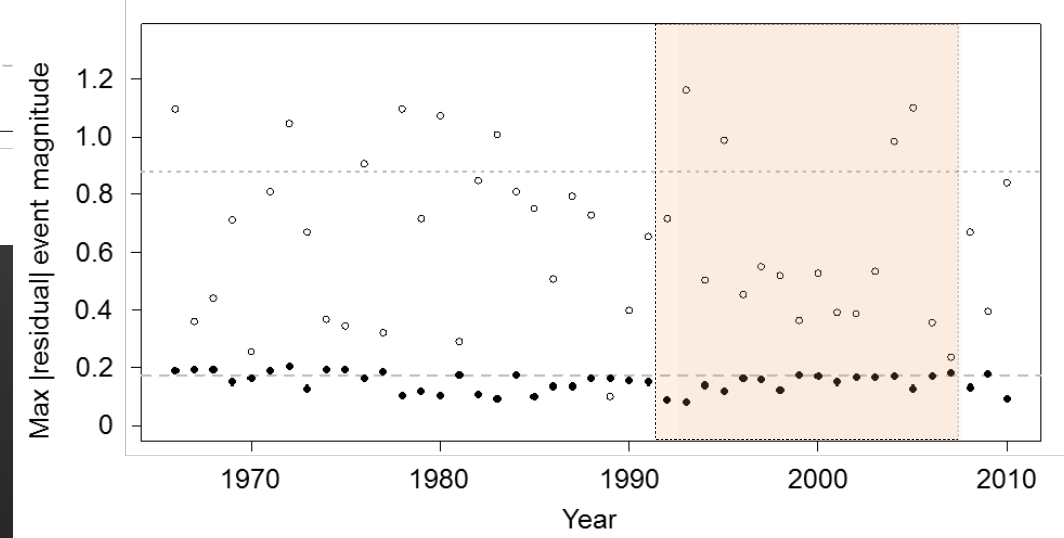
Discrete Fast Fourier Transform (DFFT) to extract characteristic signals (freq., phase, amplitude) and identify anomalous high and low flows

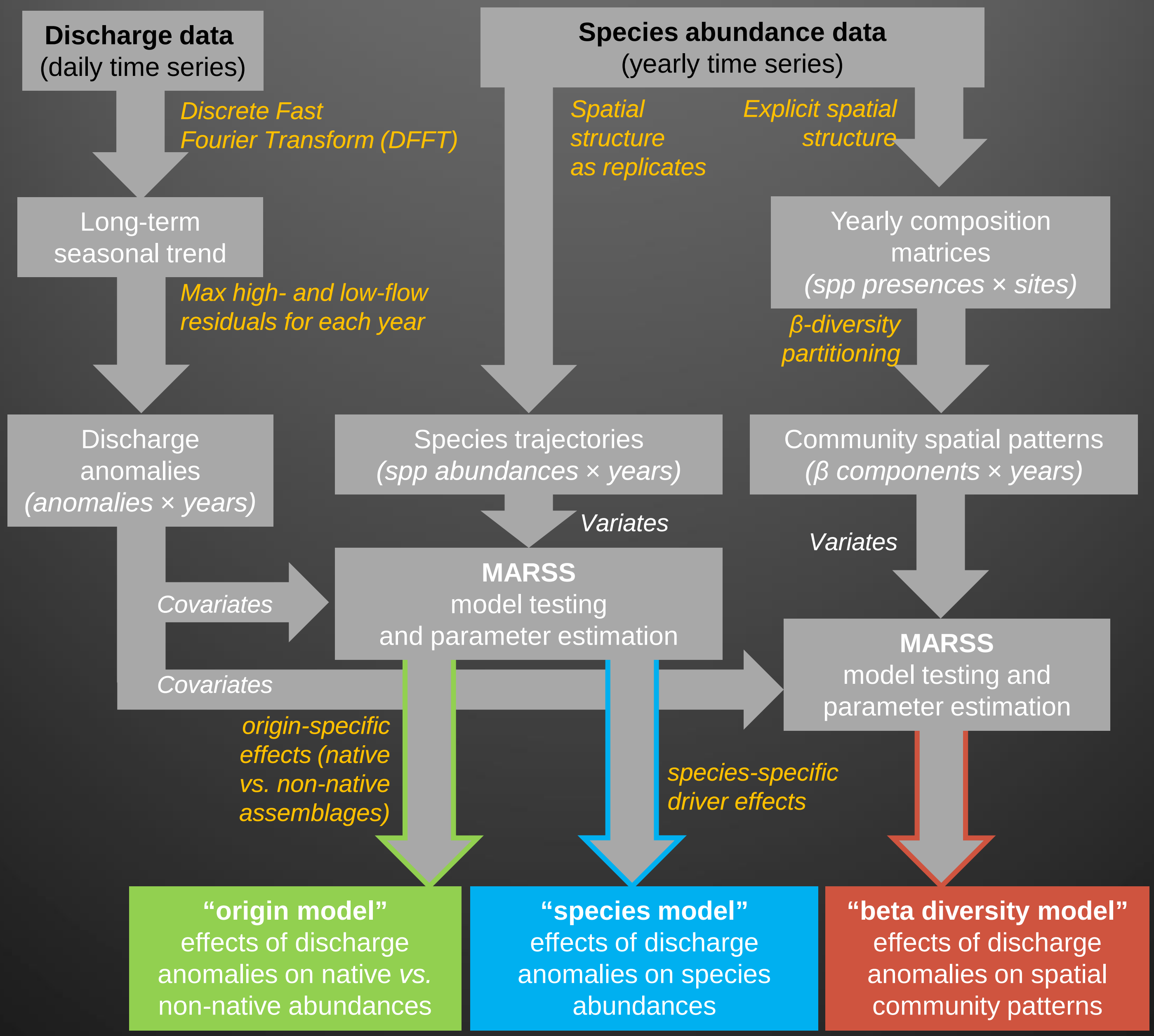
(after Sabo & Post 2008)



