Intervention models and standardized residuals for perturbation analysis

Eli Holmes

FISH 507 – Applied Time Series Analysis

12 March 2019

Big question in the finance world

What is the effect of advertising on sales?

©Budweiser

Anheuser-Busch
spends $35 million/yr
on Super Bowl ads

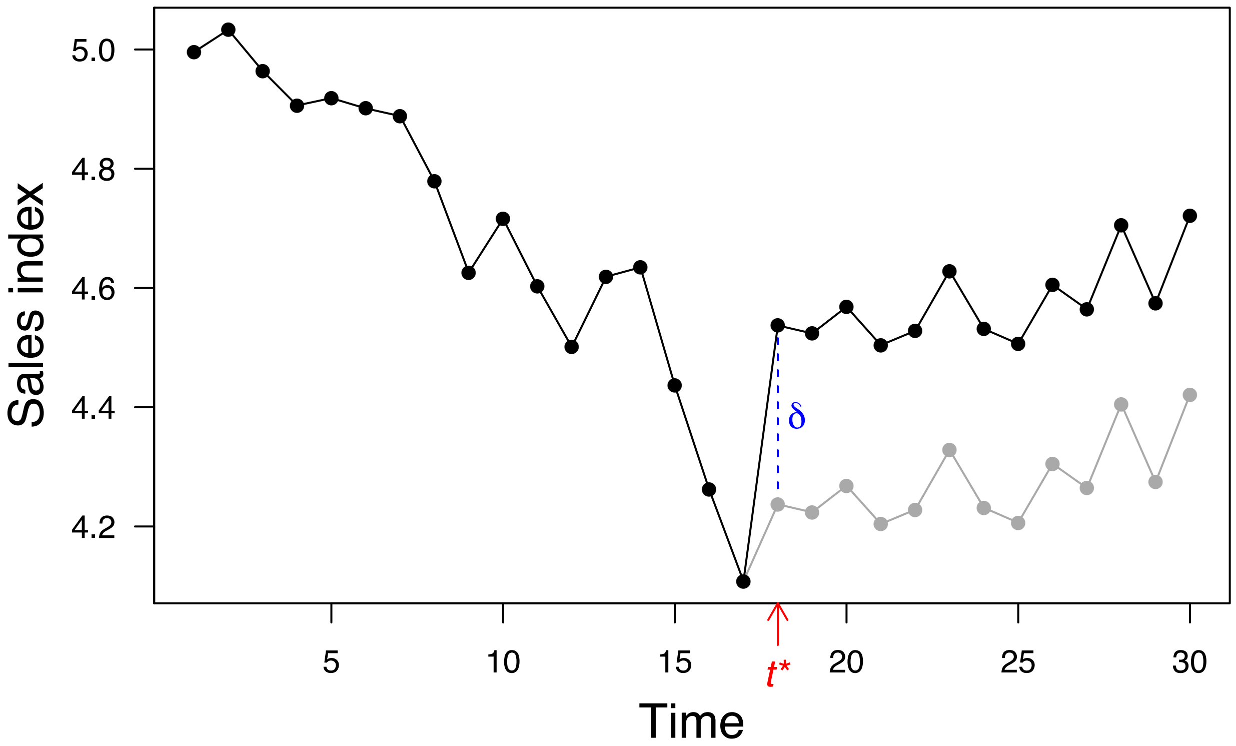
↓

$95 million/yr in revenue
(170% return!)

How do they know this?

Hartman *et al* (2015)

How much did sales change?



Model from finance world

Sales

State equation

$$x_t = x_{t-1} + \delta I_{t-h} + w_t \qquad w_t \sim \mathrm{N}(0,q)$$

Advertising effect

Indicator function

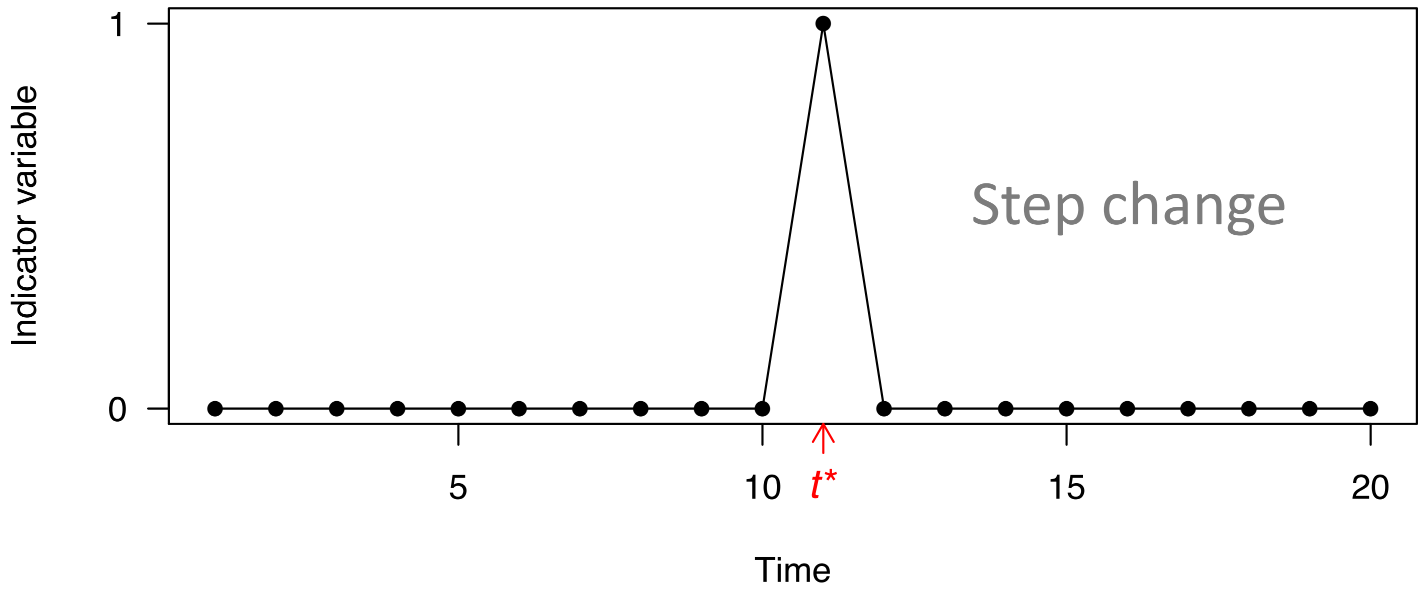
$$I_{t-h} = \begin{cases} 0 & \text{if } t-h \neq \text{event} \\ 1 & \text{if } t-h = \text{event} \end{cases}$$

Hamilton (1989), West & Harrison (1997)

