A General Regression Changepoint Test for Time Series Data

Michael Robbins
RAND Corporation

Collaborators:
Colin Gallagher, Robert Lund

June 10, 2016

Changepoints and Climate

A **changepoint** is a discontinuity in the marginal distributions of a time-ordered sequence of data $X_1, X_2, \ldots, X_n$.

Changepoints and Climate

A **changepoint** is a discontinuity in the marginal distributions of a time-ordered sequence of data $X_1, X_2, \ldots, X_n$.

For instance, imagine an abrupt shift in:

- $E[X_t]$
- $Var(X_t)$
- $P(X_t \leq x)$

at time $c \in \{1, \ldots, n\}$ (the changepoint).

Changepoints and Climate

A **changepoint** is a discontinuity in the marginal distributions of a time-ordered sequence of data $X_1, X_2, \ldots, X_n$.

For instance, imagine an abrupt shift in:

- $E[X_t]$
- $Var(X_t)$
- $P(X_t \leq x)$

at time $c \in \{1, \ldots, n\}$ (the changepoint).

Possible causes of change-points in climate data:

- Weather gauge changes/relocations
- Local climate factors (e.g., urbanization)
- Geophysical events

Changepoints and Climate

A **changepoint** is a discontinuity in the marginal distributions of a time-ordered sequence of data $X_1, X_2, \ldots, X_n$.

For instance, imagine an abrupt shift in:

- $E[X_t]$
- $Var(X_t)$
- $P(X_t \leq x)$

at time $c \in \{1, \ldots, n\}$ (the changepoint).

Possible causes of changepoints in climate data:

- Weather gauge changes/relocations
- Local climate factors (e.g., urbanization)
- Geophysical events

Problem: Even after adjusting for changepoints, data may not be stationary due to:

- Seasonality
- Trend

A **changepoint** is a discontinuity in the marginal distributions of a time-ordered sequence of data $X_1, X_2, \ldots, X_n$.

For instance, imagine an abrupt shift in:

- $E[X_t]$
- $Var(X_t)$
- $P(X_t \leq x)$
- $\beta_0, \beta_1, \ldots$??

at time $c \in \{1, \ldots, n\}$ (the changepoint).

Possible causes of changepoints in climate data:

- Weather gauge changes/relocations
- Local climate factors (e.g., urbanization)
- Geophysical events

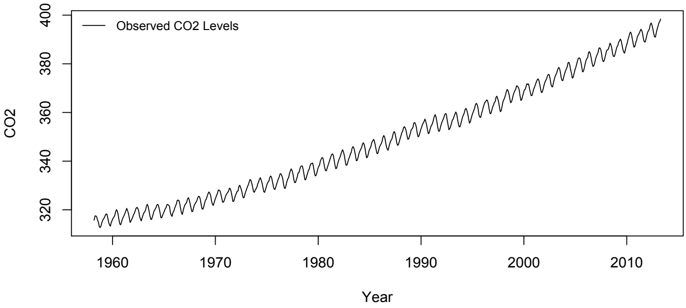
Problem: Even after adjusting for changepoints, data may not be stationary due to:

- Seasonality
- Trend

A Motivating Example

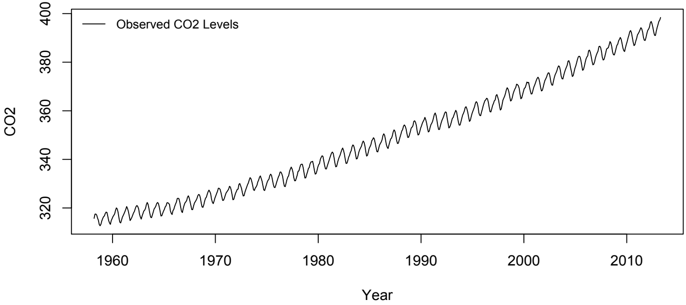
- The Kneeling Curve
  - Plots atmospheric CO_2 levels (by month).
  - Measured at **Mauna Loa, Hawaii**.
  - Monthly observations since 1958.

- The Kneeling Curve
  - Plots atmospheric CO_2 levels (by month).
  - Measured at **Mauna Loa, Hawaii**.
  - Monthly observations since 1958.