'Discharge' R package

Fish data





Discharge data
(daily time series)
Discrete Fast Fourier Transform (DFFT)
Long-term seasonal trend
Max high- and low-flow residuals for each year
Discharge anomalies
(anomalies × years)
Species abundance data
(yearly time series)
Spatial structure as replicates
Explicit spatial structure
Yearly composition matrices
(spp presences × sites)
β-diversity partitioning
Species trajectories
(spp abundances × years)
Community spatial patterns
(β components × years)
Variates
Variates
Covariates
MARSS
model testing and parameter estimation
Covariates
MARSS
model testing and parameter estimation
origin-specific effects (native vs. non-native assemblages)
species-specific driver effects
“origin model”
effects of discharge anomalies on native *vs.* non-native abundances
“species model”
effects of discharge anomalies on species abundances
“beta diversity model”
effects of discharge anomalies on spatial community patterns

Written in matrix form:

The effect of species 1 on species 2

covariates

$$\begin{bmatrix} x_{1,t} \\ x_{2,t} \\ x_{3,t} \end{bmatrix} = \underbrace{\begin{bmatrix} 1+\alpha_{1,1} & \alpha_{1,2} & \alpha_{1,3} \\ \alpha_{2,1} & 1+\alpha_{2,2} & \alpha_{2,3} \\ \alpha_{3,1} & \alpha_{3,2} & 1+\alpha_{3,3} \end{bmatrix}}_{\text{B = interaction matrix}} \begin{bmatrix} x_{1,t-1} \\ x_{2,t-1} \\ x_{3,t-1} \end{bmatrix} + \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} + \underbrace{\begin{bmatrix} c_{1,1} & c_{1,2} \\ c_{2,1} & c_{2,2} \\ c_{3,1} & c_{3,2} \end{bmatrix}}_{\text{covariate effect}} \overbrace{\begin{bmatrix} \upsilon_{1,t-1} \\ \upsilon_{2,t-1} \end{bmatrix}}^{\text{covariates}} + \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix}$$

B = interaction matrix

covariate effect

The effect of covariate 1 on species 1

Environmental variation (not from covariates)

Methods: What is a MARSS model?

Multivariate Autoregressive State-Space model

- Extensively used in econometrics
  (in ecology mainly in fisheries, rev. in Hampton et al. 2013 Ecology)
- Rely on theory about temporal correlation patterns that emerge from environmental drivers and species interactions (Ives et al. 2003 Ecol Monogr)

In the matrix form:

$\mathbf{x}_t = \mathbf{B}\mathbf{x}_{t-1} + \mathbf{C}\mathbf{c}_{t-1} + \mathbf{w}_t$, where $\mathbf{w}_t \sim \mathrm{MVN}(0,\mathbf{Q}_t)$ (Process model)

$\mathbf{y}_t = \mathbf{Z}\mathbf{x}_t + \mathbf{v}_t$, where $\mathbf{v}_t \sim \mathrm{MVN}(0,\mathbf{R}_t)$ (Observation model)

covariate data

covariate effects! (interesting)

“true” abundances
last year

“true” abundances
this year

spp interactions

“true” abundances
this year

observed abundances
this year

observation error

‘MARSS’ R-package (Holmes et al. 2014)