Model from finance world

State equation

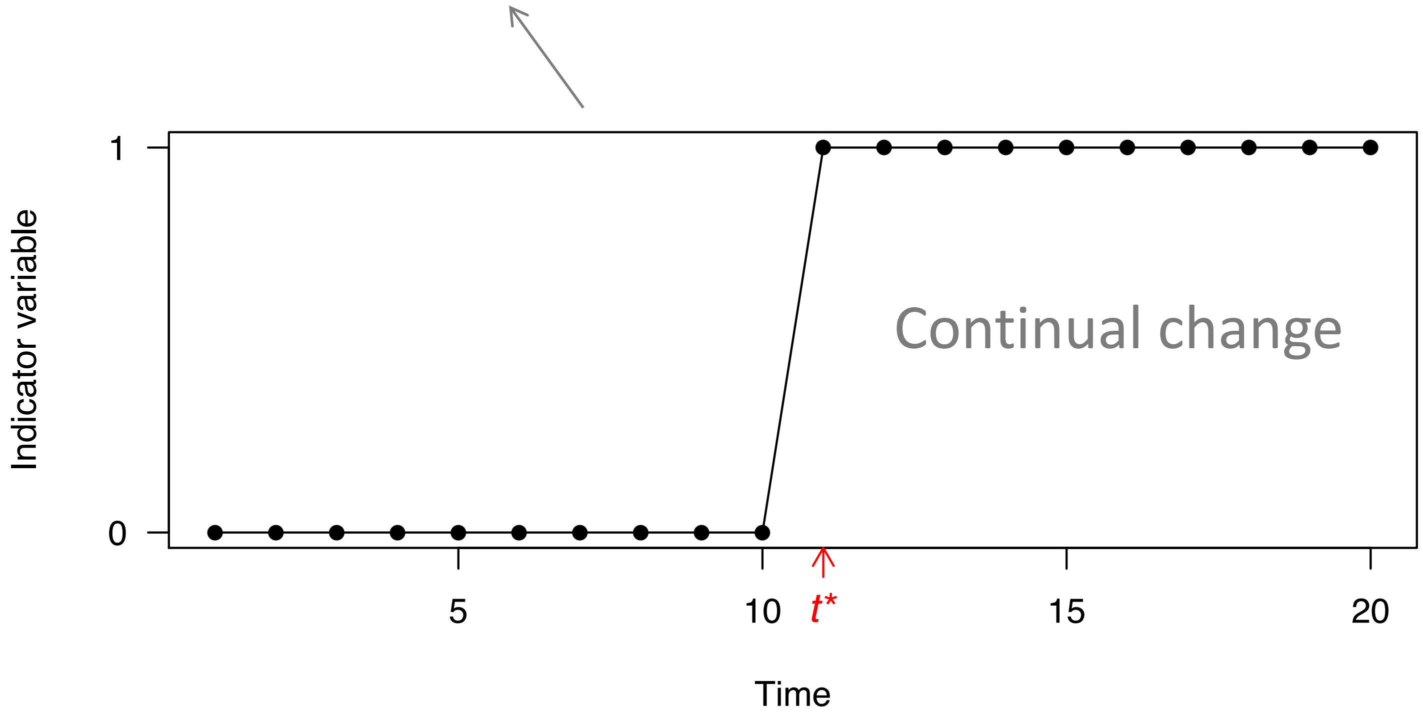
$$x_t = x_{t-1} + \delta I_{t-h} + w_t \qquad w_t \sim \mathrm{N}(0,q)$$



Model from finance world

State equation

$$x_t = x_{t-1} + \delta I_{t-h} + w_t \qquad w_t \sim \mathrm{N}(0, q)$$



Model from finance world

Sales

State equation

$$x_t = x_{t-1} + \delta E_{t-h} + w_t \qquad w_t \sim \mathrm{N}(0, q)$$

Advertising effect

Advertising expense

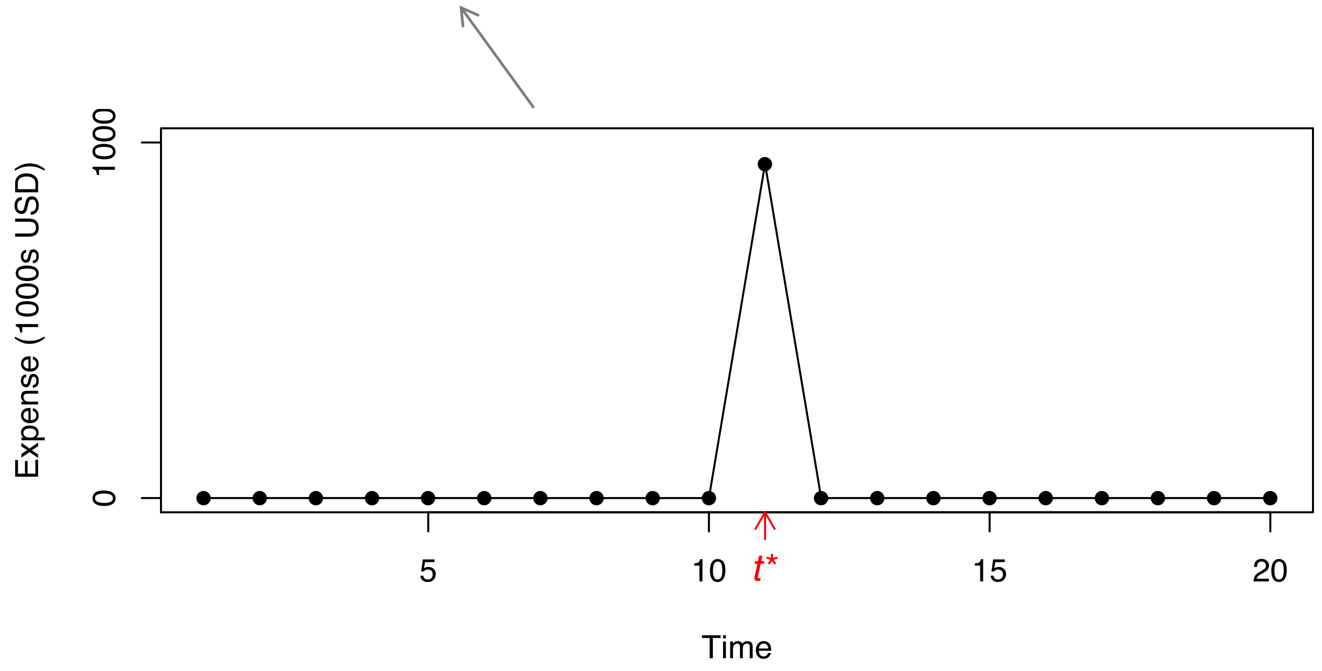
$$E_{t-h} = \begin{cases} 0 & \text{if } t-h \neq \text{event} \\ E_{t-h} & \text{if } t-h = \text{event} \end{cases}$$

Hamilton (1989), West & Harrison (1997)