A Motivating Example

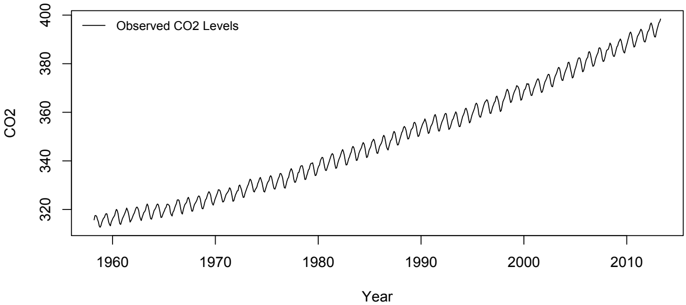
- The Kneeling Curve
  - Plots atmospheric CO_2 levels (by month).
  - Measured at **Mauna Loa, Hawaii**.
  - Monthly observations since 1958.



- Characteristics
  - Quadratic (or exponential) increasing trend.
  - Strong seasonality.

A Motivating Example

- The Kneeling Curve
  - Plots atmospheric CO_2 levels (by month).
  - Measured at **Mauna Loa, Hawaii**.
  - Monthly observations since 1958.



- Characteristics
  - Quadratic (or exponential) increasing trend.
  - Strong seasonality.
  - *Shifts in trend or in seasonality?*

Model the observed data y_t, for $1 \leq t \leq n$, as

$$y_t = \widetilde{\alpha}'\widetilde{\boldsymbol{x}}_t + \widetilde{\beta}'\widetilde{\boldsymbol{s}}_t + \widetilde{\gamma}'\widetilde{\boldsymbol{v}}_t + \epsilon_t, \tag{1}$$

A General Regression Model

Model the observed data y_t, for $1 \leq t \leq n$, as

$$y_t = \widetilde{\alpha}'\widetilde{\boldsymbol{x}}_t + \widetilde{\beta}'\widetilde{\boldsymbol{s}}_t + \widetilde{\gamma}'\widetilde{\boldsymbol{v}}_t + \epsilon_t, \qquad (1)$$

where

- $\widetilde{\boldsymbol{x}}_t$ is a vector of *functional form* predictors

A General Regression Model

Model the observed data y_t, for $1 \le t \le n$, as

$$y_t = \widetilde{\alpha}'\widetilde{\boldsymbol{x}}_t + \widetilde{\beta}'\widetilde{\boldsymbol{s}}_t + \widetilde{\gamma}'\widetilde{\boldsymbol{v}}_t + \epsilon_t, \qquad (1)$$

where

- $\widetilde{\boldsymbol{x}}_t$ is a vector of *functional form* predictors
  - i.e., $\widetilde{\boldsymbol{x}}_t = \{f_1(t/n), f_2(t/n), \ldots, f_{p_x}(t/n)\}'$
  - where $f_i(z)$ is continuous for $z \in [0, 1]$ and for $i = 1, \ldots, p_x$

A General Regression Model

Model the observed data y_t, for $1 \leq t \leq n$, as

$$y_t = \widetilde{\alpha}'\widetilde{\boldsymbol{x}}_t + \widetilde{\beta}'\widetilde{\boldsymbol{s}}_t + \widetilde{\gamma}'\widetilde{\boldsymbol{v}}_t + \epsilon_t, \tag{1}$$

where

- $\widetilde{\boldsymbol{x}}_t$ is a vector of *functional form* predictors
  - i.e., $\widetilde{\boldsymbol{x}}_t = \{f_1(t/n), f_2(t/n), \ldots, f_{p_x}(t/n)\}'$
  - where $f_i(z)$ is continuous for $z \in [0,1]$ and for $i = 1, \ldots, p_x$
- $\widetilde{\boldsymbol{s}}_t$ is a vector of *seasonal* predictors
  - Assume that $\sum_{t=1}^{T} \widetilde{\boldsymbol{s}}_t = \mathbf{0}$ for a period T.