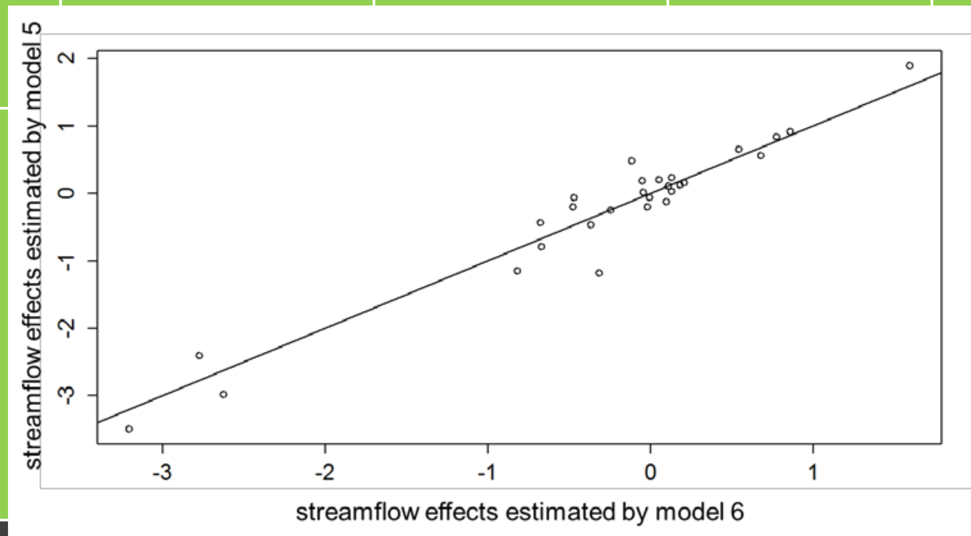
Model selection: ‘origin’ and ‘species’ models

Model					Cc-min(AICc)
ORIGIN MODEL					0
					107.4
					425.0
					539.0

i … ience long-

t

i … iffer?