Model from finance world

State equation

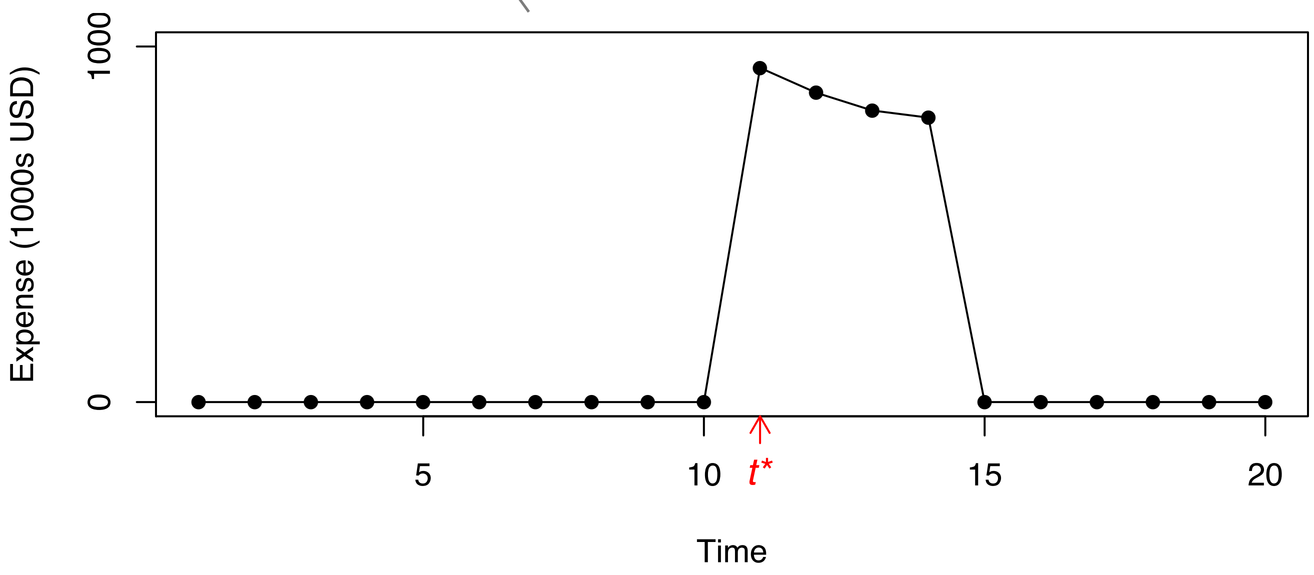
$$x_t = x_{t-1} + \delta E_{t-h} + w_t \qquad w_t \sim \mathrm{N}(0,q)$$



Model from finance world

State equation

$$x_t = x_{t-1} + \delta E_{t-h} + w_t \qquad w_t \sim \mathrm{N}(0,q)$$



What about interventions in obs?

- It is entirely possible for their to be a change (intervention) in the observations
- Field ecology (fisheries, ornithology)
- Laboratory (microscopy, genetics, chemistry)

Model for change in observation

State equation

$$x_t = x_{t-1} + w_t \qquad w_{i,t} \sim \mathrm{N}(0, q_i)$$

Observation equation

$$y_t = x_t + \delta I_{t-h} + v_t \qquad v_t \sim \mathrm{N}(0, r)$$

Intervention effect (δ) — *Indicator function* (I_{t-h})

Ecology and Evolution

Open Access

Analyzing large-scale conservation interventions with Bayesian hierarchical models: a case study of supplementing threatened Pacific salmon

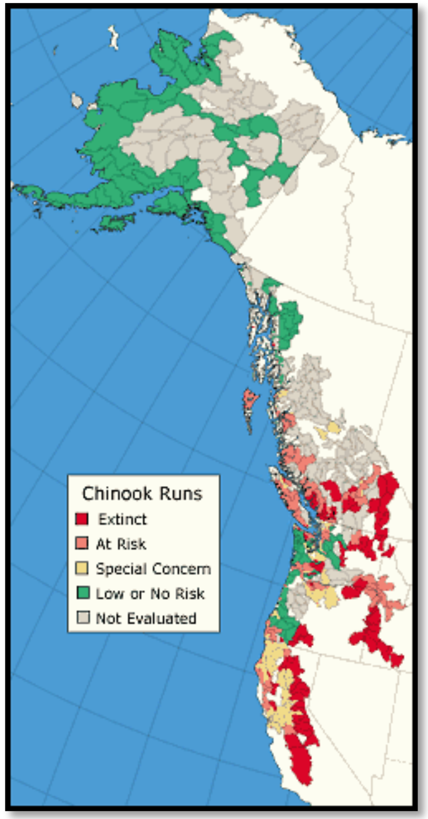
Mark D. Scheuerell[1], Eric R. Buhle[1], Brice X. Semmens[2], Michael J. Ford[3], Tom Cooney[3] & Richard W. Carmichael[4]

[1]Fish Ecology Division, Northwest Fisheries Science Center, National Marine Fisheries Service, National Oceanic and Atmospheric Administration, Seattle, Washington 98112
[2]Scripps Institute of Oceanography, University of California, San Diego, La Jolla, California 92093
[3]Conservation Biology Division, Northwest Fisheries Science Center, National Marine Fisheries Service, National Oceanic and Atmospheric Administration, Seattle, Washington 98112
[4]Northeast-Central Oregon Research and Monitoring, Oregon Department of Fish and Wildlife, Eastern Oregon University, La Grande, Oregon 97850

***Ecology and Evolution* 2015; 5(10): 2115–2125**

doi: 10.1002/ece3.1509

The salmon story



Source: State of the Salmon

- Major declines in populations across the continental U.S. & southern Canada
- Evolutionary Significant Units (ESUs) form basis for conservation & management
- 28/52 ESUs listed as *threatened* or *endangered* under U.S. Endangered Species Act
- Human (eg, dams, harvest) & natural (climate) causes have contributed to declines
- Big money business ($4 billion per decade)

Adverse effects of hatcheries

Growing evidence that hatchery fish have reduced fitness & adverse demographic effects

(eg, Araki et al. 2007, Buhle et al. 2009, Christie et al. 2014)

