

Improving Bayesian Procedures to Detect Breakpoints in Time Series Data

Dainia Higgins, Nathaniel Wilson, Madison Ell, Sarah Klingbeil
Mentor: Jeff Liebner

Coppin State University, Fullerton College, Southern Adventist University

Tuesday, July 26, 2022



Acknowledgements

Special thanks to NSF Grant number DMS-2150343, Lafayette College and Mentor Professor Jeff Liebner for making this REU (Research Experience for Undergraduates) possible.

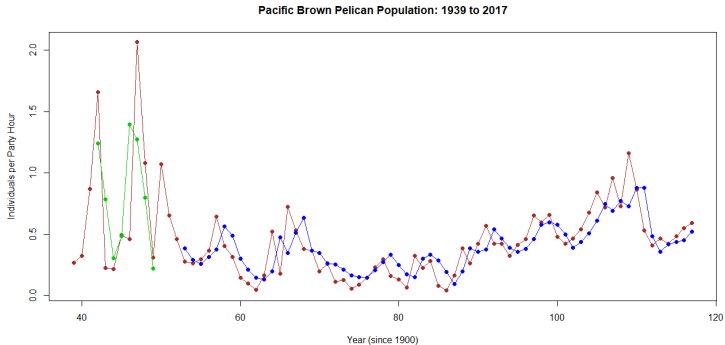
- 1 Project Overview
- 2 Progress and Results
- 3 Conclusion and Further Research

What is a Breakpoint?

Definition

Breakpoints (also known as Change Points) are points in time where the model changes.

- Breakpoints are significant changes in time series data



How do we find Breakpoints?

- 1 **Expert Opinion** : breakpoints are approximated by experts in a specific field based on historical knowledge.

What is Bai-Perron Test?

Definition

Bai-Perron Test is a general algorithm to find an optimal breakpoint set.

- ① a frequentist approach
- ② checks almost every single location for a breakpoint and returns the optimal set
- ③ requires a user to specify the number of breakpoints

What are some common types of time series models?

Auto-regressive (AR) model: each output value depends linearly upon previous values and an independent error term

AR(1)

$$x_t = \phi x_{t-1} + \epsilon_t$$

Moving average (MA) model: output value depends linearly upon previous error terms

MA(1)

$$x_t = \epsilon_t + \theta \epsilon_{t-1}$$

The Project

Bayesian Adaptive Auto-Regression (BAAR)

BAAR is a Bayesian method used to find the location and number of breakpoints in a time series.

- needs to have an input stating breakpoint places (Bai-Perron)

The Project

Bayesian Adaptive Auto-Regression (BAAR)

BAAR is a Bayesian method used to find the location and number of breakpoints in a time series.

- needs to have an input stating breakpoint places (Bai-Perron)
- A new breakpoint set is proposed at each step of the MCMC
 - birth, death, and move

The Project

Bayesian Adaptive Auto-Regression (BAAR)

BAAR is a Bayesian method used to find the location and number of breakpoints in a time series.

- **Metropolis-Hastings** (a MCMC method for receiving a sequence of random samples from the probability distribution when direct sampling is difficult) ratio determines the set's acceptance
 - Acceptance Ratio

$$ratio \approx \exp\left(\frac{-\Delta BIC}{2}\right) \frac{\pi(K_n)}{\pi(K_o)} \frac{\pi(\tau_n|K_n)}{\pi(\tau_o|K_o)} \frac{q(\tau_o K_o|\tau_n K_n)}{q(\tau_n K_n|\tau_o K_o)}$$

The Project

Bayesian Adaptive Auto-Regression (BAAR)

BAAR is a Bayesian method used to find the location and number of breakpoints in a time series.

- **Metropolis-Hastings** (a MCMC method for receiving a sequence of random samples from the probability distribution when direct sampling is difficult) ratio determines the set's acceptance
 - Acceptance Ratio

$$ratio \approx \exp\left(\frac{-\Delta BIC}{2}\right) \frac{\pi(K_n) \pi(\tau_n|K_n) q(\tau_o K_o|\tau_n K_n)}{\pi(K_o) \pi(\tau_o|K_o) q(\tau_n K_n|\tau_o K_o)}$$

- Hence, a distribution of possible breakpoints locations can be obtained

The Project

Bayesian Adaptive Moving-Average (BAMA)

BAMA is a Bayesian method used to find the location and number of breakpoints in a time series.

- Somewhat similar to BAAR in starting breakpoint and Metropolis-Hastings procedures

The Project

Bayesian Adaptive Moving-Average (BAMA)

BAMA is a Bayesian method used to find the location and number of breakpoints in a time series.

- Somewhat similar to BAAR in starting breakpoint and Metropolis-Hastings procedures
- Uses a Metropolis-Hastings procedure within a Gibbs sampler procedure to generate new coefficients

The Project

Bayesian Adaptive Moving-Average (BAMA)

BAMA is a Bayesian method used to find the location and number of breakpoints in a time series.

