

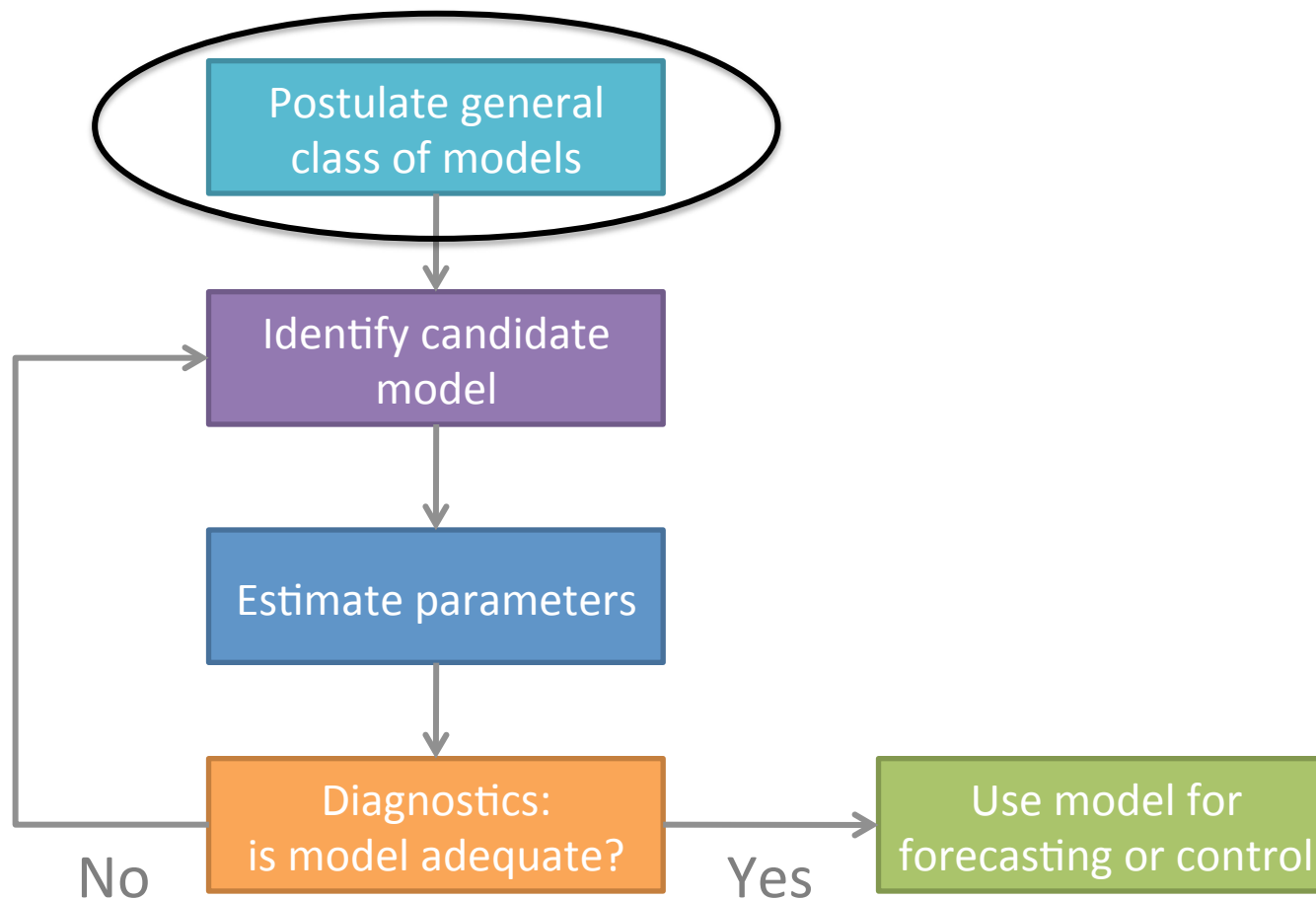
An introduction to Dynamic Factor Analysis

Mark Scheuerell

FISH 507 – Applied Time Series Analysis

9 February 2017

Iterative approach to model building



Observations of one or more states

- We have learned several forms of models for analyzing ts
- General idea is to reduce the ts to trend(s), seasonal effect(s) & stationary remainder
- When we have multiple ts, we have modeled them as observations of 1+ states (see Lec 6)

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

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

Observations of one or more states

For example, recall our analyses of Pacific harbor seals

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

3 "true" populations

$$\begin{bmatrix} x_{JF,t} \\ x_{N,t} \\ x_{S,t} \end{bmatrix} = \begin{bmatrix} x_{JF,t-1} \\ x_{N,t-1} \\ x_{S,t-1} \end{bmatrix} + \begin{bmatrix} u_{JF} \\ u_N \\ u_S \end{bmatrix} + \begin{bmatrix} w_{JF,t} \\ w_{N,t} \\ w_{S,t} \end{bmatrix}$$

$$\mathbf{y}_t = \mathbf{Z}\mathbf{x}_t + \mathbf{a} + \mathbf{v}_t$$

5 observations

$$\begin{bmatrix} y_{1,t} \\ y_{2,t} \\ y_{3,t} \\ y_{4,t} \\ y_{5,t} \end{bmatrix} = \begin{array}{c} \text{JF} \quad \text{N} \quad \text{S} \\ \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix} \end{array} \times \begin{bmatrix} x_{JF,t} \\ x_{N,t} \\ x_{S,t} \end{bmatrix} + \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \end{bmatrix} + \begin{bmatrix} v_{1,t} \\ v_{2,t} \\ v_{3,t} \\ v_{4,t} \\ v_{5,t} \end{bmatrix}$$

Finding common patterns in data

- What if our observations were not of simply 1 state, but were instead a mixture of 2+ states (eg, we sampled a haul-out in between 2 breeding sites)?

Observations of one or more states

Returning to our analyses of Pacific harbor seals

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

3 "true" populations

$$\begin{bmatrix} x_{JF,t} \\ x_{N,t} \\ x_{S,t} \end{bmatrix} = \begin{bmatrix} x_{JF,t-1} \\ x_{N,t-1} \\ x_{S,t-1} \end{bmatrix} + \begin{bmatrix} u_{JF} \\ u_N \\ u_S \end{bmatrix} + \begin{bmatrix} w_{JF,t} \\ w_{N,t} \\ w_{S,t} \end{bmatrix}$$

$$\mathbf{y}_t = \mathbf{Z}\mathbf{x}_t + \mathbf{a} + \mathbf{v}_t$$

5 observations

$$\begin{bmatrix} y_{1,t} \\ y_{2,t} \\ y_{3,t} \\ y_{4,t} \\ y_{5,t} \end{bmatrix} = \begin{array}{c} JF \quad N \quad S \\ \begin{bmatrix} 0.8 & 0.2 & 0 \\ 0.2 & 0.7 & 0.1 \\ 0 & 0.9 & 0.1 \\ 0 & 0.3 & 0.7 \\ 0 & 0.1 & 0.9 \end{bmatrix} \end{array} \times \begin{bmatrix} x_{JF,t} \\ x_{N,t} \\ x_{S,t} \end{bmatrix} + \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \end{bmatrix} + \begin{bmatrix} v_{1,t} \\ v_{2,t} \\ v_{3,t} \\ v_{4,t} \\ v_{5,t} \end{bmatrix}$$

Finding common patterns in data

- What if our observations were not of simply 1 state, but were instead a mixture of 2+ states (eg, we sampled a haul-out in between 2 breeding sites)?
- What if some unknown & unmeasured environmental drivers created common patterns among our ts (eg, adult salmon in the North Pacific)?

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

We know how many observations we have, but we don't know the number of states!

$$\mathbf{y}_t = \mathbf{Z}\mathbf{x}_t + \mathbf{a} + \mathbf{v}_t$$

So, what are the dimensions of \mathbf{Z} ?