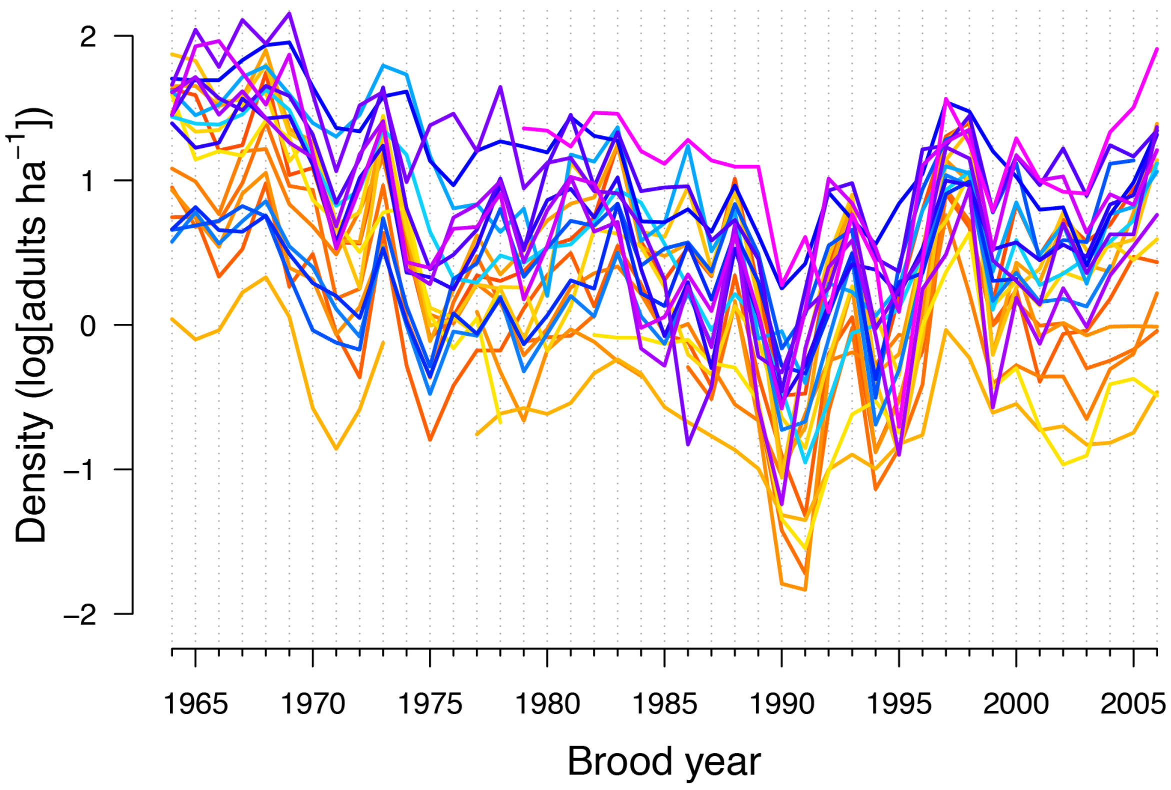
The big picture

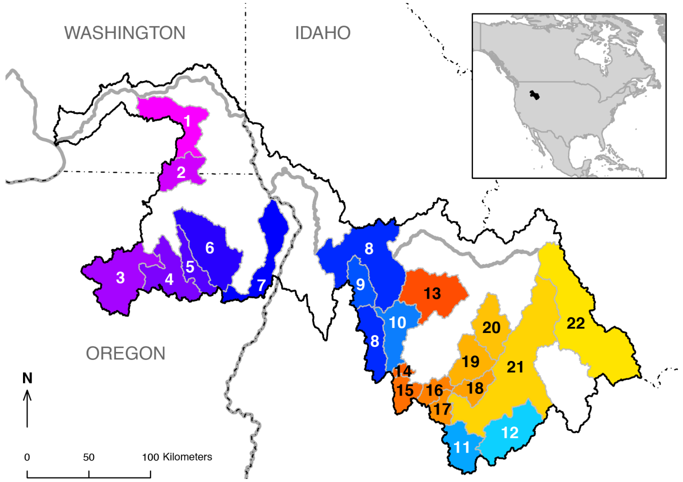
Question

What is the effect of hatchery supplementation on Snake River spring/summer Chinook salmon at

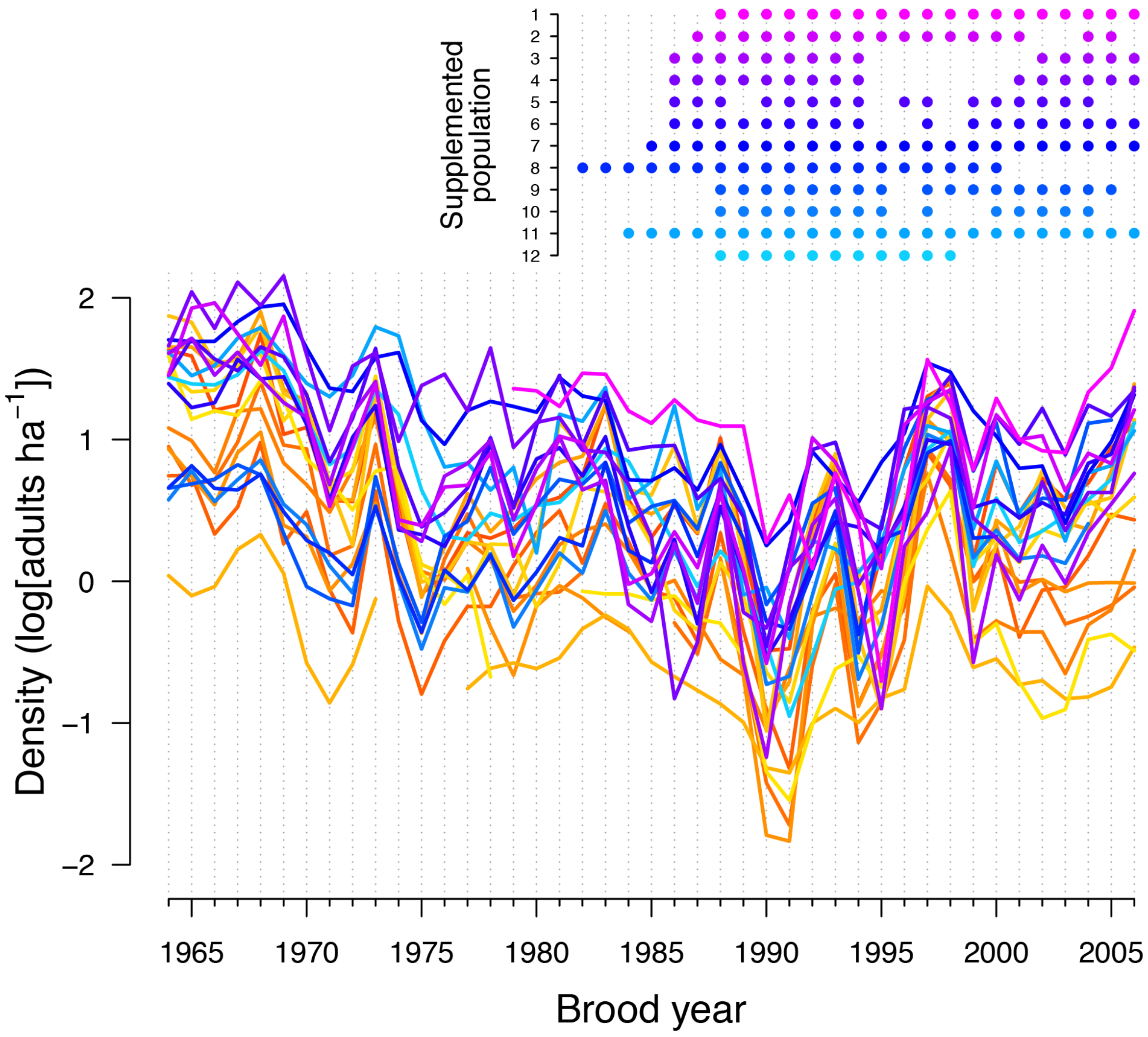
1) population level, and
2) broader ESU scale?

