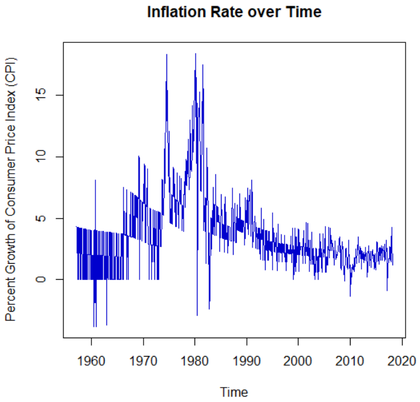
Developing a Bayesian procedure to detect breakpoints in time series

Kathryn Haglich, Sarah Neitzel, Amy Pitts
Mentor: Jeff Liebner

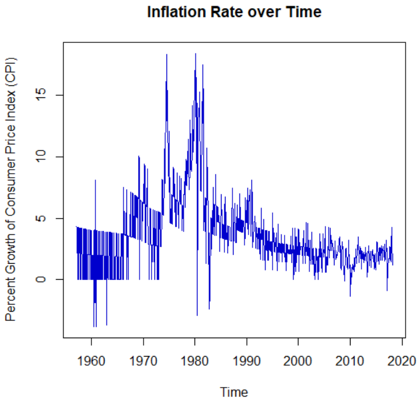
Lafayette College, Unity College, Marist College

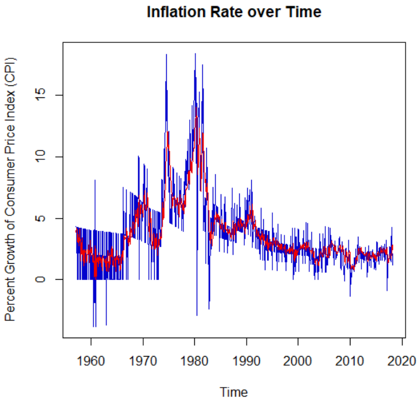
Wednesday, June 13, 2018

What is the problem?

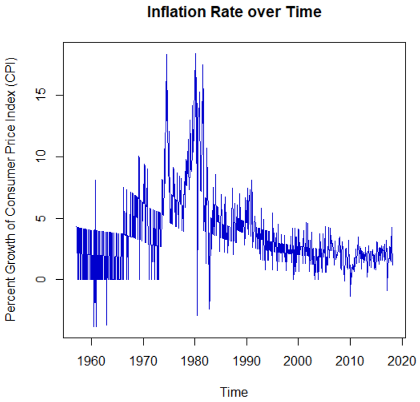


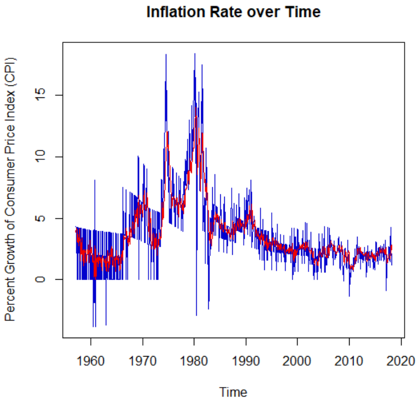
What is the problem?





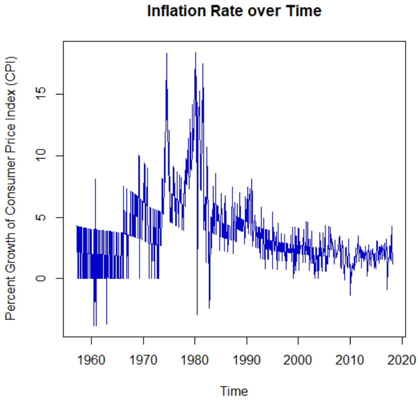
What is the problem?