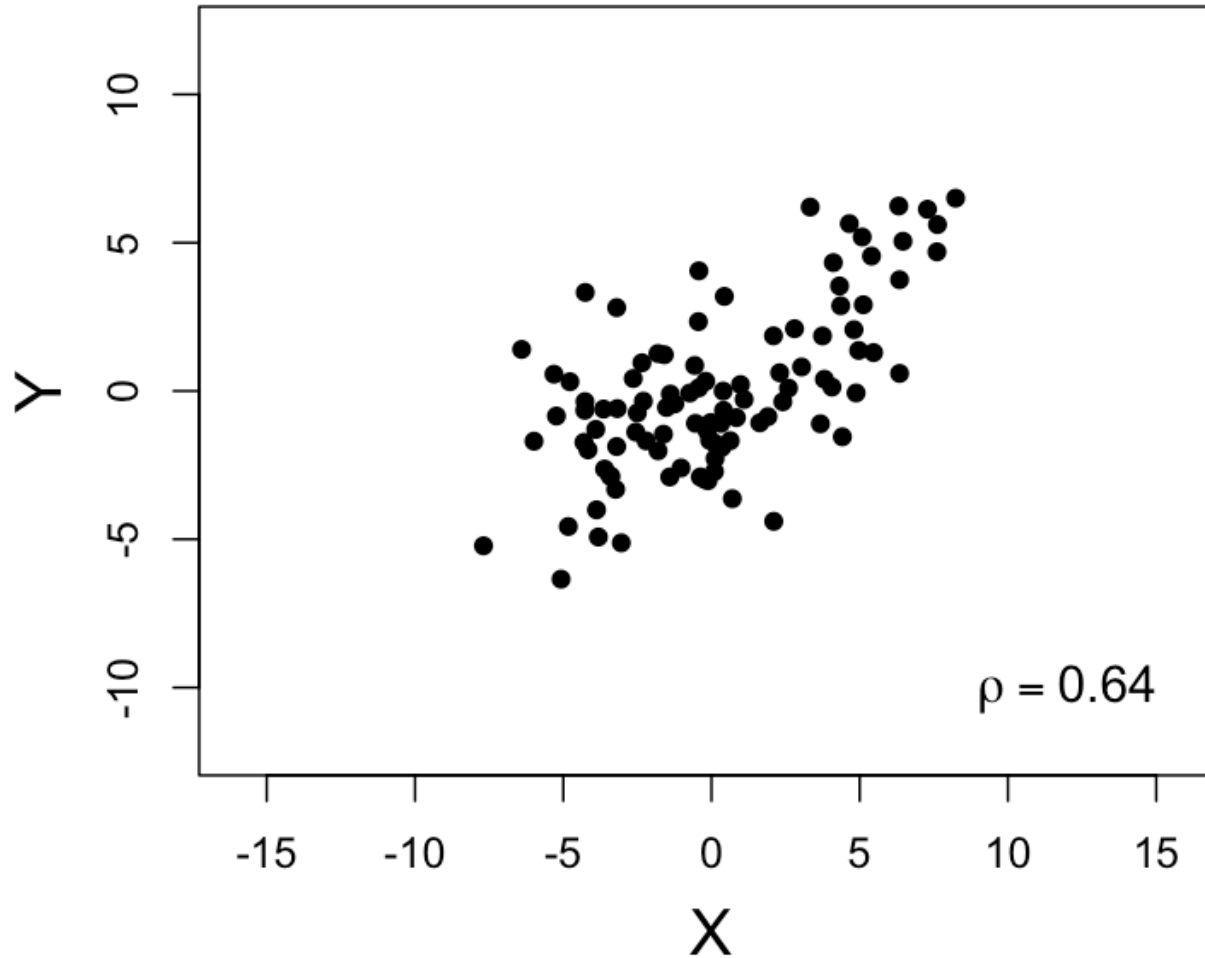
Finding common patterns in data

- What if our observations were not of simply 1 state, but were instead a mixture of 2+ states (eg, we sampled a haul-out in between 2 breeding sites)?
- What if some unknown & unmeasured environmental drivers created common patterns among our ts (eg, adult salmon in the North Pacific)?
- Perhaps we could use a few common trends to describe much/most of the variance in the observations?
- Dynamic Factor Analysis (DFA) is an approach to ts modeling that does just that

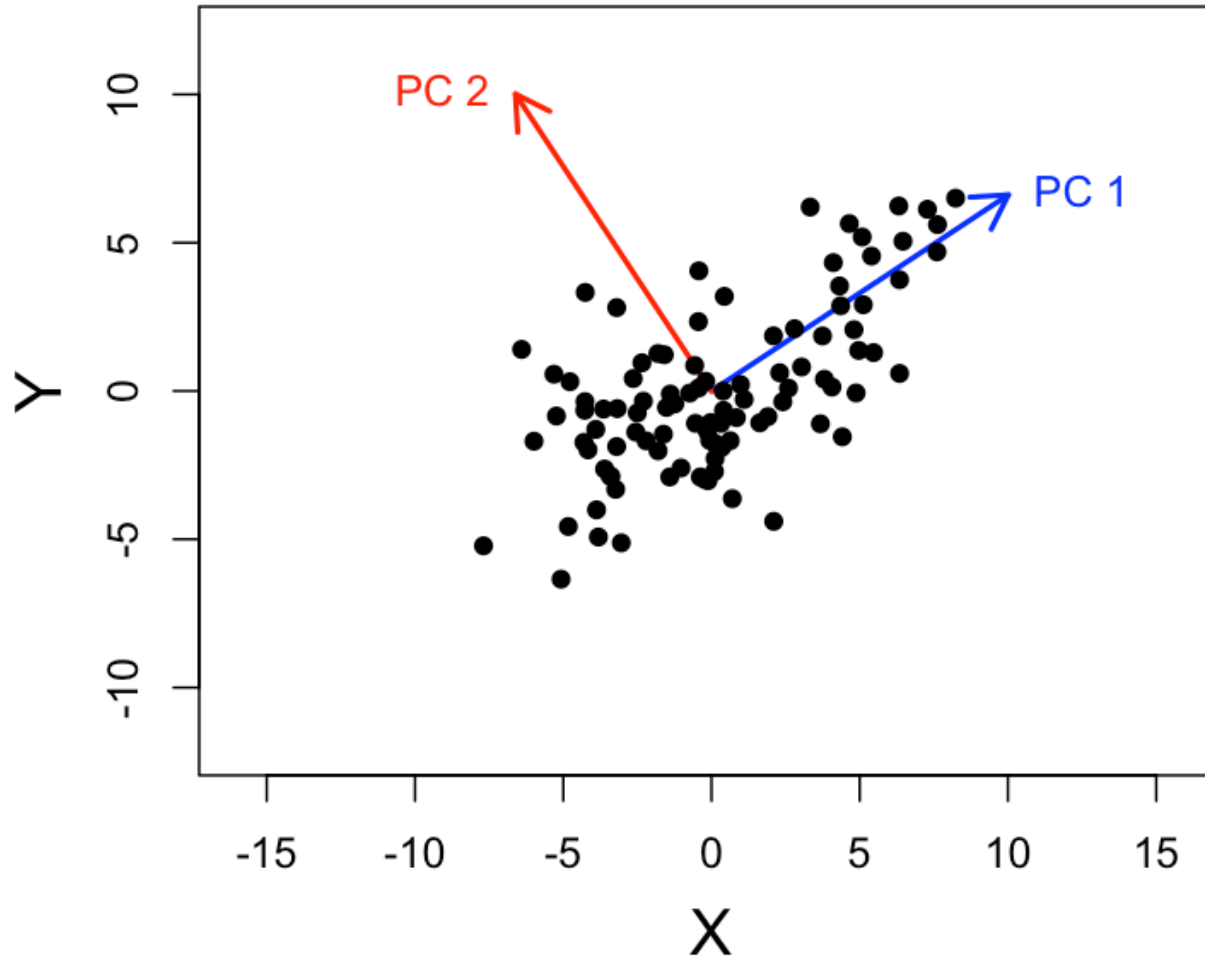
Let's start with PCA

- PCA stands for Principal Component Analysis
- Goal is to reduce some correlated variables to fewer uncorrelated values
- Number of principal components is generally less than the number of original variables