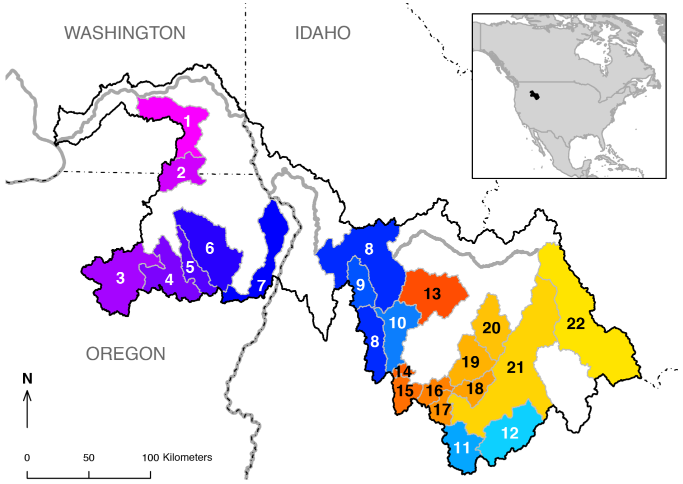
Time series of spawner density





Time series of supplementation





Model for supplementation

True density

State equation

$$x_{i,t} = x_{i,t-1} + \alpha_t + \delta_i I_{i,t-h} + w_{i,t} \qquad w_{i,t} \sim \mathrm{N}(0, q_i)$$

Common year effect

Model for supplementation

True density

<u>State equation</u>

$$x_{i,t} = x_{i,t-1} + \alpha_t + \delta_i I_{i,t-h} + w_{i,t} \qquad w_{i,t} \sim \mathrm{N}(0, q_i)$$

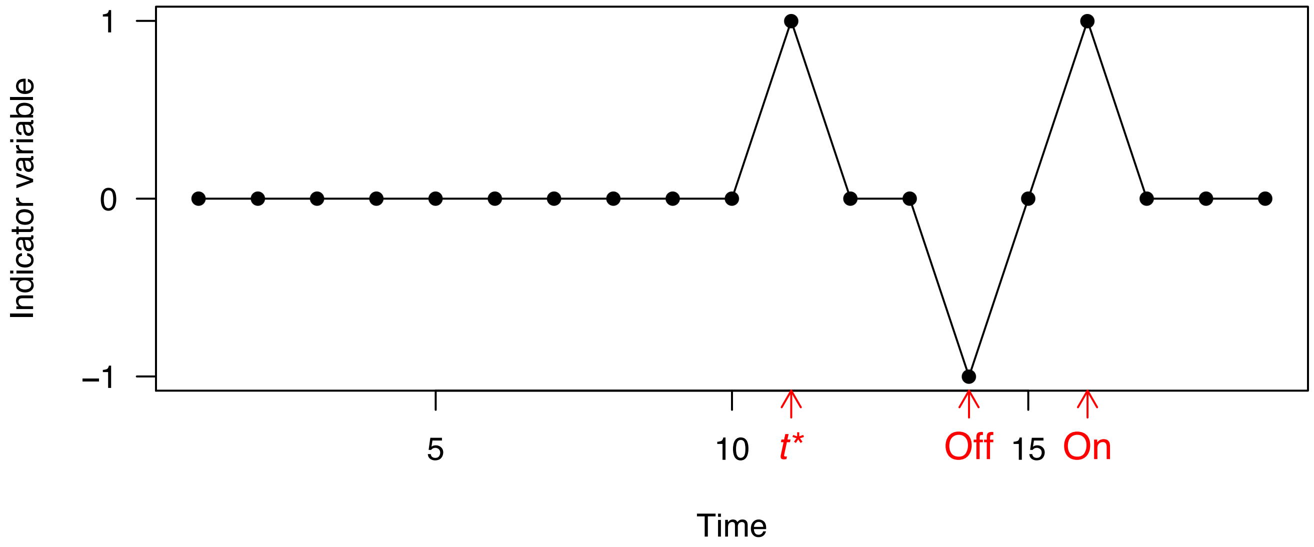
Supplementation effect

Indicator function

Model for supplementation

State equation

$$x_{i,t} = x_{i,t-1} + \alpha_t + \delta_i I_{i,t-h} + w_{i,t} \qquad w_{i,t} \sim \mathrm{N}(0, q_i)$$



Model for supplementation

True density

State equation

$$x_{i,t} = x_{i,t-1} + \alpha_t + \delta_i I_{i,t-h} + w_{i,t} \qquad w_{i,t} \sim \mathrm{N}(0, q_i)$$

Supplementation effect

Indicator function

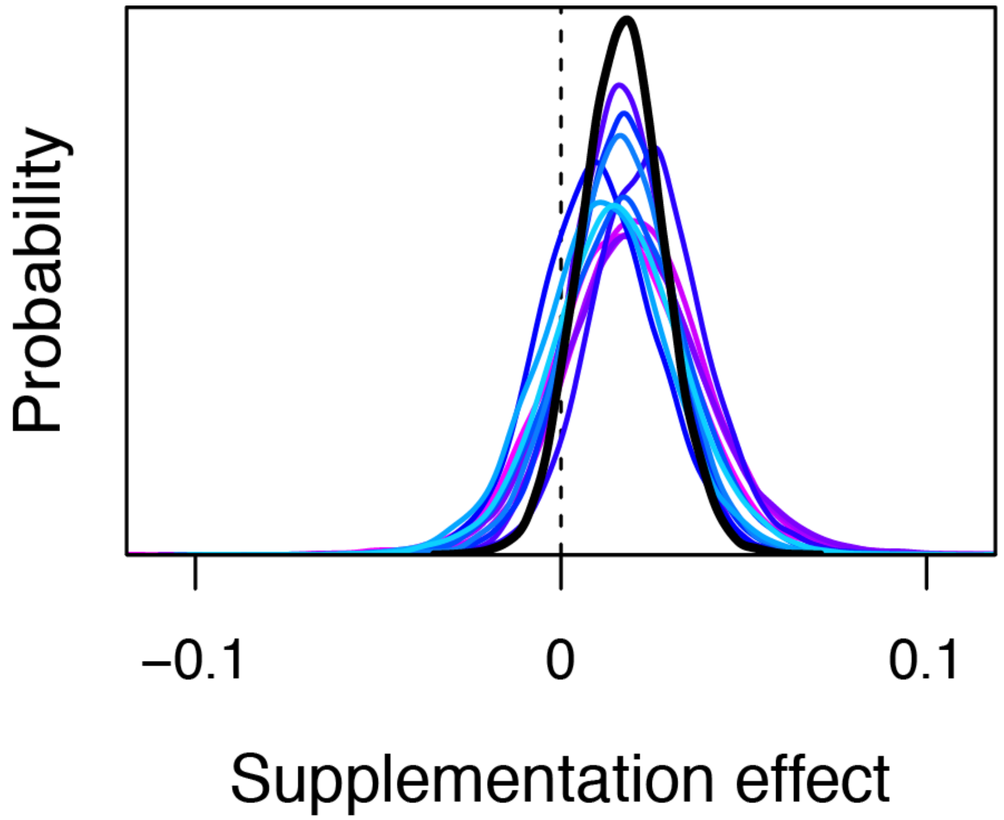
Observation equation

$$y_t = x_t + v_t \qquad v_t \sim \mathrm{N}(0, r)$$

Observed density

Distribution of intervention sizes

ESU-level:	Mean	95% CI	Pr(+)
	0.033	(-0.077, 0.15)	0.73



Summary

- Intervention models are used in many fields
- Intervention models can take many forms