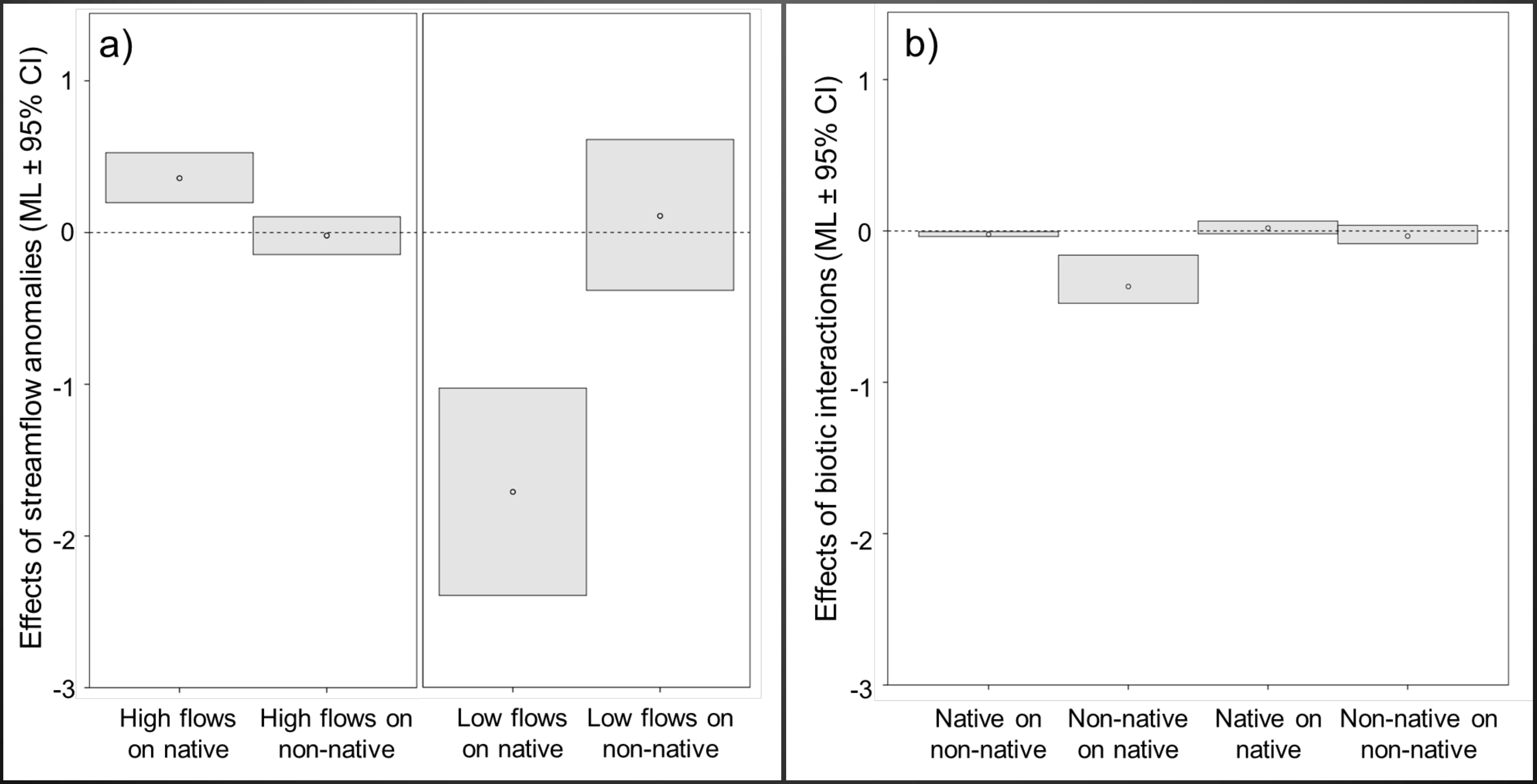


Model		Biotic interactions	Obs. errors	AICc	AICc-min(AICc)
SPECIES MODEL	5	YES, species-specific	species-specific	1104.2	133.5
	6	none	species-specific	1104.6	133.9
	7	YES, species-specific	constant	1559.3	588.6
	8	none	constant	1538.2	567.5

Origin model

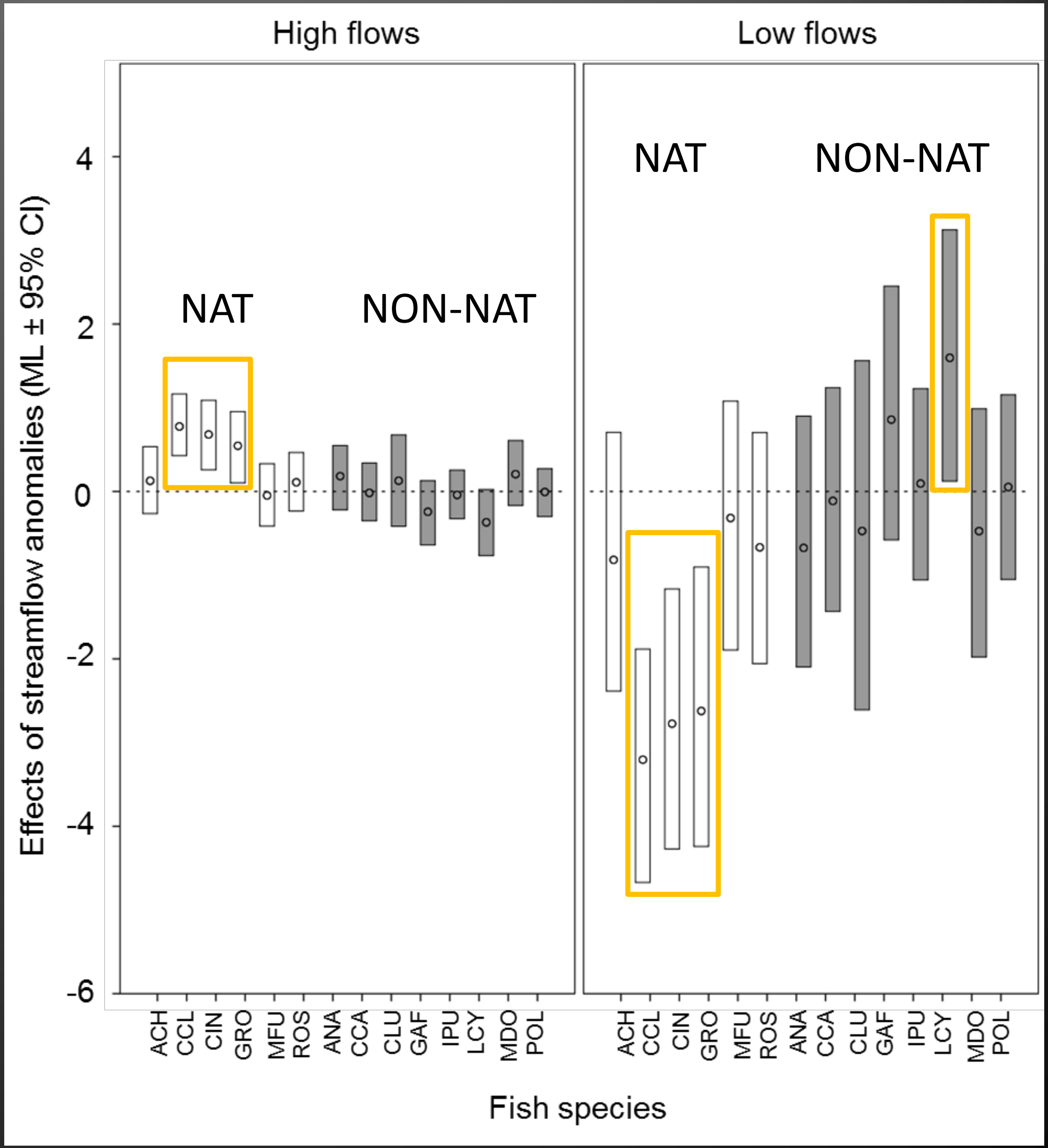
Abiotic drivers

Biotic drivers



Species model (only abiotic)

	Code	English name	Latin name	FUni
Native	ACH	Longfin dace	*Agosia chrysogaster*	0.783
	CCL	Desert sucker	*Catostomus clarki*	0.822
	CIN	Sonora sucker	*Catostomus insignis*	1.000
	GRO	Roundtail chub	*Gila robusta*	0.900
	MFU	Spike dace	*Meda fulgida*	0.860
	ROS	Speckled dace	*Rhinichthys osculus*	0.919
	ANA	Yellow bullhead	*Ameiurus natalis*	0.860
	CCA	Common carp	*Cyprinus carpio*	0.783



Trait differences between NAT and NON-NAT

(iii) Were differences in responses between the native and the non-native assemblage explained by differences in their biological traits?