A General Regression Model

Model the observed data y_t, for $1 \leq t \leq n$, as

$$y_t = \widetilde{\alpha}'\widetilde{\boldsymbol{x}}_t + \widetilde{\beta}'\widetilde{\boldsymbol{s}}_t + \widetilde{\gamma}'\widetilde{\boldsymbol{v}}_t + \epsilon_t, \tag{1}$$

where

- $\widetilde{\boldsymbol{x}}_t$ is a vector of *functional form* predictors
  - i.e., $\widetilde{\boldsymbol{x}}_t = \{f_1(t/n), f_2(t/n), \ldots, f_{p_x}(t/n)\}'$
  - where $f_i(z)$ is continuous for $z \in [0, 1]$ and for $i = 1, \ldots, p_x$
- $\widetilde{\boldsymbol{s}}_t$ is a vector of *seasonal* predictors
  - Assume that $\sum_{t=1}^{T} \widetilde{\boldsymbol{s}}_t = \mathbf{0}$ for a period T.
- $\widetilde{\boldsymbol{v}}_t$ is a vector of (stationary) random design points
  - $E\left[\widetilde{\boldsymbol{v}}_t\right] = \mathbf{0}$ and $\mathbf{Var}\left[\widetilde{\boldsymbol{v}}_t\right] = \boldsymbol{\Sigma}$

A General Regression Model

Model the observed data y_t, for $1 \leq t \leq n$, as

$$y_t = \widetilde{\alpha}'\widetilde{\boldsymbol{x}}_t + \widetilde{\beta}'\widetilde{\boldsymbol{s}}_t + \widetilde{\gamma}'\widetilde{\boldsymbol{v}}_t + \epsilon_t, \tag{1}$$

where

- $\widetilde{\boldsymbol{x}}_t$ is a vector of *functional form* predictors
  - i.e., $\widetilde{\boldsymbol{x}}_t = \{f_1(t/n), f_2(t/n), \ldots, f_{p_x}(t/n)\}'$
  - where $f_i(z)$ is continuous for $z \in [0, 1]$ and for $i = 1, \ldots, p_x$
- $\widetilde{\boldsymbol{s}}_t$ is a vector of *seasonal* predictors
  - Assume that $\sum_{t=1}^{T} \widetilde{\boldsymbol{s}}_t = \mathbf{0}$ for a period T.
- $\widetilde{\boldsymbol{v}}_t$ is a vector of (stationary) random design points
  - $E[\widetilde{\boldsymbol{v}}_t] = \mathbf{0}$ and $\mathbf{Var}[\widetilde{\boldsymbol{v}}_t] = \boldsymbol{\Sigma}$
- The ϵ_t are *stationary* mean zero errors.

A General Regression Model

Model the observed data y_t, for $1 \leq t \leq n$, as

$$y_t = \widetilde{\alpha}'\widetilde{\boldsymbol{x}}_t + \widetilde{\boldsymbol{\beta}}'\widetilde{\boldsymbol{s}}_t + \widetilde{\gamma}'\widetilde{\boldsymbol{v}}_t + \epsilon_t, \qquad (1)$$

where

- $\widetilde{\boldsymbol{x}}_t$ is a vector of *functional form* predictors
  - i.e., $\widetilde{\boldsymbol{x}}_t = \{f_1(t/n), f_2(t/n), \ldots, f_{p_x}(t/n)\}'$
  - where $f_i(z)$ is continuous for $z \in [0, 1]$ and for $i = 1, \ldots, p_x$
- $\widetilde{\boldsymbol{s}}_t$ is a vector of *seasonal* predictors
  - Assume that $\sum_{t=1}^{T} \widetilde{\boldsymbol{s}}_t = \mathbf{0}$ for a period T.
- $\widetilde{\boldsymbol{v}}_t$ is a vector of (stationary) random design points
  - $E[\widetilde{\boldsymbol{v}}_t] = \mathbf{0}$ and $\mathbf{Var}[\widetilde{\boldsymbol{v}}_t] = \boldsymbol{\Sigma}$
- The ϵ_t are *stationary* mean zero errors.
- $\widetilde{\alpha}$, $\widetilde{\boldsymbol{\beta}}$, and $\widetilde{\gamma}$ are vectors of regression coefficients.

Back to the Example

We assume the following model for the Mauna Loa CO_2 data:

$$\text{CO2}_t = \alpha_0 + \alpha_1 \left(\frac{t}{n}\right) + \alpha_2 \left(\frac{t}{n}\right)^2$$

Back to the Example

We assume the following model for the Mauna Loa CO_2 data:

$$\begin{aligned} \mathrm{CO2}_t &= \alpha_0 + \alpha_1 \left(\frac{t}{n}\right) + \alpha_2 \left(\frac{t}{n}\right)^2 \\ &\quad + \sum_{j=1}^{4} \left[\beta_{1,j} \cos\left(\frac{2\pi j t}{12}\right) + \beta_{2,j} \sin\left(\frac{2\pi j t}{12}\right) \right] \end{aligned}$$