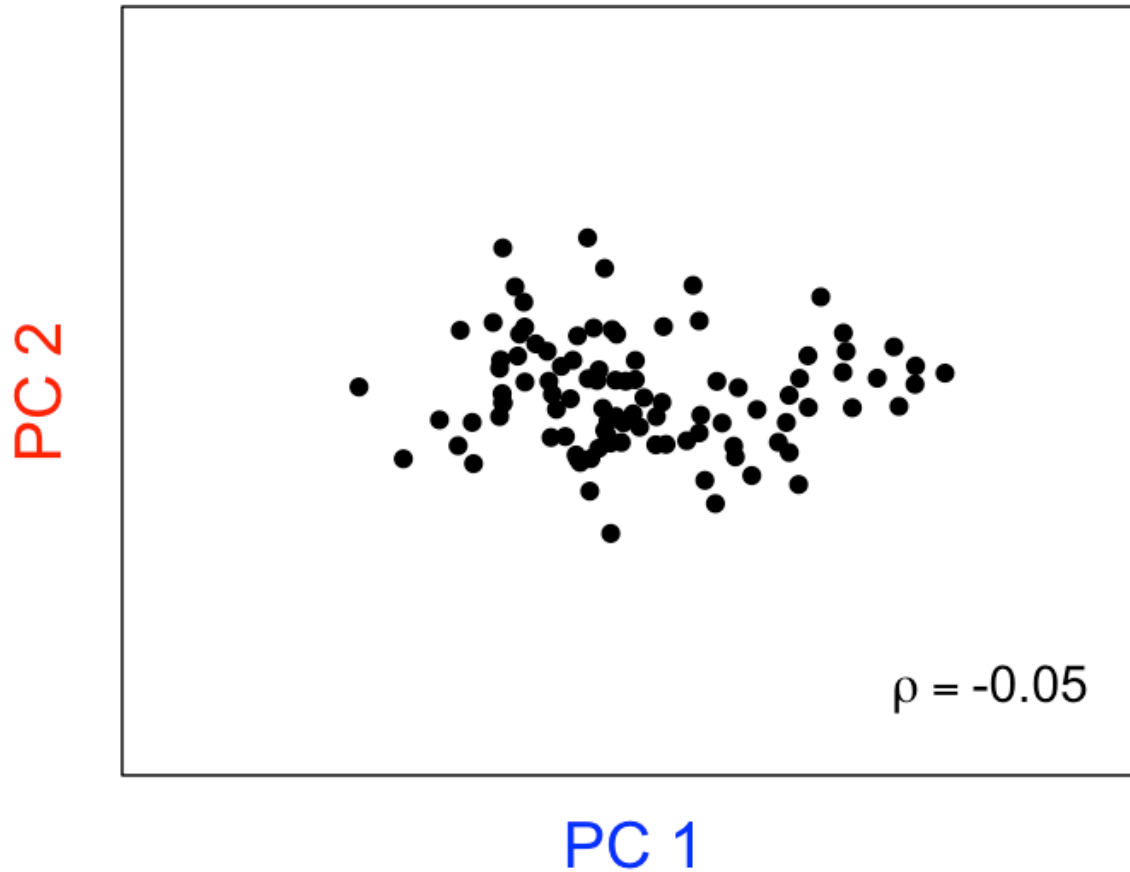
A graphical example



Adding in the first 2 PC's



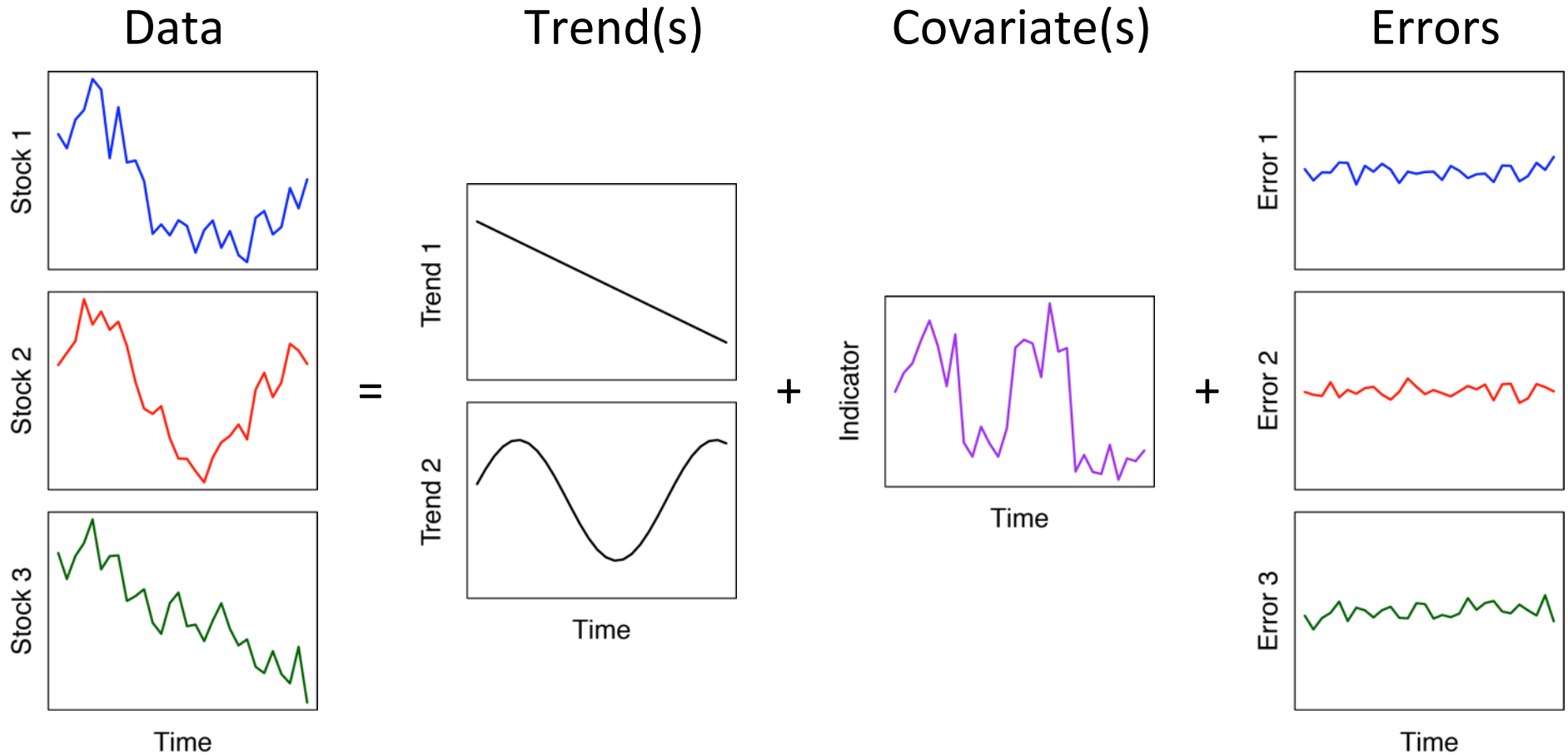
And rotating the basis



What exactly *is* DFA?

- It's like PCA for time series
- DFA can indicate whether there are any:
 - 1) underlying common patterns/trends in the time series,
 - 2) interactions between the response variables, and
 - 3) what the effects of explanatory variables are.
- The mathematics are a bit complex—for details, see Zuur et al. (2003) *Environmetrics*

DFA in common terms



DFA in matrix form

State equation

$$\mathbf{x}_t = \mathbf{x}_{t-1} + \mathbf{w}_t$$

$$\mathbf{w}_t \sim \text{MVN}(\mathbf{0}, \mathbf{Q})$$

Common trends over time

Observation equation

$$\mathbf{y}_t = \mathbf{Z}\mathbf{x}_t + \mathbf{a} + \mathbf{v}_t$$

$$\mathbf{v}_t \sim \text{MVN}(\mathbf{0}, \mathbf{R})$$

Relate trends (\mathbf{x}) to observations (\mathbf{y}) via \mathbf{Z}

DFA with covariates

State equation

$$\mathbf{x}_t = \mathbf{x}_{t-1} + \mathbf{w}_t$$

$$\mathbf{w}_t \sim \text{MVN}(\mathbf{0}, \mathbf{Q})$$

Common trends over time

Observation equation

$$\mathbf{y}_t = \mathbf{Z}\mathbf{x}_t + \mathbf{a} + \mathbf{D}\mathbf{d}_t + \mathbf{v}_t$$

$$\mathbf{v}_t \sim \text{MVN}(\mathbf{0}, \mathbf{R})$$

Relate trends (\mathbf{x}) to observations (\mathbf{y}) via \mathbf{Z} ,
and covariates (\mathbf{d}) to \mathbf{y} via \mathbf{D}

Examining the \mathbf{Z} matrix

Observation equation

$$\mathbf{y}_t = \mathbf{Z}\mathbf{x}_t + \mathbf{a} + \mathbf{D}\mathbf{d}_t + \mathbf{v}_t$$

$$\mathbf{Z} = \begin{array}{c} \text{1 state?} \\ \left[\begin{array}{c} z_{11} \\ z_{12} \\ z_{13} \\ z_{14} \\ z_{15} \end{array} \right] \end{array} \quad \mathbf{Z} = \begin{array}{c} \text{2 states?} \\ \left[\begin{array}{cc} z_{11} & z_{21} \\ z_{12} & z_{22} \\ z_{13} & z_{23} \\ z_{14} & z_{24} \\ z_{15} & z_{25} \end{array} \right] \end{array} \quad \mathbf{Z} = \begin{array}{c} \text{3 states?} \\ \left[\begin{array}{ccc} z_{11} & z_{21} & z_{31} \\ z_{12} & z_{22} & z_{32} \\ z_{13} & z_{23} & z_{33} \\ z_{14} & z_{24} & z_{34} \\ z_{15} & z_{25} & z_{35} \end{array} \right] \end{array}$$

We will use model selection criteria to decide

Relationship between PCA & DFA

- Similarity with PCA can be seen via

$$\text{Cov}(\mathbf{y}_t) = \mathbf{Z}\mathbf{Z}^\top + \mathbf{R}$$

- In PCA, however, \mathbf{R} is constrained to be diagonal
- Not so in DFA

Various forms for R

Diagonal & equal

$$\begin{bmatrix} \sigma & 0 & 0 & 0 \\ 0 & \sigma & 0 & 0 \\ 0 & 0 & \sigma & 0 \\ 0 & 0 & 0 & \sigma \end{bmatrix}$$

Equal variance & covariance

$$\begin{bmatrix} \sigma & \gamma & \gamma & \gamma \\ \gamma & \sigma & \gamma & \gamma \\ \gamma & \gamma & \sigma & \gamma \\ \gamma & \gamma & \gamma & \sigma \end{bmatrix}$$

Diagonal & unequal

$$\begin{bmatrix} \sigma_1 & 0 & 0 & 0 \\ 0 & \sigma_2 & 0 & 0 \\ 0 & 0 & \sigma_3 & 0 \\ 0 & 0 & 0 & \sigma_4 \end{bmatrix}$$

Spp-specific variances & covariances

$$\begin{bmatrix} \sigma_{chin} & 0 & 0 & 0 \\ 0 & \sigma_{coho} & 0 & \gamma_{coho} \\ 0 & 0 & \sigma_{pink} & 0 \\ 0 & \gamma_{coho} & 0 & \sigma_{coho} \end{bmatrix}$$

Some caveats in fitting DFA models

State equation

$$\mathbf{x}_t = \mathbf{x}_{t-1} + \mathbf{w}_t$$

$$\mathbf{w}_t \sim \text{MVN}(\mathbf{0}, \mathbf{Q})$$

Observation equation

$$\mathbf{y}_t = \mathbf{Z}\mathbf{x}_t + \mathbf{a} + \mathbf{D}\mathbf{d}_t + \mathbf{v}_t$$

$$\mathbf{v}_t \sim \text{MVN}(\mathbf{0}, \mathbf{R})$$

Infinite combinations!