standardized residuals

Detection of outliers and structural breaks using standardized residuals

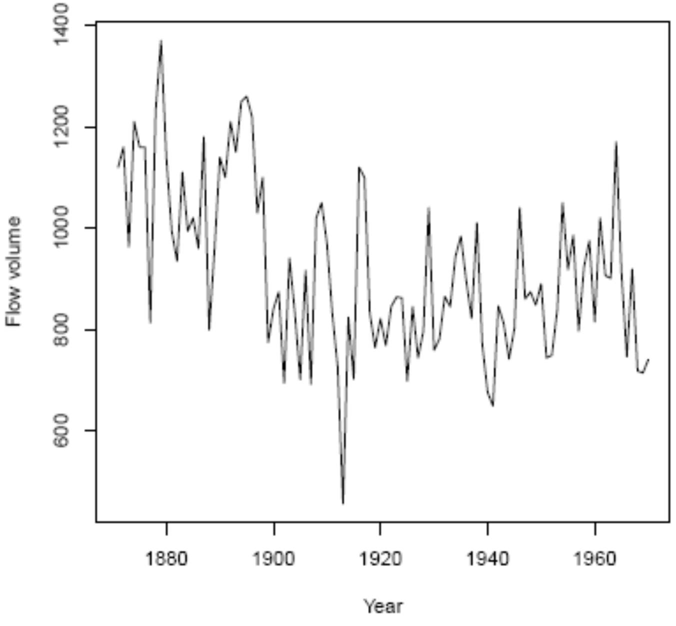
See the chapter on outlier and structural break detection in the HWS (MARSS User Guide)

de Jong, P. and Penzer, J. 1998. Diagnosing shocks in time series. Journal of the American Statistical Association 93:796-806.

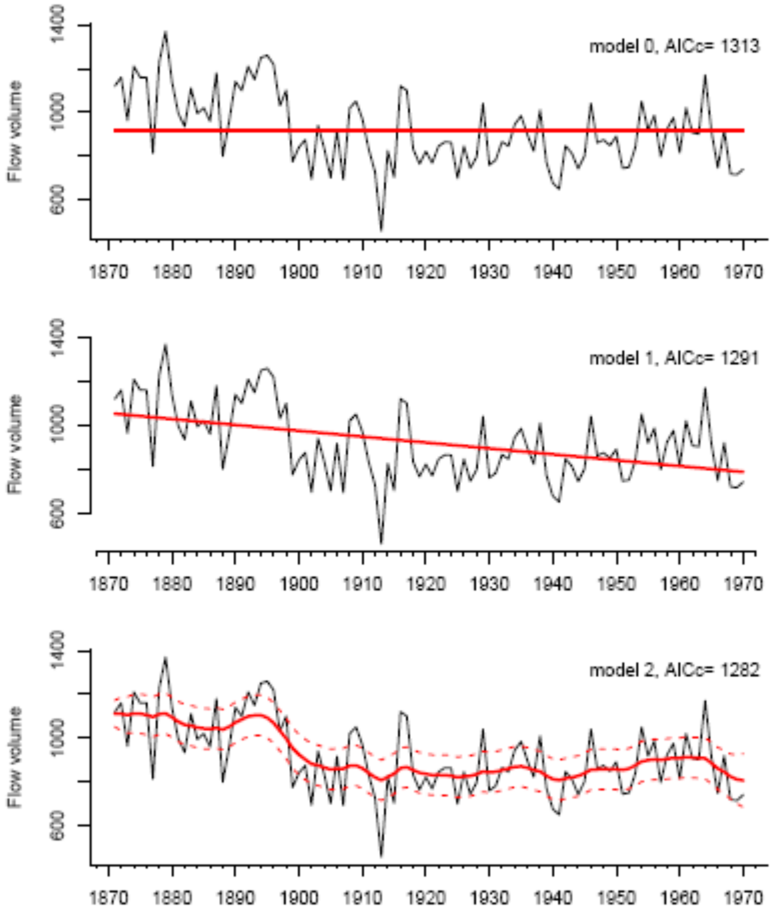
Durbin and Koopman. 2012. Time series analysis by state-space methods. Chapter 2, Section 12

Back to the Nile River data

River flow by year



Three different models



Observation outlier detection

Observation outlier: observation (data) at time t is different than what you would expect given the model.

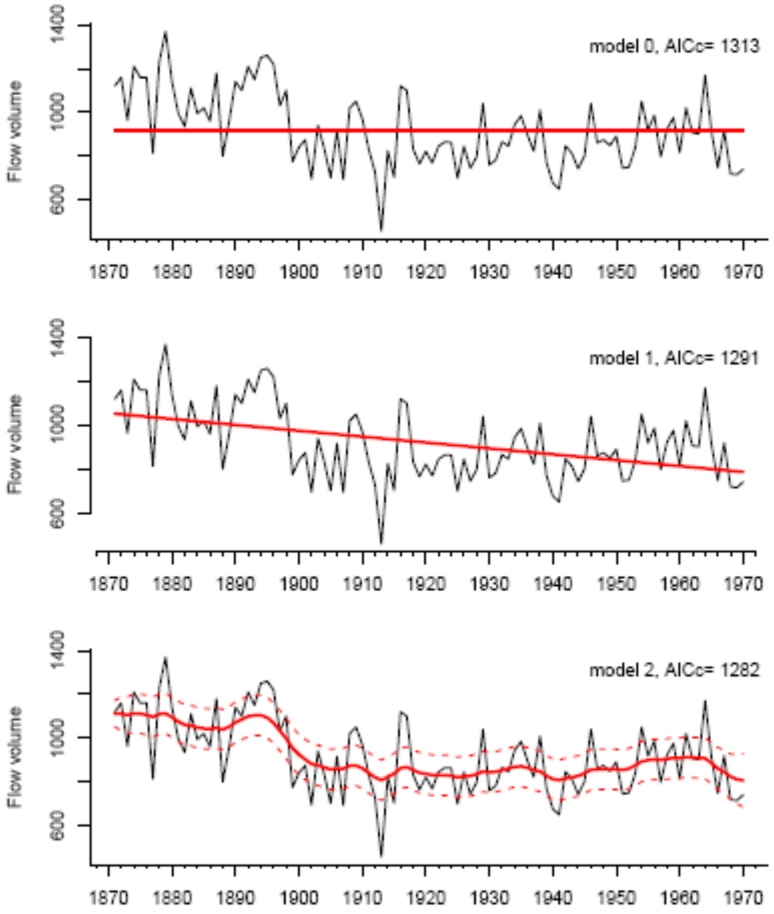
obs. residual = data – fitted value

$$\hat{v}_t = y_t - \hat{y}_{t|T}$$

$$e_t = \frac{1}{\sqrt{\mathrm{var}(\hat{v}_t)}}\hat{v}_t$$

we standardize by the estimated variance and get a t-distributed standardized residual

Three different models



This idea hinges on v(t) being normal so that means it hinges on the model being able to fit the data (= put a line through the data)