Back to the Example

We assume the following model for the Mauna Loa CO_2 data:

$$\begin{aligned} \text{CO2}_t &= \alpha_0 + \alpha_1 \left(\frac{t}{n}\right) + \alpha_2 \left(\frac{t}{n}\right)^2 \\ &+ \sum_{j=1}^{4} \left[\beta_{1,j} \cos\left(\frac{2\pi jt}{12}\right) + \beta_{2,j} \sin\left(\frac{2\pi jt}{12}\right)\right] \\ &+ \gamma(\text{ENSO}_{t-12}) + \epsilon_t, \end{aligned}$$

where ENSO_t is the El Niño/Southern Oscillation index at time t.

Back to the Example

We assume the following model for the Mauna Loa CO_2 data:

$$
\begin{aligned}
\text{CO2}_t &= \alpha_0 + \alpha_1 \left(\frac{t}{n}\right) + \alpha_2 \left(\frac{t}{n}\right)^2 \\
&\quad + \sum_{j=1}^{4} \left[\beta_{1,j} \cos\left(\frac{2\pi j t}{12}\right) + \beta_{2,j} \sin\left(\frac{2\pi j t}{12}\right) \right] \\
&\quad + \gamma(\text{ENSO}_{t-12}) + \epsilon_t,
\end{aligned}
$$

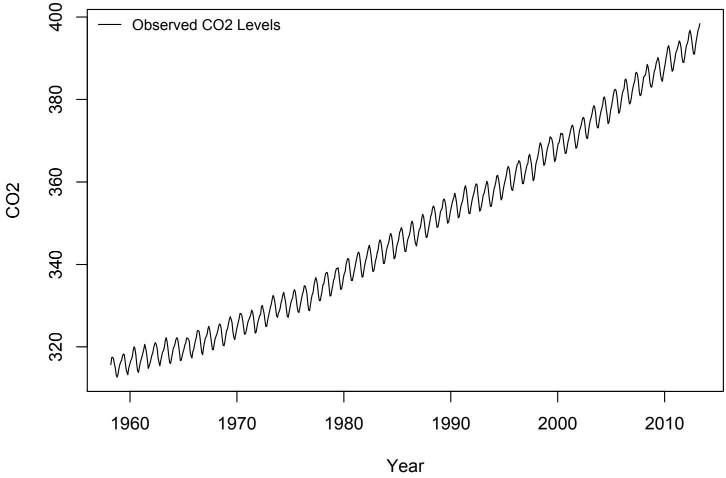
where ENSO_t is the El Niño/Southern Oscillation index at time t.

An equivalent representation is

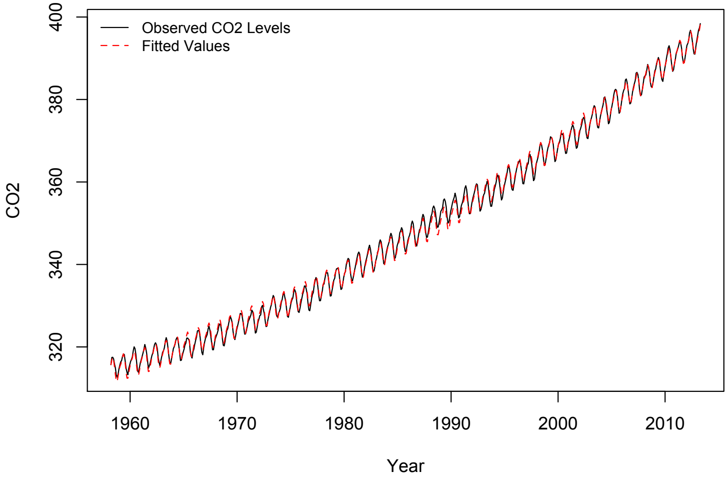
$$
\begin{aligned}
\text{CO2}_t &= \alpha_0 + \alpha_1 t + \alpha_2 t^2 \\
&\quad + \sum_{j=1}^{4} \left[\beta_{1,j} \cos\left(\frac{2\pi j t}{12}\right) + \beta_{2,j} \sin\left(\frac{2\pi j t}{12}\right) \right] \\
&\quad + \gamma(\text{ENSO}_{t-12}) + \epsilon_t,
\end{aligned}
$$