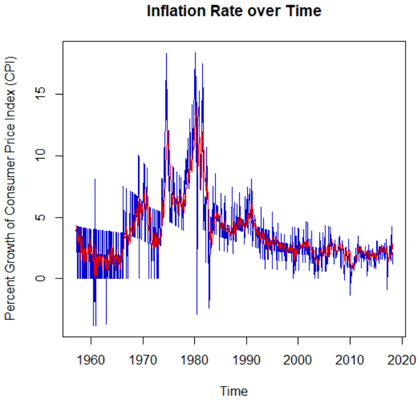




$$Y = \beta_0 + \sum_{i=1}^{n} \beta_i Y_{t-i} + \epsilon_{t-i}$$

What is the problem?



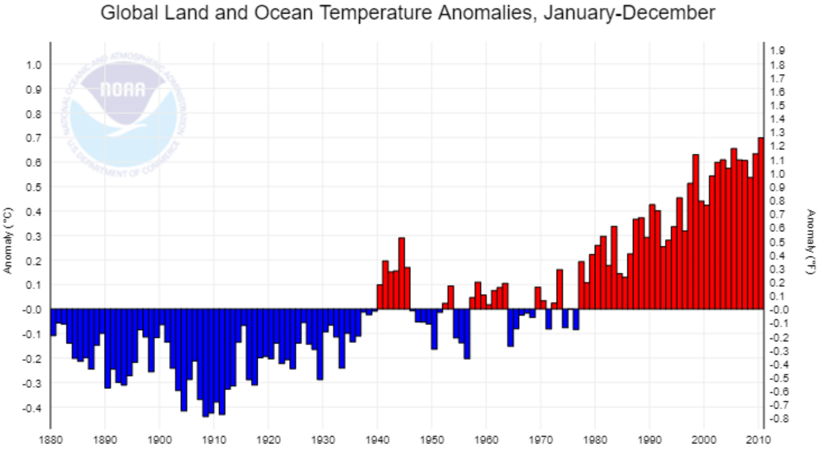


9 coefficients for autoregressive model:
1: 0.210967388; 2: 0.246292153; 3: 0.095668132;
4: -0.006722439; 5: 0.086106156; 6: 0.083077353;
7: 0.021451801; 8: 0.089686225; 9: 0.097260118

Goal

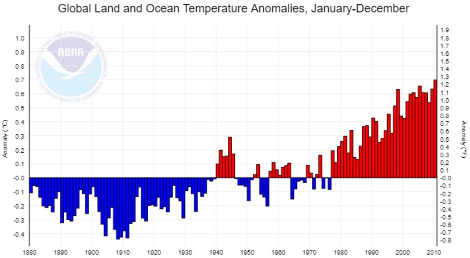
To develop a better quantitative method for locating break points in time series data.

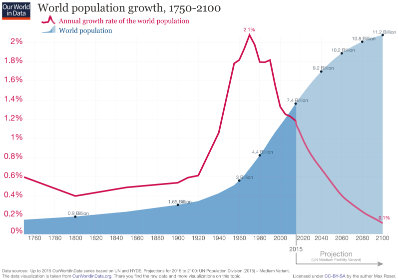
Other Data



Images courtesy of NOAA and Our World in Data.

Other Rate





Images courtesy of NOAA and Our World in Data.

Current Strategies

Bai-Perron Test (1998):

- Frequentist approach to identify significant changes in a time series.
- Search through every single possible break point and determine the best model off of the residual sum of squares.
- Conditions on the number of breaks specified by the user the output is a single model.
- It has multiple downsides however.

Current Strategies

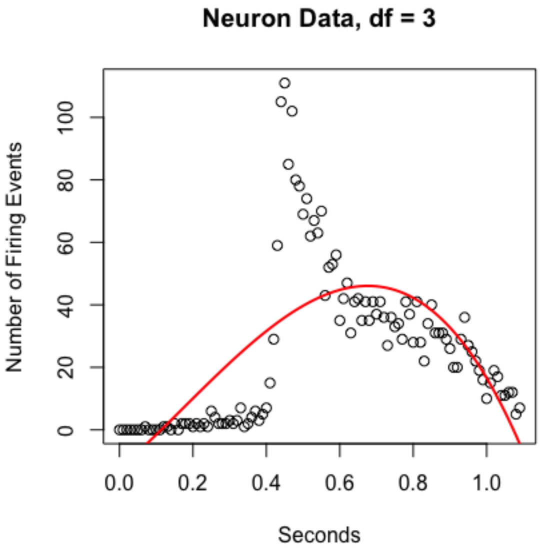
Curve Fitting with Splines

- Uses cubic polynomial piece wise functions that are connected at certain points, called knots.
- How many knots and where they are located determine the fit.