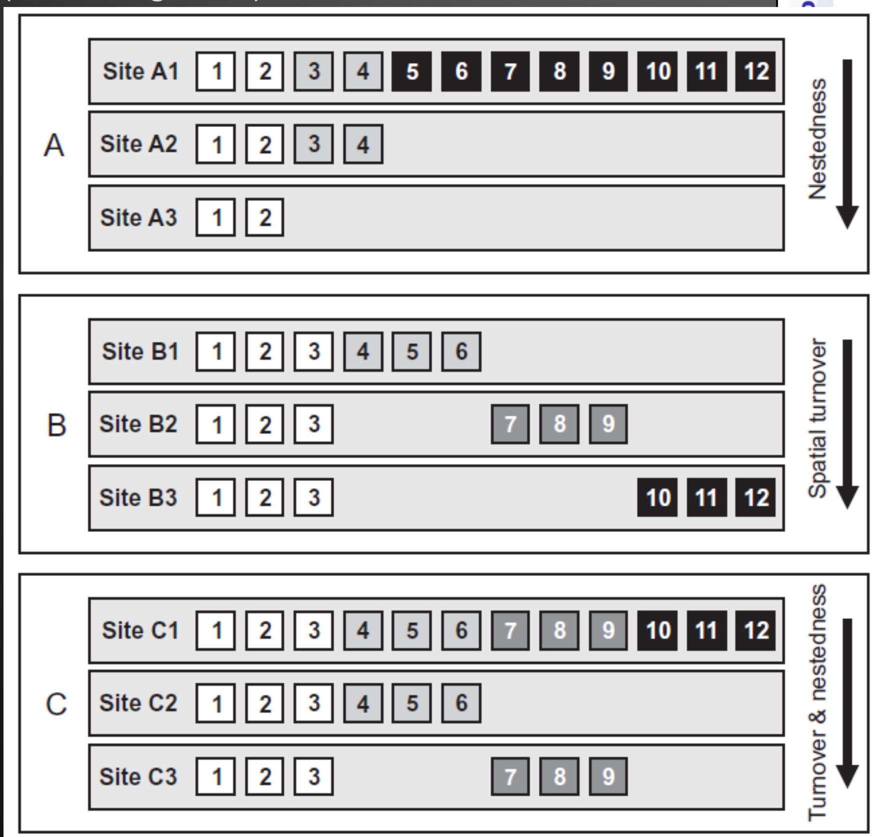
SIMPER (SIMilarity PERcentages) test, native *vs.* non-native species
Based on the Fish Traits database (Frimpong & Angermeier 2009)

Trait type	Trait	Native	Non-native	Diss/SD	Contrib%
body size	Max total length	4.12	9.41	1.17	3.02
habitat preference	preference for lentic systems	2.08	10.94	1.62	3.88
	preference for moderate current	11.11	4.17	1.20	2.63
	preference for clay or silt substrate	4.76	8.93	1.02	3.08
life history	length of the spawning season	8.15	6.38	1.24	2.00
	longevity	4.90	8.82	1.10	2.41
	serial or batch spawner	9.52	5.36	1.01	3.19
trophic ecology	fishes and large macroinvertebrates	2.78	10.42	1.11	4.12
	detritus	9.52	5.36	1.03	3.18

Spatial patterns: beta diversity model

beta diversity partitioning

(after Baselga 2010)