Back to the Example

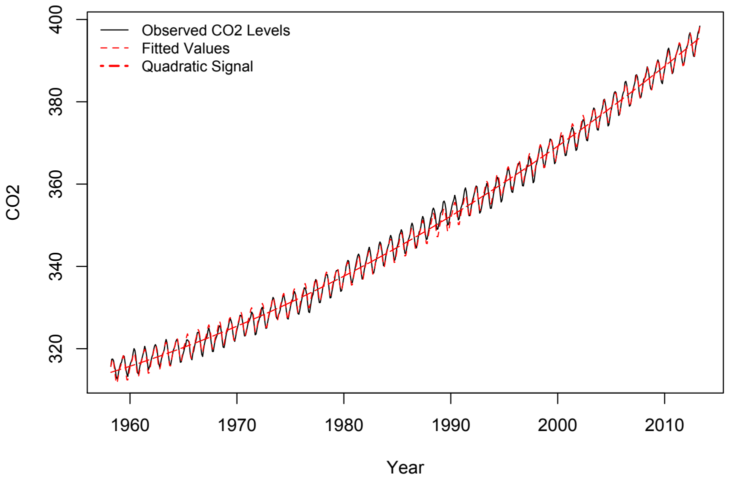
The Mauna Loa CO_2 data:



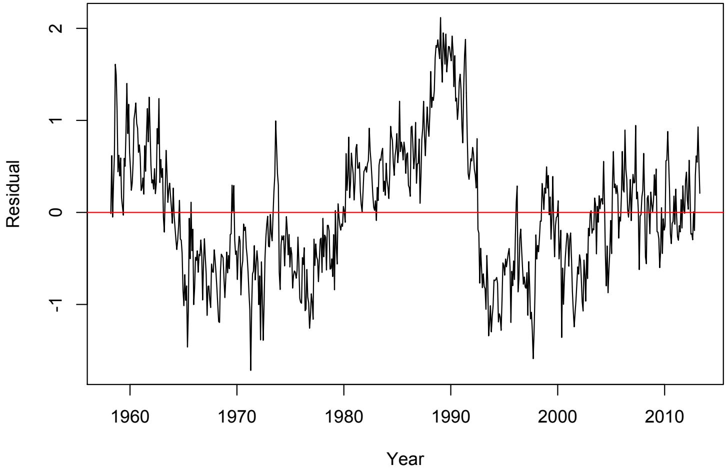
The Mauna Loa CO_2 data with fit:



The Mauna Loa CO_2 data with fit & trend:

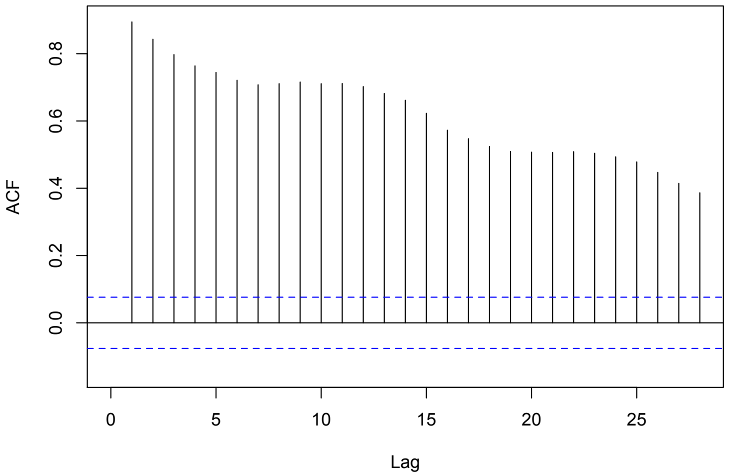


The OLS residuals:



Back to the Example

The ACF of the OLS residuals:



A General Regression *Changepoint* Model

Let $\widetilde{\boldsymbol{x}}_t = (\boldsymbol{x}_t', (\boldsymbol{x}_t^*)')'$, $\widetilde{\boldsymbol{s}}_t = (\boldsymbol{s}_t', (\boldsymbol{s}_t^*)')'$ and $\widetilde{\boldsymbol{x}}_t = (\boldsymbol{x}_t', (\boldsymbol{x}_t^*)')'$.