Observation residual in the context of state-space models

$$\hat{v}_t = y_t - \hat{y}_{t|T}$$

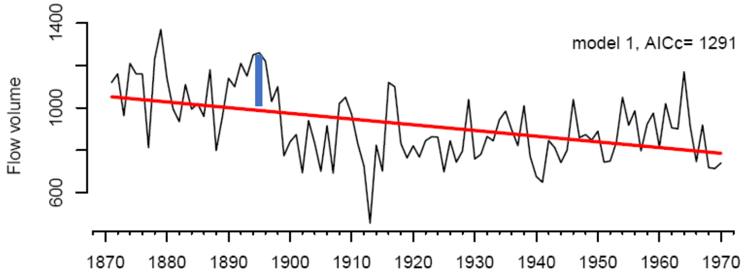
$$e_t = \frac{1}{\sqrt{\mathrm{var}(\hat{v}_t)}} \hat{v}_t$$

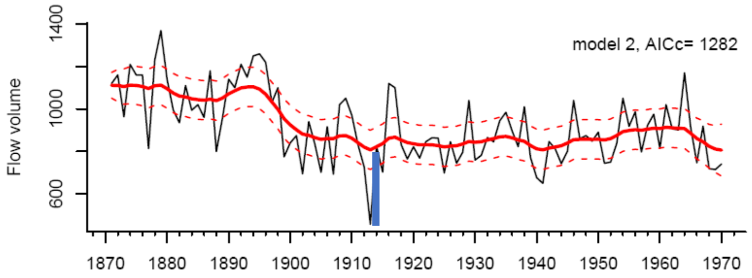
obs. residual = data – fitted value

$$\hat{y}_{t|T} = \hat{Z}\tilde{x}_{t|T} + \hat{a}$$

you need to standardize by the variance of that, which is a bit hairy but algorithms for computing it are worked out.

for a linear regression, 'fitted y' is easy.

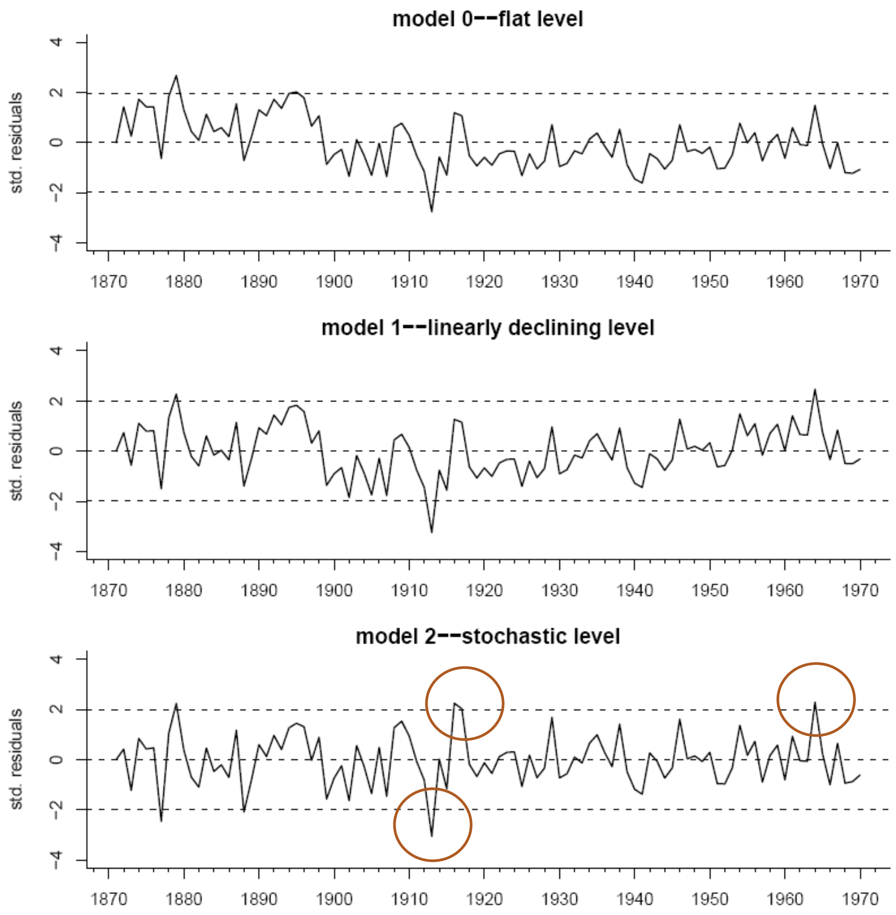




for a state-space model, there isn't one 'fitted y'. 'fitted y' has a distribution.

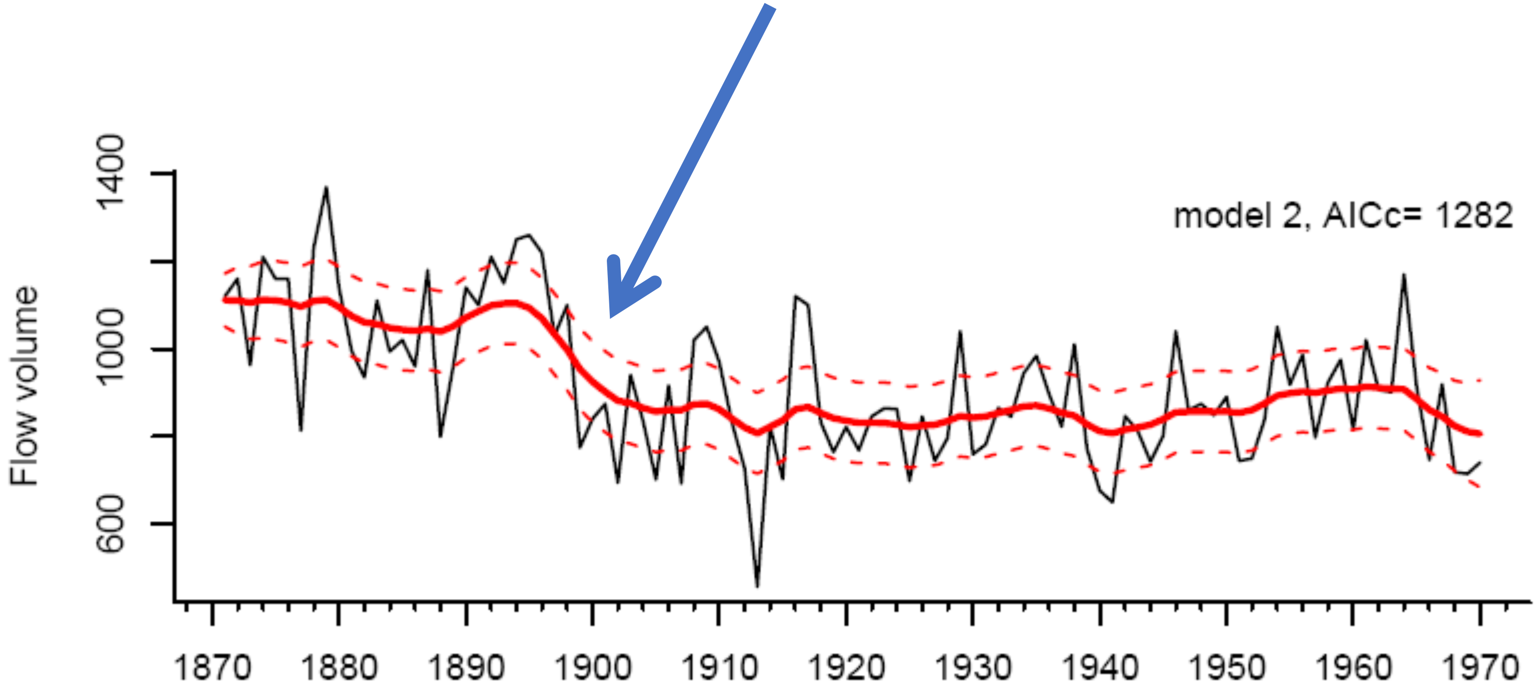
```
resids.0=residuals(kem.0)$std.residuals
resids.1=residuals(kem.1)$std.residuals
resids.2=residuals(kem.2)$std.residuals
```

Note, the standard concerns regarding setting test levels for multiple tests exist



“Structural break detection” aka testing state outliers

Idea is to test whether observed changes in the stochastic state (in this example level) were more unusual than you would expect given the estimated MAR model for the state.