Some caveats in fitting DFA models

State equation

$$\mathbf{x}_t = \mathbf{x}_{t-1} + \mathbf{w}_t$$

$$\mathbf{w}_t \sim \text{MVN}(\mathbf{0}, \mathbf{Q})$$

1) Set $\mathbf{Q} = \text{Identity}$

Observation equation

$$\mathbf{y}_t = \mathbf{Z}\mathbf{x}_t + \mathbf{a} + \mathbf{D}\mathbf{d}_t + \mathbf{v}_t$$

$$\mathbf{v}_t \sim \text{MVN}(\mathbf{0}, \mathbf{R})$$

2) Constrain portions of \mathbf{Z} and \mathbf{a}

Constraining the \mathbf{a} vector

Observation equation

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

Constraining portions of \mathbf{a} (eg, $n = 5$; $m = 3$)

$$\mathbf{a} = \begin{bmatrix} a_i \\ a_i \\ a_i \\ a_i \\ a_i \end{bmatrix}$$

Constraining the \mathbf{a} vector

Observation equation

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

Constraining portions of \mathbf{a} (eg, $n = 5$; $m = 3$)

$$\mathbf{a} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ a_i \\ a_i \end{bmatrix}$$

in first m rows of \mathbf{a} , $a_i = 0$

Note: This approach causes the EM algorithm to take a very long time to converge

Soln: We will demean our data and set $\mathbf{a} = \mathbf{0}$

Constraining the \mathbf{Z} matrix

Observation equation

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

Constraining portions of \mathbf{Z} (eg, $n = 5; m = 3$)

$$\mathbf{Z} = \begin{bmatrix} z_{ij} & z_{ij} & z_{ij} \\ z_{ij} & z_{ij} & z_{ij} \\ z_{ij} & z_{ij} & z_{ij} \\ z_{ij} & z_{ij} & z_{ij} \\ z_{ij} & z_{ij} & z_{ij} \end{bmatrix}$$

Constraining the \mathbf{Z} matrix

Observation equation

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

Constraining portions of \mathbf{Z} (eg, $n = 5; m = 3$)

$$\mathbf{Z} = \begin{bmatrix} z_{ij} & 0 & 0 \\ z_{ij} & z_{ij} & 0 \\ z_{ij} & z_{ij} & z_{ij} \\ z_{ij} & z_{ij} & z_{ij} \\ z_{ij} & z_{ij} & z_{ij} \end{bmatrix} \quad \text{in } m-1 \text{ rows of } \mathbf{Z}, z_{ij} = 0 \text{ if } j > i$$

Rotation matrix **H**

We arbitrarily constrained **Z** to obtain just 1 of ∞ solutions

$$(1) \quad \begin{aligned} \mathbf{x}_t &= \mathbf{x}_{t-1} + \mathbf{w}_t \\ \mathbf{y}_t &= \mathbf{Z}\mathbf{x}_t + \mathbf{a} + \mathbf{D}\mathbf{d}_t + \mathbf{v}_t \end{aligned}$$

Rotation matrix \mathbf{H}

If \mathbf{H} is $m \times m$ non-singular matrix, these 2 models are equivalent

$$(1) \quad \begin{aligned} \mathbf{x}_t &= \mathbf{x}_{t-1} + \mathbf{w}_t \\ \mathbf{y}_t &= \mathbf{Z}\mathbf{x}_t + \mathbf{a} + \mathbf{D}\mathbf{d}_t + \mathbf{v}_t \end{aligned}$$

We need to choose appropriate \mathbf{H} — we'll use “varimax”

$$(2) \quad \begin{aligned} \mathbf{H}\mathbf{x}_t &= \mathbf{H}\mathbf{x}_{t-1} + \mathbf{H}\mathbf{w}_t \\ \mathbf{y}_t &= \mathbf{Z}\mathbf{H}^{-1}\mathbf{x}_t + \mathbf{a} + \mathbf{D}\mathbf{d}_t + \mathbf{v}_t \end{aligned}$$

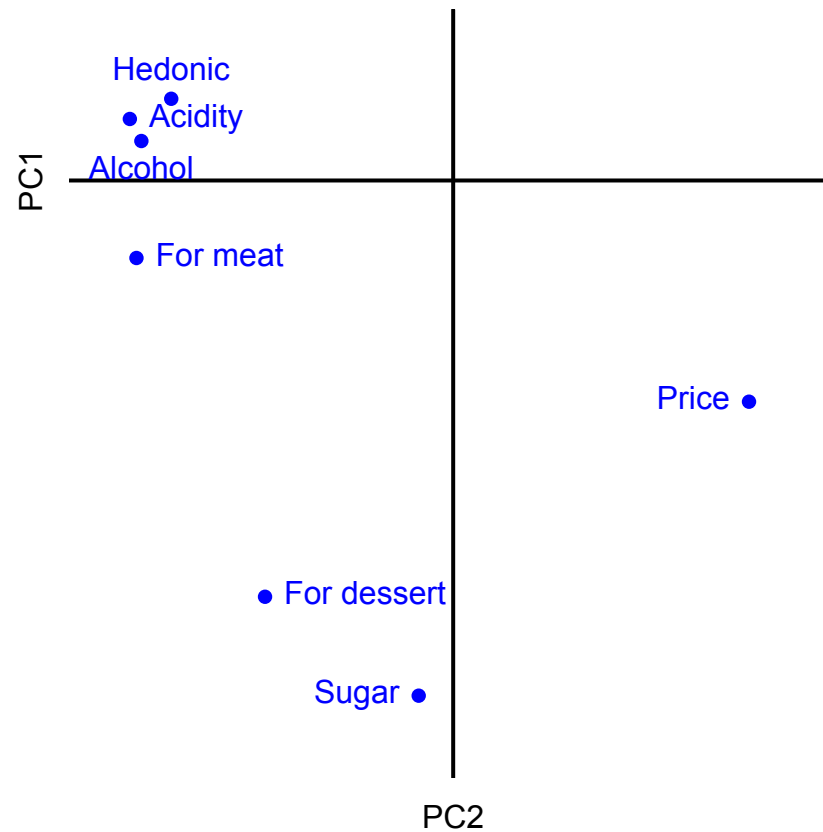
Varimax rotation for H

- A “simple” solution means each factor has a small number of large loadings, and a large number of (near) zero loadings
- After varimax rotation, each original variable tends to be associated with one (or a small number) of factors
- Varimax searches for a linear combination of original factors that *maximizes* the variance of the loadings

$$\max \sum (z_{ij}^2 - \bar{z}_{ij}^2)^2$$

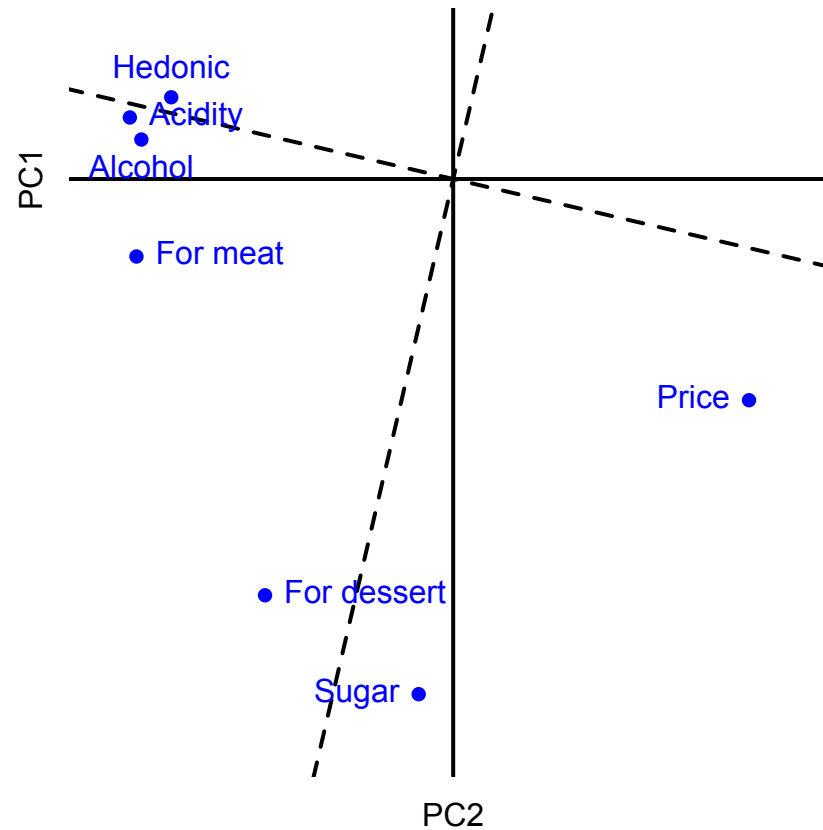
Varimax rotation for H

For example, PCA on wine descriptions



Varimax rotation for H

After varimax rotation



A note of caution for model selection

- Some people compare data support for models with and without covariates, and varying numbers of trends
- But, this has to be done in a well reasoned manner