Global Ecology and Biogeography, (Global Ecol. Biogeogr.) (2010) **19**, 134–143

MACROECOLOGICAL METHODS

Partitioning the turnover and nestedness components of beta diversity

Journal of Biogeography (J. Biogeogr.) (2014) **41**, 749–761

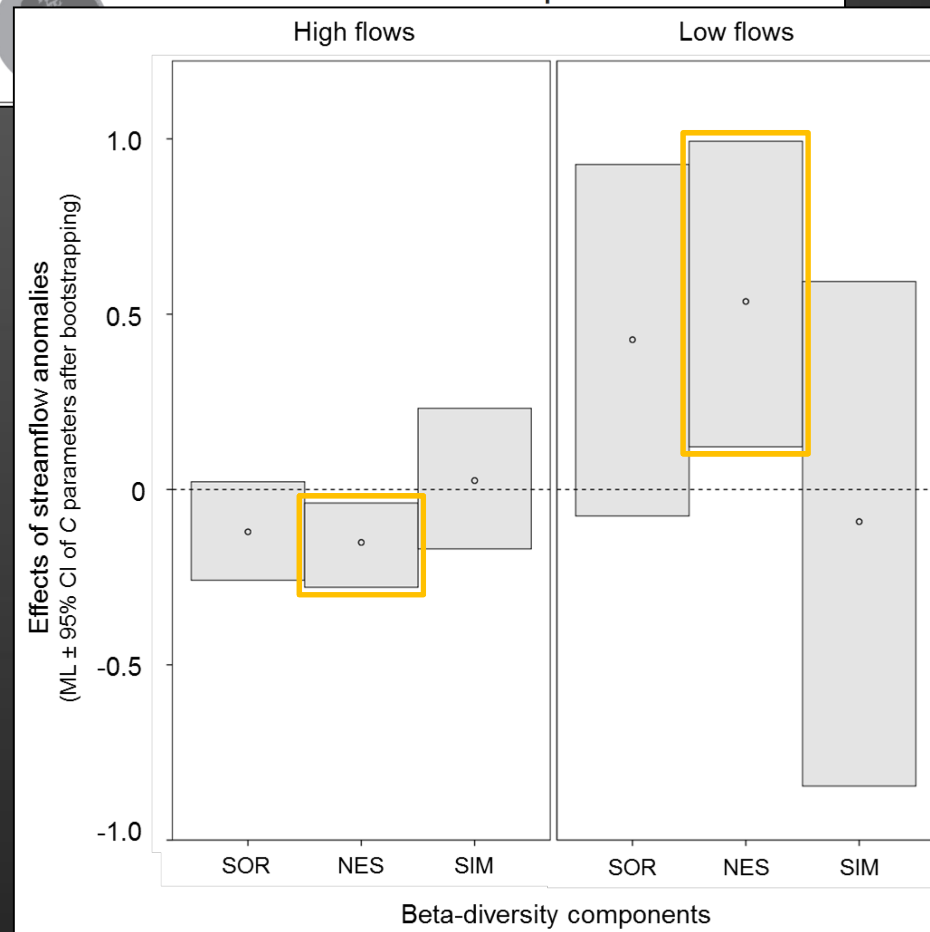
ORIGINAL ARTICLE

Partitioning taxon, phylogenetic and functional beta diversity into replacement and richness difference

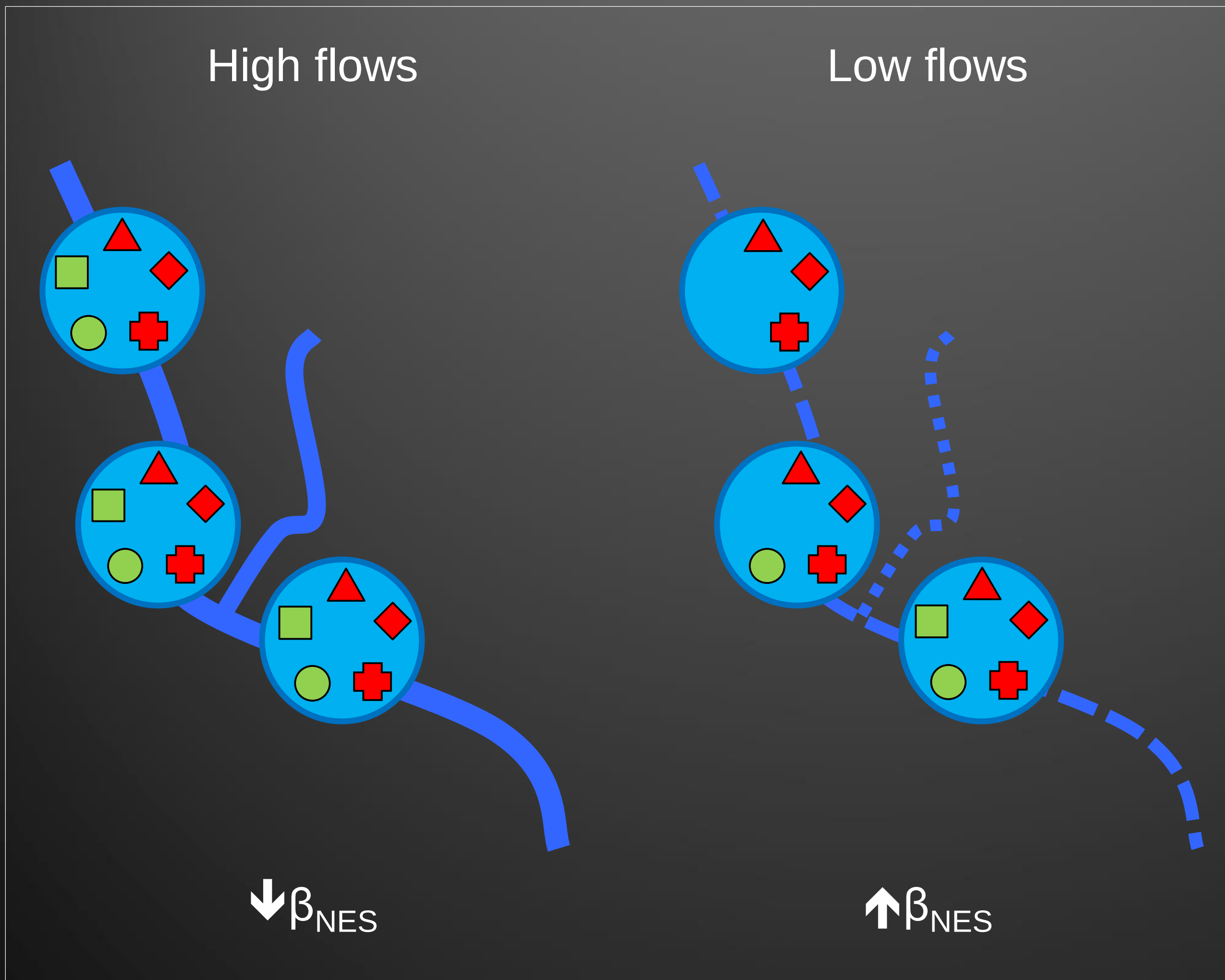
Global Ecology and Biogeography, (Global Ecol. Biogeogr.) (2014) **23**, 1324–1334