Update the model in (1) to allow for a changepoint in the trend:

$$\begin{aligned} y_t &= (\boldsymbol{\alpha} \qquad)'\boldsymbol{x}_t + (\boldsymbol{\beta} \qquad)'\boldsymbol{s}_t + (\boldsymbol{\gamma} \qquad)'\boldsymbol{v}_t \\ &\quad + (\boldsymbol{\alpha}^*)'\boldsymbol{x}_t^* + (\boldsymbol{\beta}^*)'\boldsymbol{s}_t^* + (\boldsymbol{\gamma}^*)'\boldsymbol{v}_t^* + \epsilon_t, \end{aligned}$$

A General Regression *Changepoint* Model

Let $\widetilde{\boldsymbol{x}}_t = (\boldsymbol{x}_t', (\boldsymbol{x}_t^*)')'$, $\widetilde{\boldsymbol{s}}_t = (\boldsymbol{s}_t', (\boldsymbol{s}_t^*)')'$ and $\widetilde{\boldsymbol{x}}_t = (\boldsymbol{x}_t', (\boldsymbol{x}_t^*)')'$.

Update the model in (1) to allow for a changepoint in the trend:

$$\begin{aligned} y_t &= (\boldsymbol{\alpha} + \boldsymbol{\delta}_{\mathrm{x},t})' \boldsymbol{x}_t + (\boldsymbol{\beta} + \boldsymbol{\delta}_{\mathrm{s},t})' \boldsymbol{s}_t + (\boldsymbol{\gamma} + \boldsymbol{\delta}_{\mathrm{v},t})' \boldsymbol{v}_t \\ &\quad + (\boldsymbol{\alpha}^*)' \boldsymbol{x}_t^* + (\boldsymbol{\beta}^*)' \boldsymbol{s}_t^* + (\boldsymbol{\gamma}^*)' \boldsymbol{v}_t^* + \epsilon_t, \end{aligned}$$

where

$$\boldsymbol{\delta}_t = (\boldsymbol{\delta}_{\mathrm{x},t}', \boldsymbol{\delta}_{\mathrm{s},t}', \boldsymbol{\delta}_{\mathrm{v},t}')' = \begin{cases} \mathbf{0}, & 1 \le t \le c, \\ \boldsymbol{\Delta}, & c < t \le n, \end{cases}$$

for some *unknown* changepoint time c, where $1 \le c < n$.

Let $\widetilde{\boldsymbol{x}}_t = (\boldsymbol{x}_t', (\boldsymbol{x}_t^*)')'$, $\widetilde{\boldsymbol{s}}_t = (\boldsymbol{s}_t', (\boldsymbol{s}_t^*)')'$ and $\widetilde{\boldsymbol{x}}_t = (\boldsymbol{x}_t', (\boldsymbol{x}_t^*)')'$.

Update the model in (1) to allow for a changepoint in the trend:

$$\begin{aligned} y_t \;&=\; (\boldsymbol{\alpha} + \boldsymbol{\delta}_{\mathrm{x},t})'\boldsymbol{x}_t + (\boldsymbol{\beta} + \boldsymbol{\delta}_{\mathrm{s},t})'\boldsymbol{s}_t + (\boldsymbol{\gamma} + \boldsymbol{\delta}_{\mathrm{v},t})'\boldsymbol{v}_t \\ &\quad + (\boldsymbol{\alpha}^*)'\boldsymbol{x}_t^* + (\boldsymbol{\beta}^*)'\boldsymbol{s}_t^* + (\boldsymbol{\gamma}^*)'\boldsymbol{v}_t^* + \epsilon_t, \end{aligned}$$

where

$$\boldsymbol{\delta}_t = (\boldsymbol{\delta}_{\mathrm{x},t}', \boldsymbol{\delta}_{\mathrm{s},t}', \boldsymbol{\delta}_{\mathrm{v},t}')' = \begin{cases} \mathbf{0}, & 1 \le t \le c, \\ \boldsymbol{\Delta}, & c < t \le n, \end{cases}$$

for some *unknown* changepoint time c, where $1 \le c < n$.

That is,

- Coefs. corresponding to $\boldsymbol{x}_t$, $\boldsymbol{s}_t$, and $\boldsymbol{v}_t$ are allowed to change;

A General Regression *Changepoint* Model

Let $\widetilde{\boldsymbol{x}}_t = (\boldsymbol{x}_t', (\boldsymbol{x}_t^*)')'$, $\widetilde{\boldsymbol{s}}_t = (\boldsymbol{s}_t', (\boldsymbol{s}_t^*)')'$ and $\widetilde{\boldsymbol{x}}_t = (\boldsymbol{x}_t', (\boldsymbol{x}_t^*)')'$.