This is an undetermined random walk

$$\mathbf{y}_t = \mathbf{Z}\mathbf{x}_t + \mathbf{a} + \mathbf{D}\mathbf{d}_t + \mathbf{v}_t$$

This is a predetermined covariate

- Unless \mathbf{d}_t is very highly correlated with \mathbf{y}_t , then the trend-only models will be selected



Population coherence and environmental impacts across spatial scales: a case study of Chinook salmon

JAN OHLBERGER,^{1,†} MARK D. SCHEUERELL,² AND DANIEL E. SCHINDLER¹

¹*School of Aquatic and Fishery Sciences, University of Washington, Seattle, Washington 98195 USA*

²*Fish Ecology Division, Northwest Fisheries Science Center, National Marine Fisheries Service, National Oceanic and Atmospheric Administration, Seattle, Washington 98112 USA*

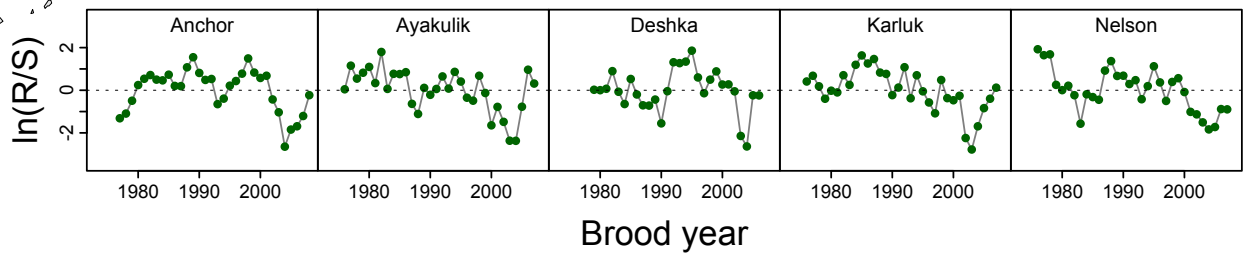
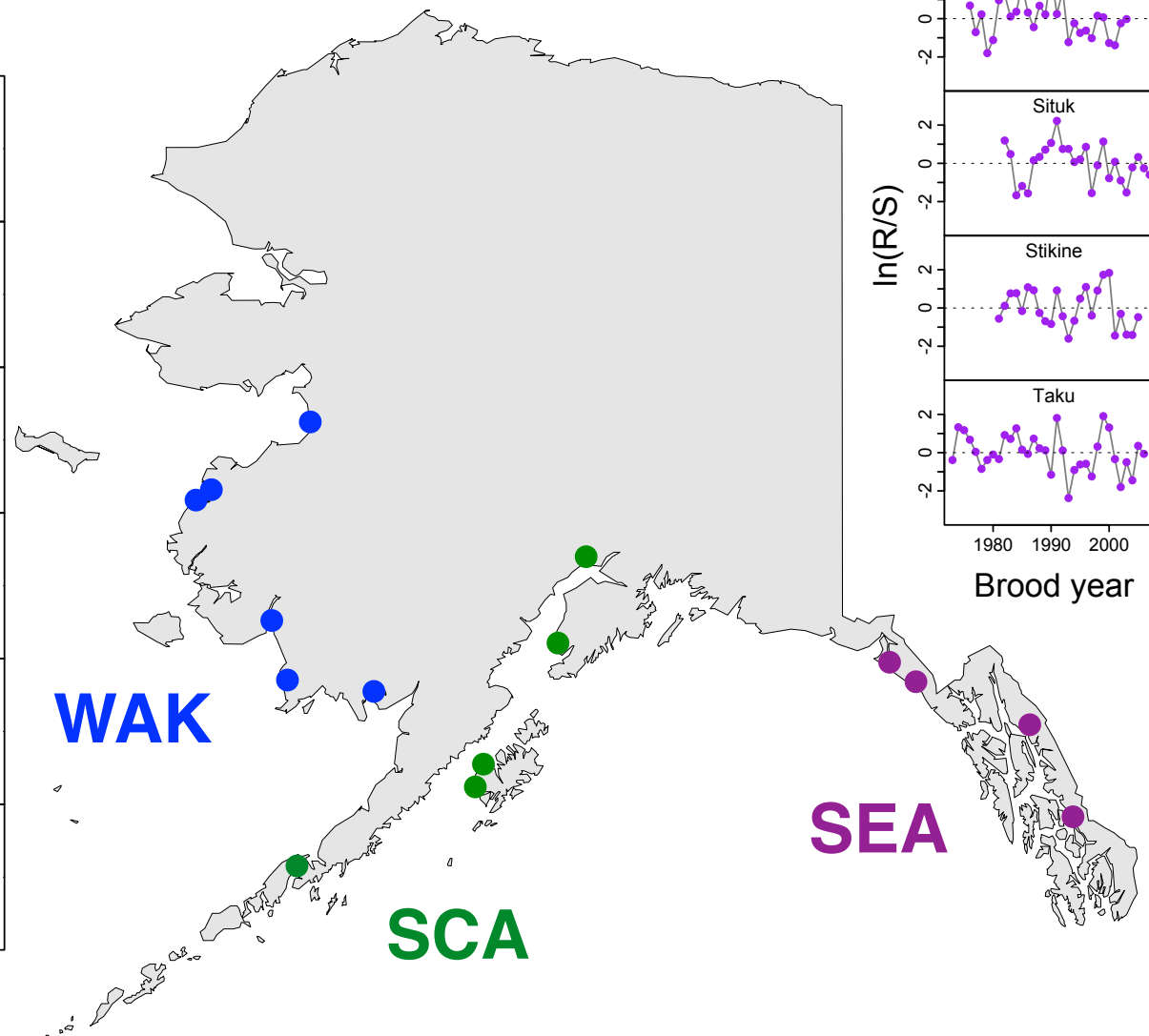
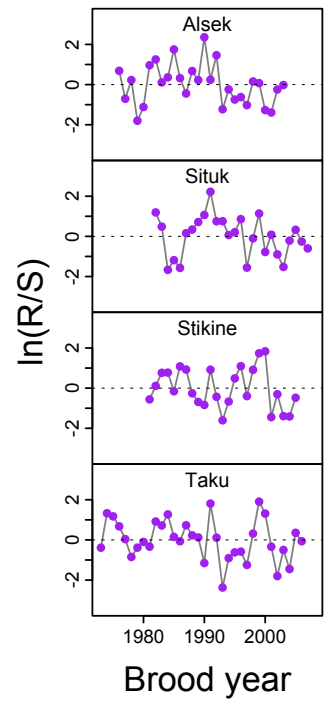
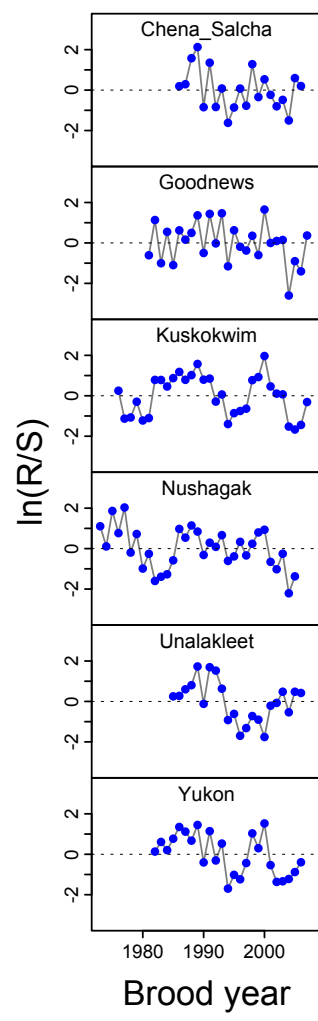
Citation: Ohlberger J., M. D. Scheuerell, and D. E. Schindler. 2016. Population coherence and environmental impacts across spatial scales: a case study of Chinook salmon. *Ecosphere* 7(4):e01333. [10.1002/ecs2.1333](https://doi.org/10.1002/ecs2.1333)

Arctic-Yukon-Kuskokwim salmon

- 1) Poor returns of Chinook & chum in AYK region over past decade have led to severe restrictions on commercial & subsistence harvest
- 2) This has also led to repeated disaster declarations by the state and federal governments (**nobody** fished in 2012!).
- 3) In response, native regional organizations, state and federal agencies formed an innovative partnership to cooperatively address problems (AYK SSI)

The data

- 15 stocks of Alaskan Chinook
- Brood years 1976-2005
- Response is $\log(\text{Recruits}/\text{Spawner})$
- Covariates lagged by 1-5 years