- Somewhat similar to BAAR in starting breakpoint and Metropolis-Hastings procedures
- Uses a Metropolis-Hastings procedure within a Gibbs sampler procedure to generate new coefficients
- Can handle seasonal/cyclical datasets

The Project

Bayesian Adaptive Moving-Average (BAMA)

BAMA is a Bayesian method used to find the location and number of breakpoints in a time series.

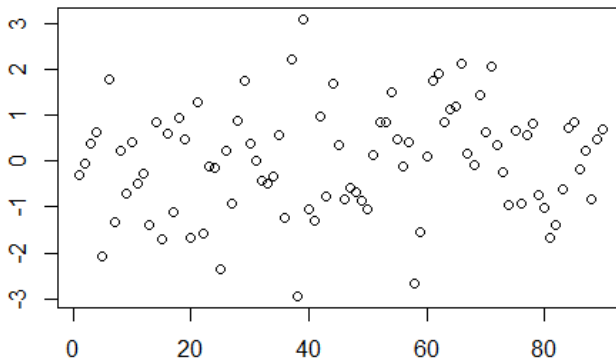
- Somewhat similar to BAAR in starting breakpoint and Metropolis-Hastings procedures
- Uses a Metropolis-Hastings procedure within a Gibbs sampler procedure to generate new coefficients
- Can handle seasonal/cyclical datasets
- Improves the Bayesian adaptive algorithm compatibility with various kinds of datasets

- 1 Project Overview
- 2 Progress and Results
- 3 Conclusion and Further Research

Stress Test

Example

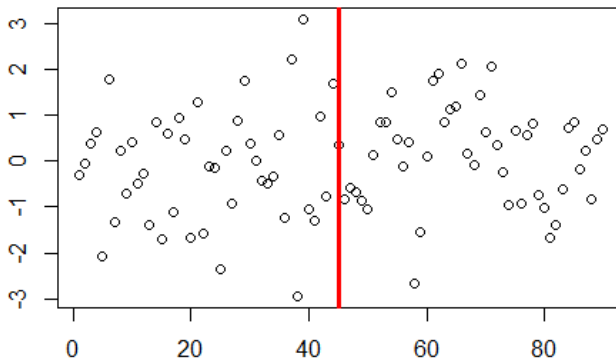
Where do you think a Breakpoint is located?



Stress Test

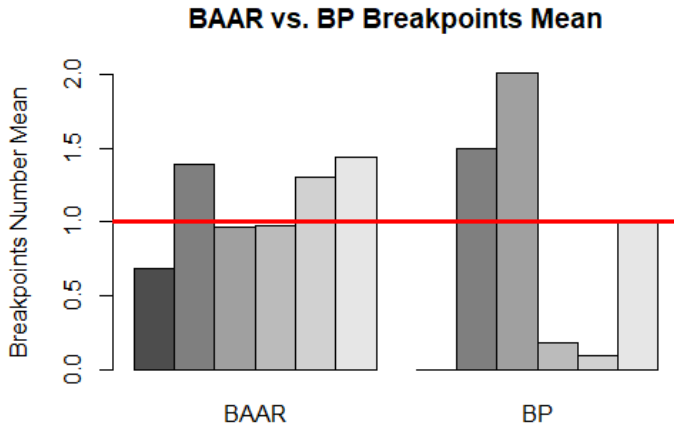
Example

How did you do?



Stress Test

Simulation results

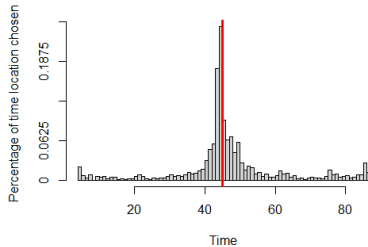


Standard Deviation=1,Breakpoint at 45

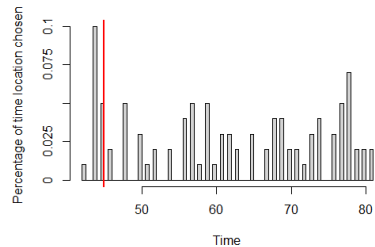
Stress Test

Simulation results

BAAR Breakpoints



Bai-Perron Breakpoints



- 1 Project Overview
- 2 Progress and Results
- 3 Conclusion and Further Research

Conclusion

- BAAR correctly identifies breakpoints with greater accuracy than existing algorithms (i.e. Bai-Perron)
- BAMA still a work in progress, in part due to COVID outbreak

Before JMM Plans

- 1 Conduct stress testing on BAMA
- 2 Run BAAR simulations on World Health Organization (WHO) dataset of suicide rates among 15-24 year olds in the United States of America
- 3 Pick a case study on real data use on BAMA
- 4 If time permits, author an R package containing the BAAR and BAMA algorithms

Further Research

- Improve starting point of the BAAR algorithm
- Expand BAAR and BAMA techniques to more complicated ARMA and ARIMA models

References

Bai, J. and Perron, P., (2003). *Computation and analysis of multiple structural change models*. Journal of applied econometrics, 18(1), pp.1-22.

Any Questions?