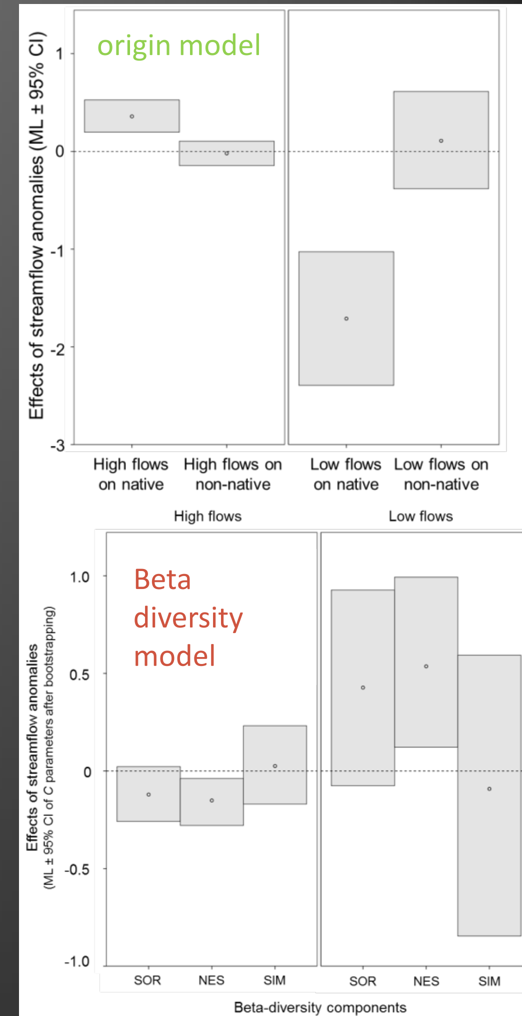
MACROECOLOGICAL METHODS

Interpreting the replacement and richness difference components of



Beta diversity model





Non-natives
Natives

Ongoing debate: deterministic or stochastic?

- "Classical" fish studies: biotic = deterministic, abiotic = stochastic…
  (e.g., Grossman 1982, Rahel et al. 1984, Grossman et al. 1985, Moyle & Vondracek 1985…)

Vol. 126, No. 2 The American Naturalist August 1985

STOCHASTICITY AND ASSEMBLAGE ORGANIZATION IN AN INDIANA STREAM FISH ASSEMBLAGE

In 1982, Grossman et al. purported to demonstrate that a stream fish assemblage was probably affected more by unpredictable environmental disturbances than by interactive processes such as competition. The main purpose of that paper was to

STOCHASTIC OR DETERMINISTIC REGULATION OF ASSEMBLAGE STRUCTURE? IT MAY DEPEND ON HOW THE ASSEMBLAGE IS DEFINED

The relative importance of deterministic versus stochastic processes in structuring communities is of current interest to ecologists (Connor and Simberloff 1979; Sale 1980*a*; Gilpin and Diamond 1982; Matthews 1982). As summarized by Grossman et al. (1982), the determinist school of thought maintains that assem-