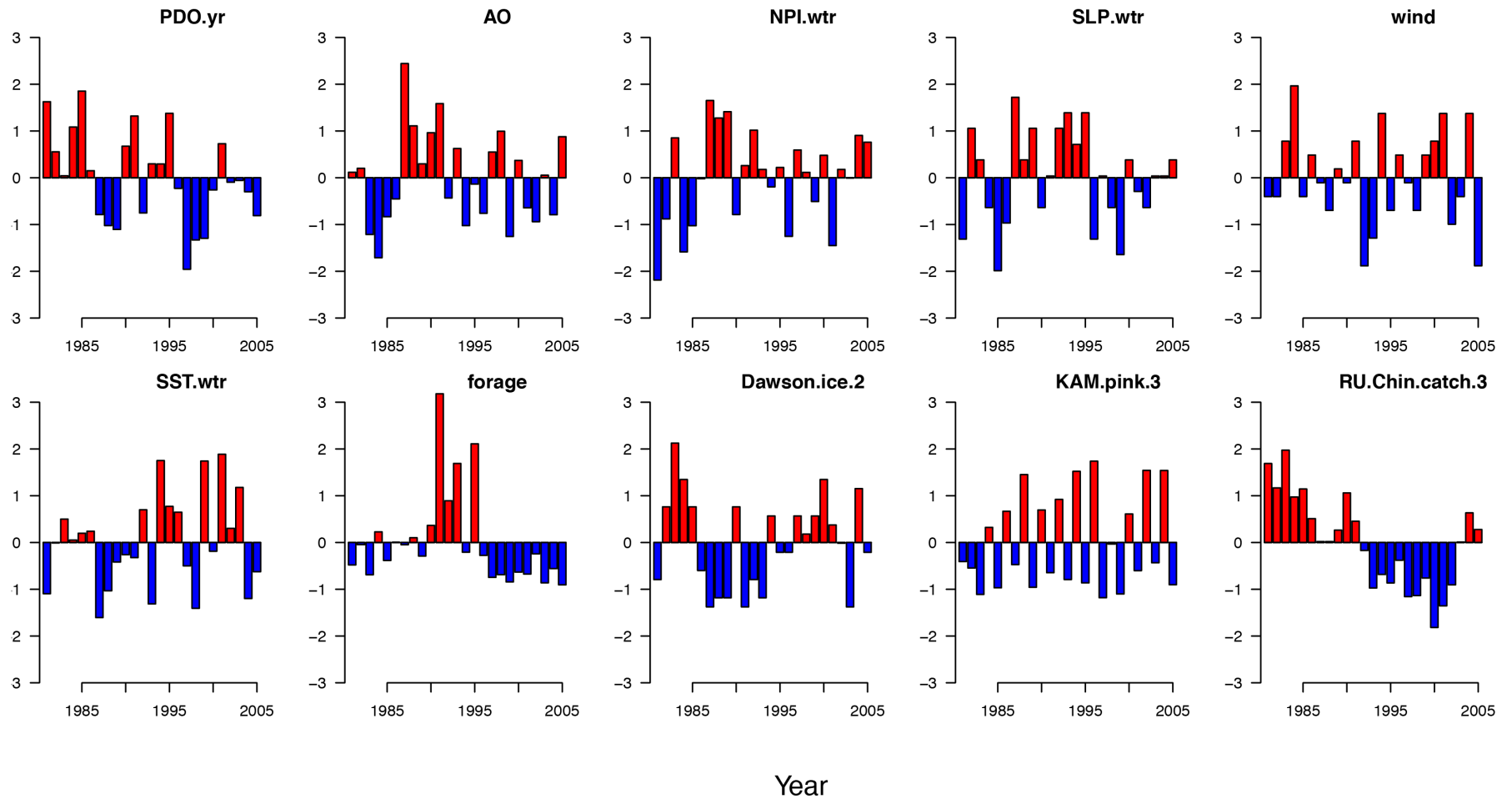


Environmental indicators

Table 2. Environmental indicators used in the DFA. Indicators were hypothesized to affect Chinook salmon at specific ages, and were lagged by the appropriate year, relative to brood year.

Indicator	Acronym	Year					Source
		1	2	3	4	5	
Air temperature on land	TEMP	X	X				1, 2
Ice-out dates	ICE	X	X				3
Sea level pressure	SLP		X				2
Sea surface temperature	SST		X				2
Strong winds index	SWI		X				2
Arctic Oscillation	AO		X				2
Pacific Decadal Oscillation	PDO		X				2
North Pacific Index	NPI		X				2
North Pacific Gyre Oscillation	NPGO		X				4
Bogoslof region pollock biomass	BOG		X				5
Russian Chinook catch	RUS			X	X	X	5
Kamchatka pink abundance	KAM			X			6
BSAI Chinook bycatch	BSAI		X	X	X		6

Examples of some indicators



The analysis

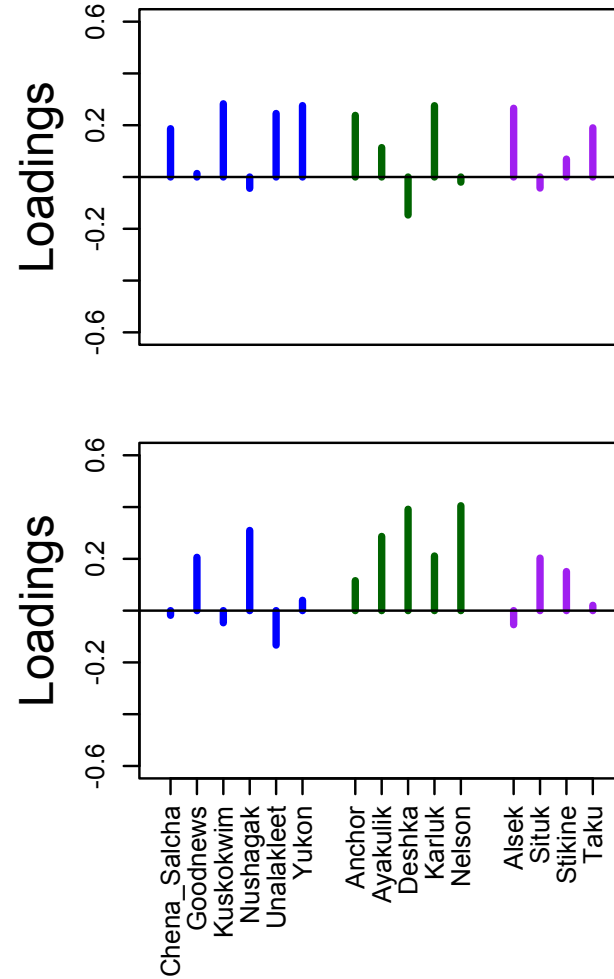
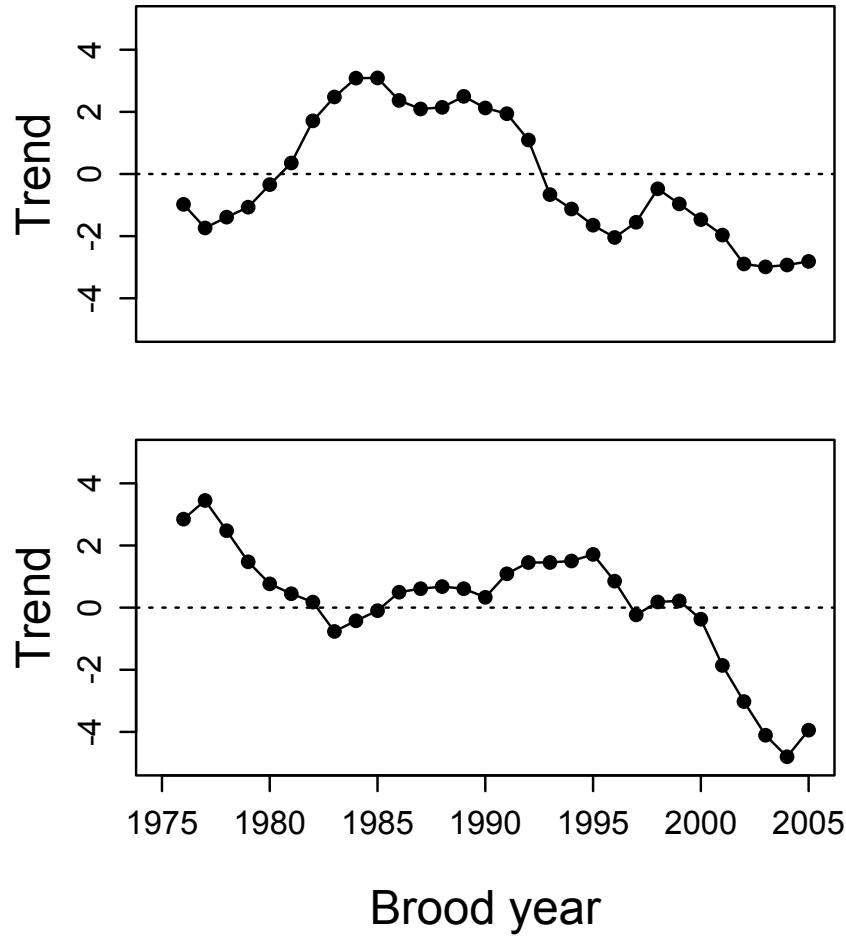
- Varied the number of states/trends from 1-3
- Varied forms of R to try:
 - 1) Diagonal and equal,
 - 2) Diagonal and unequal,
 - 3) Equal variances and covariances.
- Used AICc to select “best” model