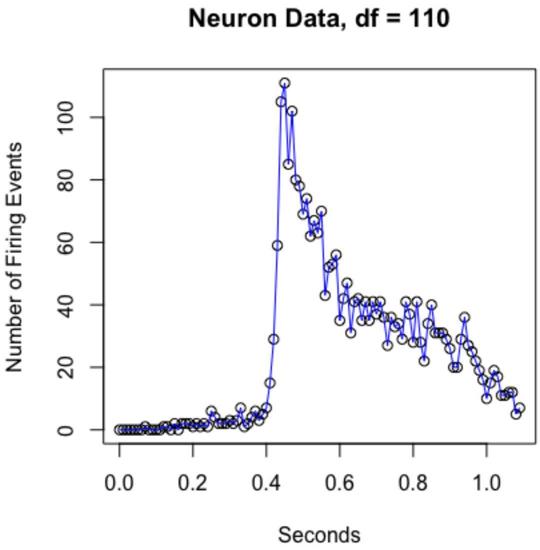
Generalized B Spline Function

$f(x) = \sum_{j=1}^{k} \beta_j B_{jk}(x)$

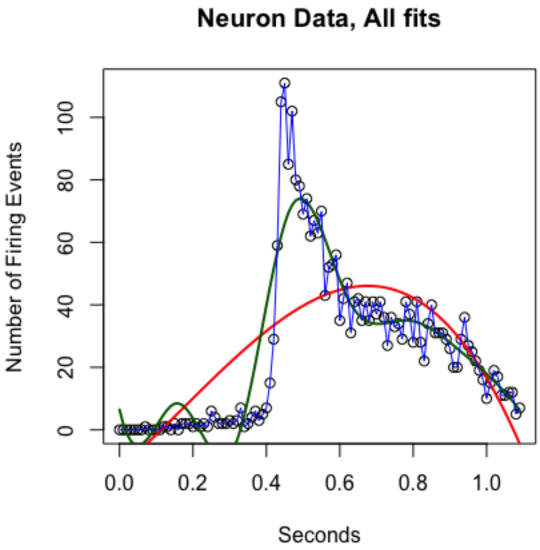
Current Strategies



Current Strategies



Current Strategies



Current Strategies

Bayesian Adaptive Regression Splines (BARS):

- A stochastic process that proposes new set of knots for the splines.
- This process involves adding, subtracting, and moving knots to obtain a draw from the possible fits to the model.
- The optimal fit is obtained by averaging the different fits.

Project Outline

Project Format:

1. Initial Knot/Breakpoint Problem
2. Modification of BARS to propose a new set of knots
3. Get $\theta = (K, \tau, \beta, \sigma)$ and find parameter distributions
4. Evaluate new proposed fit using Metropolis-Hastings
5. Repeat steps 2 through 4

Initial Breakpoint Problem

Determining the initial placement of the first or first couple of knots

Previous Approaches:

- random
- middle placement

New Approach: Bai-Perron (frequentist method)

- Places the initial breakpoints strategically to improve the model.

Modification of BARS

Modification: Autoregressive (AR) rather then cubic splines

AR(1) Model

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \epsilon_t$$

Modification: Adding, Subtracting, and Moving

- Adding, subtracting, and moving breakpoints all have their own probability of occurring.
- Proposed addition will be based on probability and current location of other breakpoints.
    - Changes are necessary because BARS proposes breakpoints that are too close together which is not appropriate for most time series.

Going Forward

- Develop adding and subtracting algorithms for BARS modification
- Figure out the algorithms for the best initial breakpoint(s) position(s)
- Lots of coding and simulations to R!!!