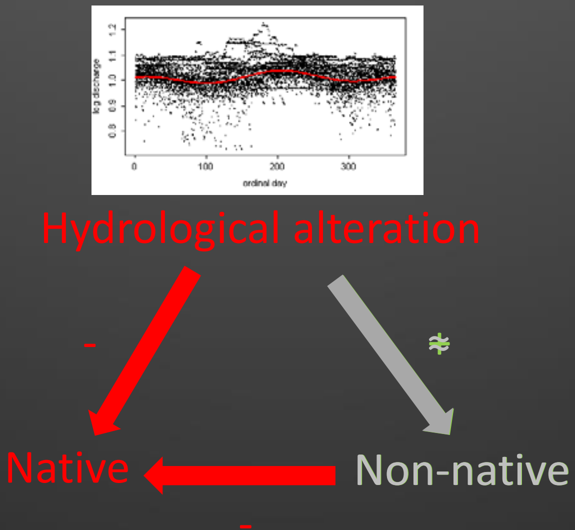
  - …false dichotomy? Communities increasingly driven by discharge variation as the stochastic component of environmental variation increases (Grossman & Sabo 2010)

- Environmental variation is key and may underlie the commonly-observed negative co-variation between native and non-native faunas in highly variable environments

Implications

- ✓ Non-natives as a symptom of hydrologic alteration (i.e. “passengers, not drivers” of native decline)
  [e.g. similar to Cottonwood vs. Saltcedar, Stromberg et al. 2007 GEB]

- ✓ Invasion may not be addressed without better water management



Is this a Southwest-wide pattern?

A different approach to parameter estimation for hidden state problems: Expectation-Maximization algorithms

Holmes, E. E. 2013. Derivation of the EM algorithm for ***constrained*** and unconstrained multivariate autoregressive state-space (MARSS) models. Technical Report. arXiv:1302.3919 [stat.ME]

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}_t = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}_{t-1} + \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}_t, \quad \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}_t \sim MVN\left(\begin{bmatrix} u_1 \\ u_2 \end{bmatrix}, \begin{bmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{bmatrix} \right)$$

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}_t = \begin{bmatrix} z_{11} & z_{12} \\ z_{21} & z_{22} \\ z_{31} & z_{32} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}_t + \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}_t, \quad \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}_t \sim MVN\left(\begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}, \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \right)$$

$$\alpha + \beta_1 p_1 + \beta_2 p_2 + \dots$$

What's the point?

1) It can make **multivariate** AR **state-space** model fitting problems tractable by being considerably faster and more stable
2) For many of the problems we work on, other approaches grind to a halt

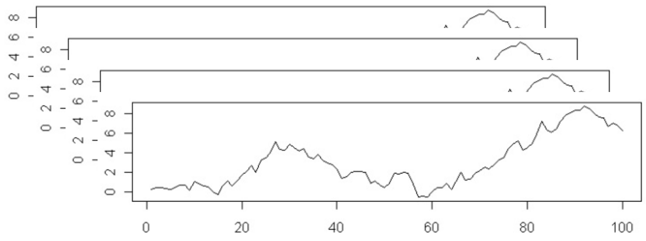
MARSS model

Multivariate autoregressive "random walk"

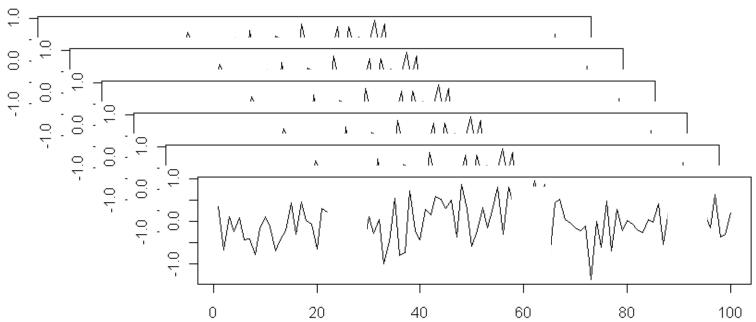
$$\mathbf{x}_t = \mathbf{B}\mathbf{x}_{t-1} + \mathbf{u} + \mathbf{w}_t, \text{ where } \mathbf{w}_t \sim \text{MVN}(0, \mathbf{Q})$$

$$\mathbf{y}_t = \mathbf{Z}\mathbf{x}_t + \mathbf{a} + \mathbf{v}_t, \text{ where } \mathbf{v}_t \sim \text{MVN}(0, \mathbf{R})$$

Multivariate with noise



hidden



observe

written out....

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}_t = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}_{t-1} + \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}_t, \quad \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}_t \sim MVN\left(\begin{bmatrix} u_1 \\ u_2 \end{bmatrix}, \begin{bmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{bmatrix} \right)$$

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}_t = \begin{bmatrix} z_{11} & z_{12} \\ z_{21} & z_{22} \\ z_{31} & z_{32} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}_t + \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}_t, \quad \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}_t \sim MVN\left(\begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}, \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \right)$$

"We have not succeeded in answering all our problems. The answers we have found only serve to raise a whole set of new questions. In some ways we feel we are as confused as ever, but we believe we are confused on a higher level and about more important things."

Earl C. Kelley, 1951, 'The Workshop Way of Learning'