Statewide results

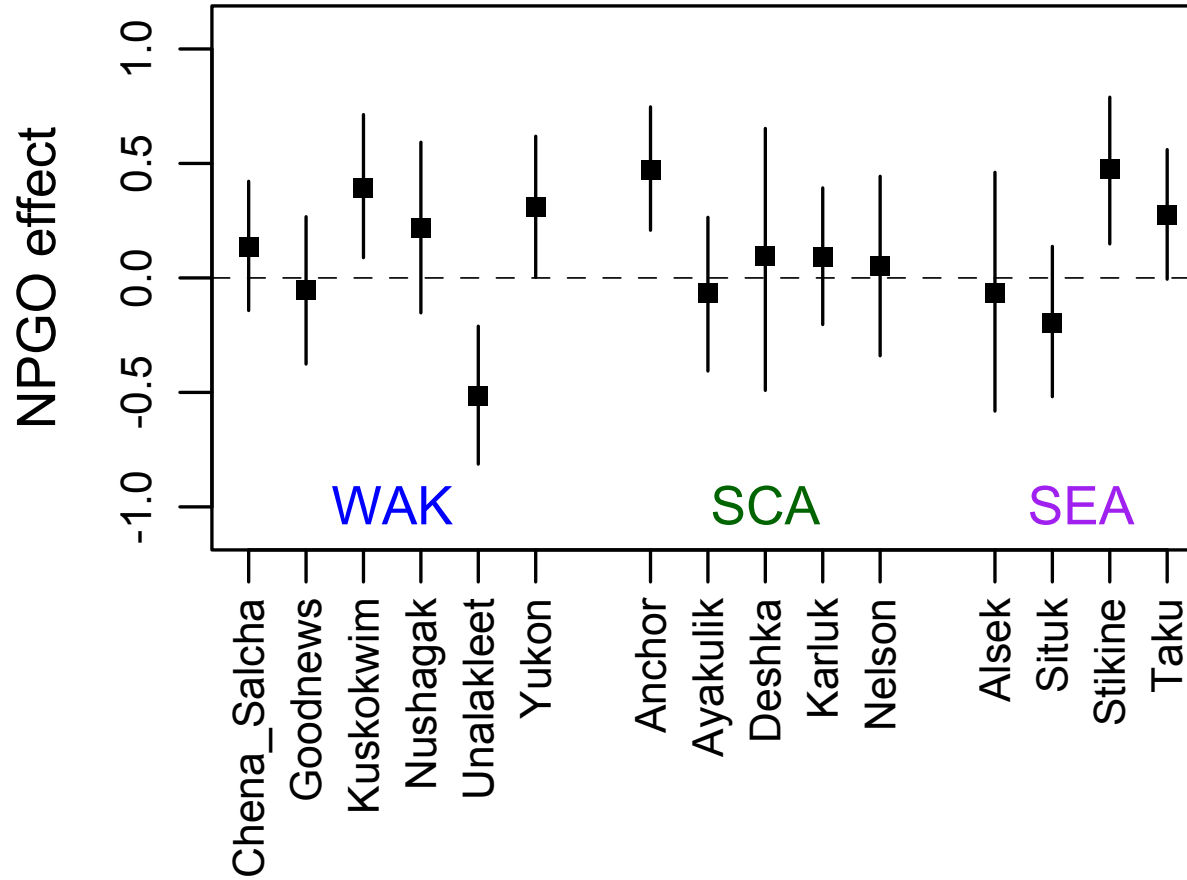
The most parsimonious model had:

- 1) 2 common trends
- 2) North Pacific Gyre Oscillation

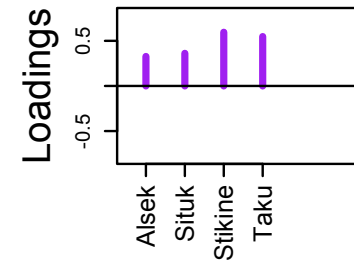
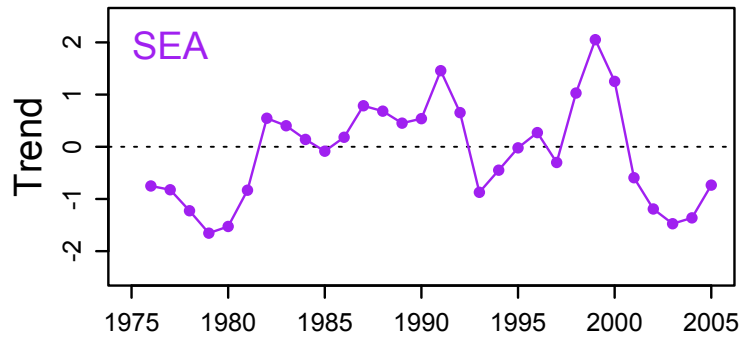
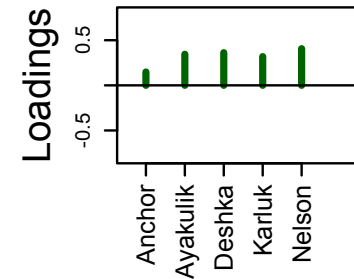
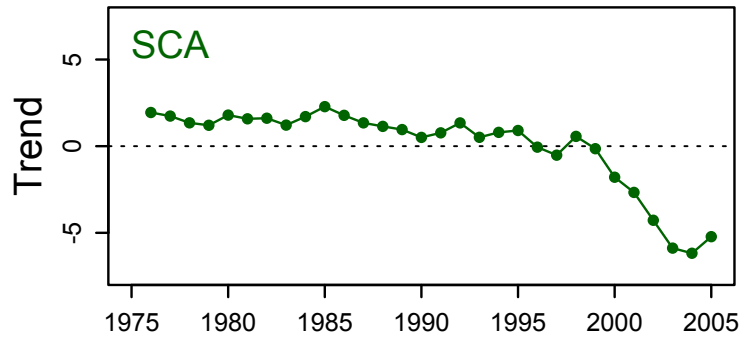
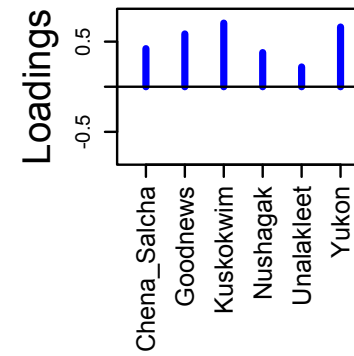
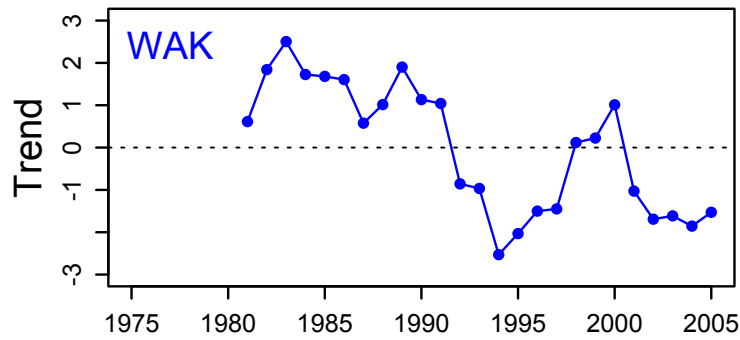
Statewide results