"Structural break detection" aka testing state outliers

State outlier: estimated state at time t+1 is different than what you would expect given the model.

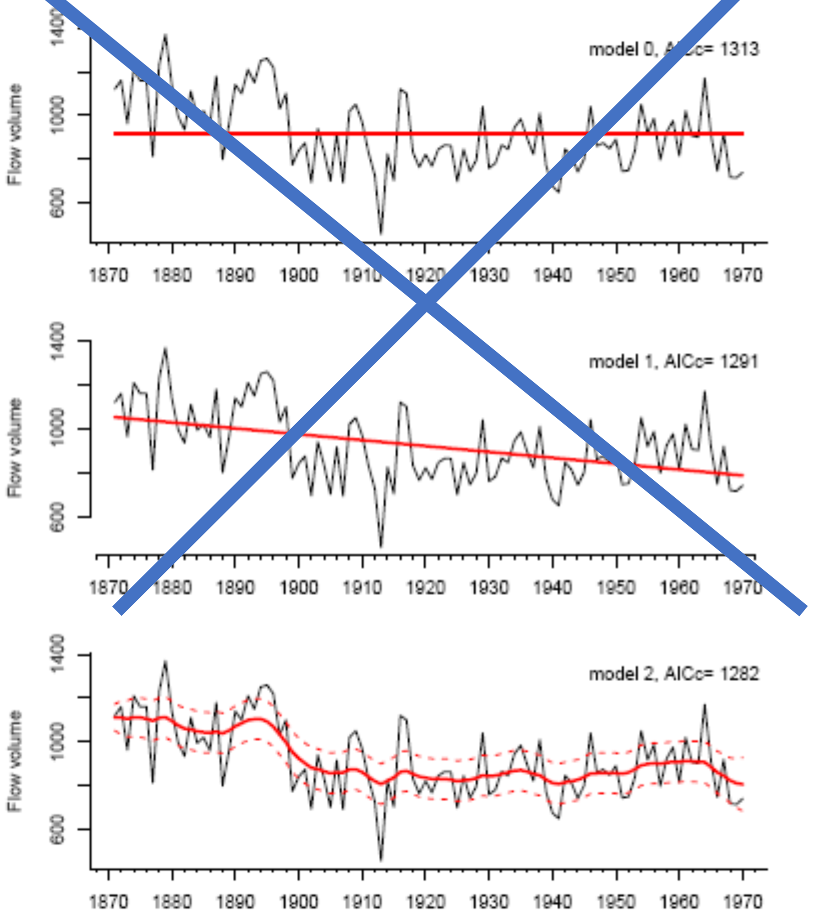
state. residual =

$$\hat{w}_t = \tilde{x}_{t|T} - \tilde{x}_{t-1|T}$$

$$f_t = \frac{1}{\sqrt{\mathrm{var}(\hat{w}_t)}}\hat{w}_t$$

we standardize by the estimated variance and get a t-distributed standardized residual

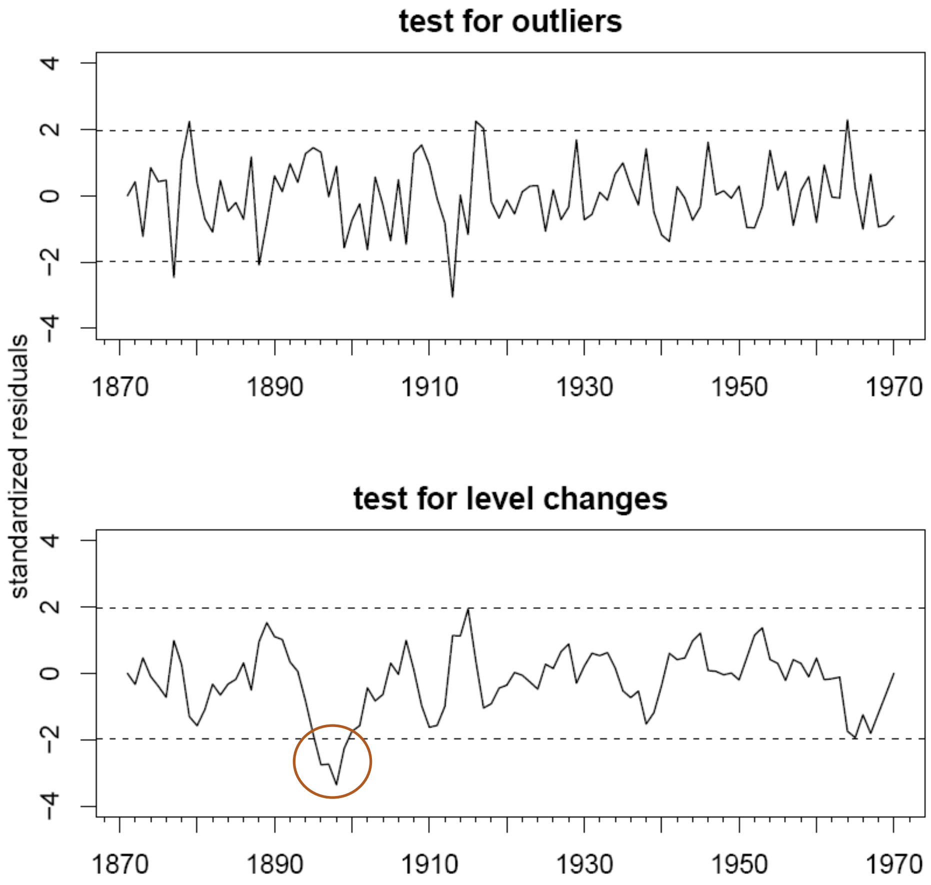
Three different models



Again this idea hinges on w(t) being normal so that means it hinges on the model being able to fit the data (= put a line through the data)

```
resids.0=residuals(kem.0)$std.residuals
resids.1=residuals(kem.1)$std.residuals
resids.2=residuals(kem.2)$std.residuals
```

Note, the standard concerns regarding setting test levels for multiple tests exist



Summary

- Residual analysis is a diagnostic tool to look for observation or state outliers and evidence of times when the underlying model is violated, but there is no cause involved.
- Intervention analysis is more suited to a mechanistic analysis of changes/breaks that may or may not have occurred.