Statewide results – covariate effect



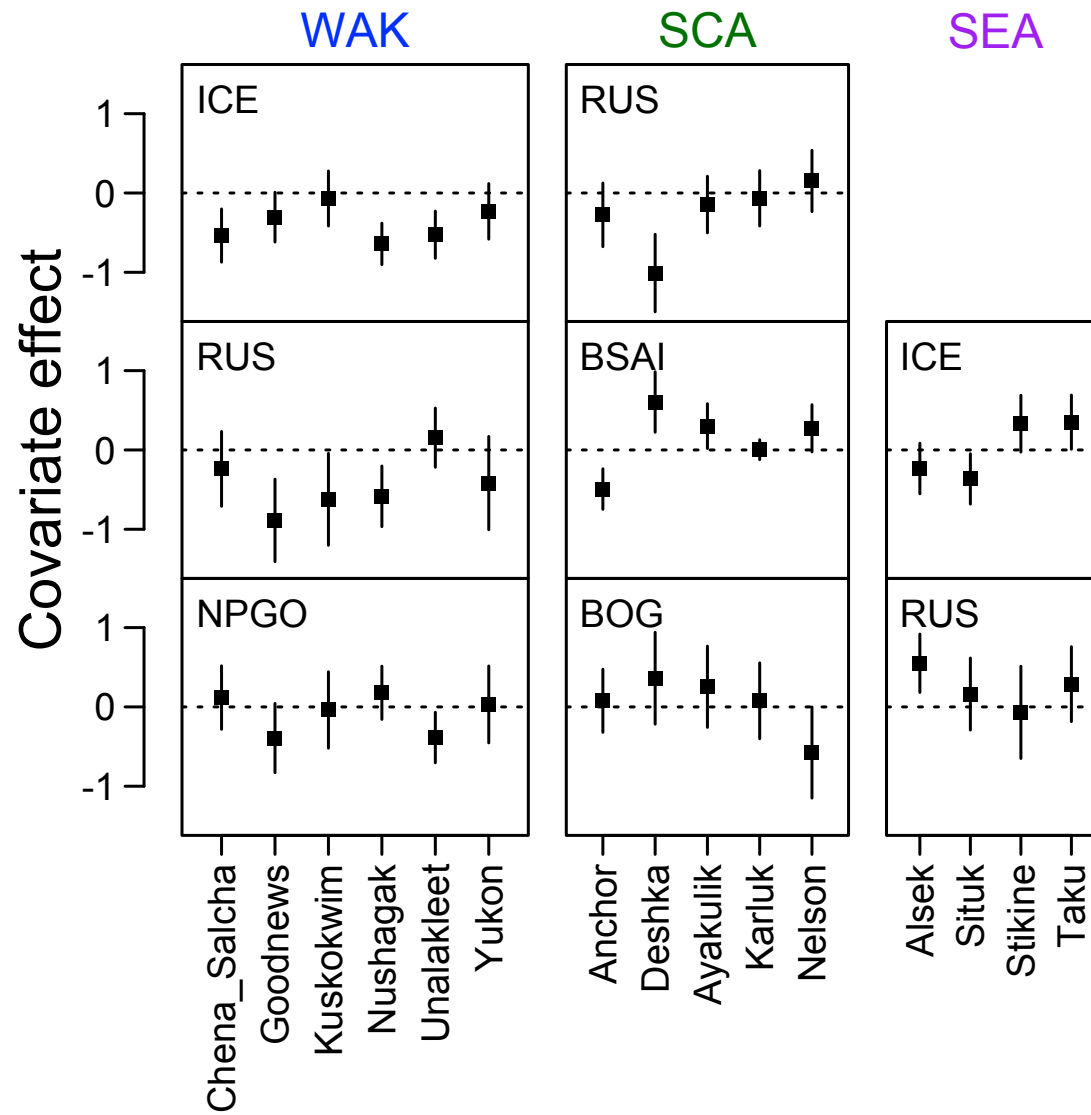
Regional results



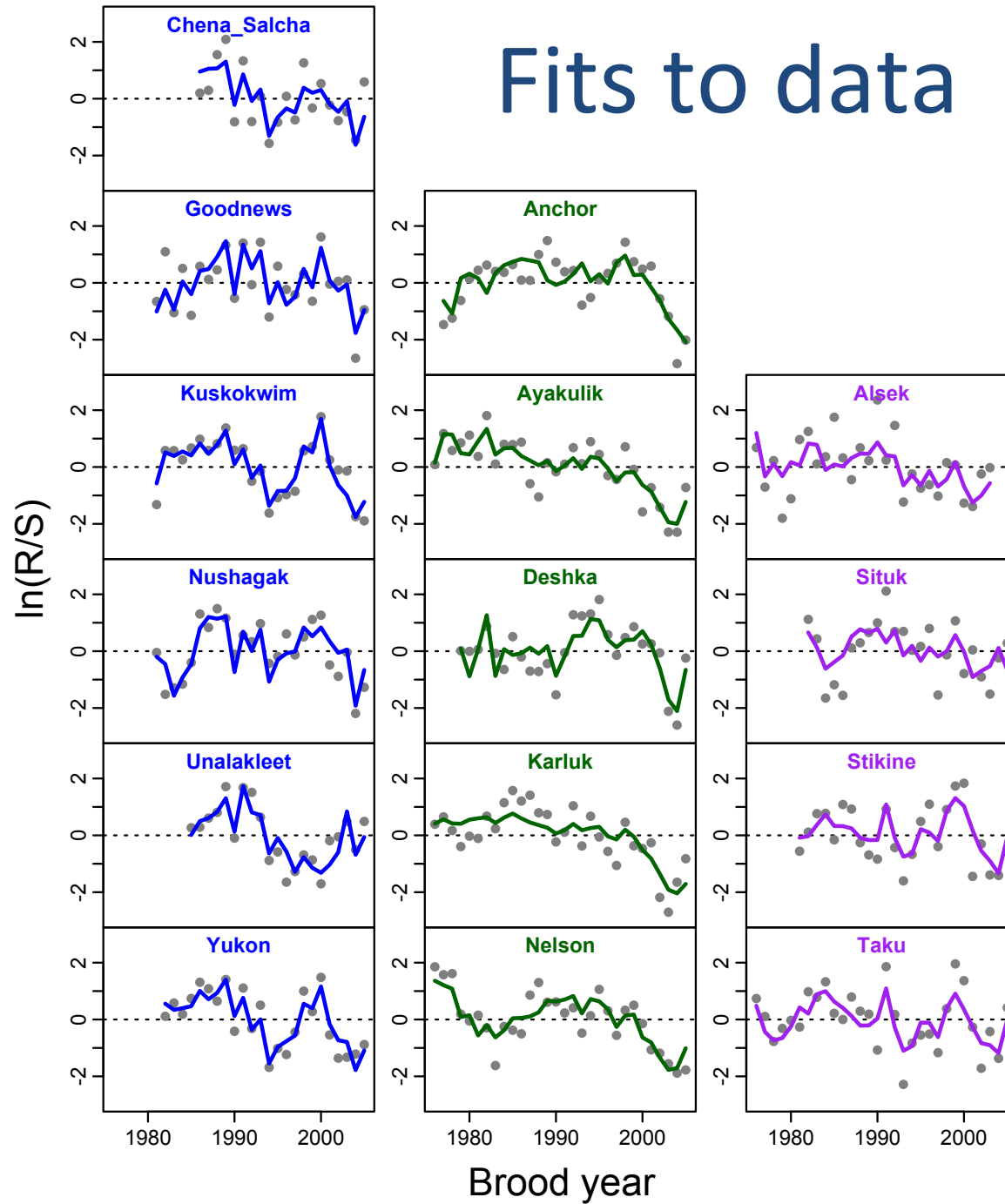
Brood year

Stocks

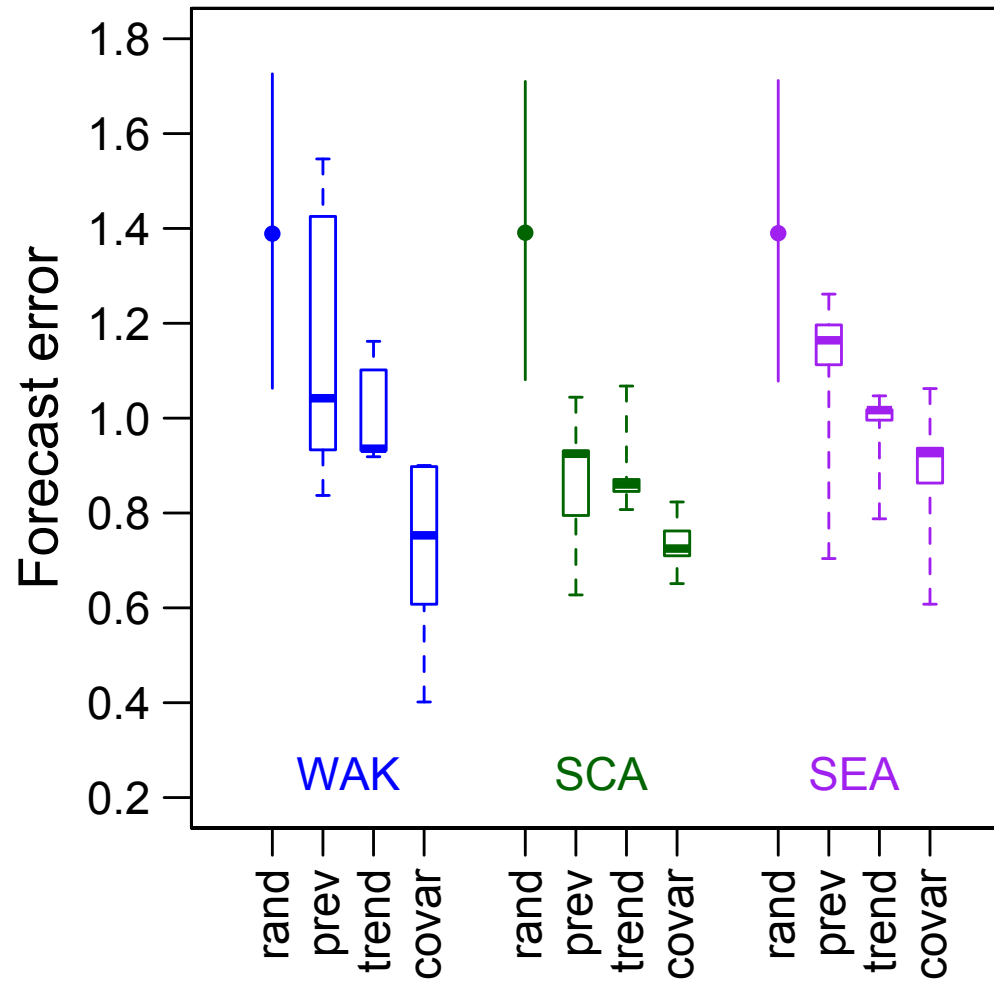
Regional results – covariate effect



Fits to data



Forecasting ability



Topics for lab

- Fitting DFA models without covariates
- Fitting DFA models with covariates
- Doing factor rotations