Update the model in (1) to allow for a changepoint in the trend:

$$\begin{aligned} y_t &= (\boldsymbol{\alpha} + \boldsymbol{\delta}_{\mathrm{x},t})' \boldsymbol{x}_t + (\boldsymbol{\beta} + \boldsymbol{\delta}_{\mathrm{s},t})' \boldsymbol{s}_t + (\boldsymbol{\gamma} + \boldsymbol{\delta}_{\mathrm{v},t})' \boldsymbol{v}_t \\ & \quad + (\boldsymbol{\alpha}^*)' \boldsymbol{x}_t^* + (\boldsymbol{\beta}^*)' \boldsymbol{s}_t^* + (\boldsymbol{\gamma}^*)' \boldsymbol{v}_t^* + \epsilon_t, \end{aligned}$$

where

$$\boldsymbol{\delta}_t = (\boldsymbol{\delta}_{\mathrm{x},t}', \boldsymbol{\delta}_{\mathrm{s},t}', \boldsymbol{\delta}_{\mathrm{v},t}')' = \begin{cases} \mathbf{0}, & 1 \leq t \leq c, \\ \boldsymbol{\Delta}, & c < t \leq n, \end{cases}$$

for some *unknown* changepoint time c, where $1 \leq c < n$.

That is,

- Coefs. corresponding to $\boldsymbol{x}_t$, $\boldsymbol{s}_t$, and $\boldsymbol{v}_t$ are allowed to change;
- Coefs. corresponding to $\boldsymbol{x}_t^*$, $\boldsymbol{s}_t^*$, and $\boldsymbol{v}_t^*$ may *not* change;

A General Regression *Changepoint* Model

Let $\widetilde{\boldsymbol{x}}_t = (\boldsymbol{x}_t', (\boldsymbol{x}_t^*)')'$, $\widetilde{\boldsymbol{s}}_t = (\boldsymbol{s}_t', (\boldsymbol{s}_t^*)')'$ and $\widetilde{\boldsymbol{x}}_t = (\boldsymbol{x}_t', (\boldsymbol{x}_t^*)')'$.

Update the model in (1) to allow for a changepoint in the trend:

$$\begin{aligned} y_t &= (\boldsymbol{\alpha} + \boldsymbol{\delta}_{\mathrm{x},t})' \boldsymbol{x}_t + (\boldsymbol{\beta} + \boldsymbol{\delta}_{\mathrm{s},t})' \boldsymbol{s}_t + (\boldsymbol{\gamma} + \boldsymbol{\delta}_{\mathrm{v},t})' \boldsymbol{v}_t \\ &\quad + (\boldsymbol{\alpha}^*)' \boldsymbol{x}_t^* + (\boldsymbol{\beta}^*)' \boldsymbol{s}_t^* + (\boldsymbol{\gamma}^*)' \boldsymbol{v}_t^* + \epsilon_t, \end{aligned}$$

where

$$\boldsymbol{\delta}_t = (\boldsymbol{\delta}_{\mathrm{x},t}', \boldsymbol{\delta}_{\mathrm{s},t}', \boldsymbol{\delta}_{\mathrm{v},t}')' = \begin{cases} \mathbf{0}, & 1 \le t \le c, \\ \boldsymbol{\Delta}, & c < t \le n, \end{cases}$$

for some *unknown* changepoint time c, where $1 \le c < n$.

That is,

- Coefs. corresponding to $\boldsymbol{x}_t$, $\boldsymbol{s}_t$, and $\boldsymbol{v}_t$ are allowed to change;
- Coefs. corresponding to $\boldsymbol{x}_t^*$, $\boldsymbol{s}_t^*$, and $\boldsymbol{v}_t^*$ may *not* change;
- Designed to optimize flexibility when modeling a changepoint.

A General Regression *Changepoint* Model

Let $\widetilde{\boldsymbol{x}}_t = (\boldsymbol{x}_t', (\boldsymbol{x}_t^*)')'$, $\widetilde{\boldsymbol{s}}_t = (\boldsymbol{s}_t', (\boldsymbol{s}_t^*)')'$ and $\widetilde{\boldsymbol{x}}_t = (\boldsymbol{x}_t', (\boldsymbol{x}_t^*)')'$.

Update the model in (1) to allow for a changepoint in the trend:

$$\begin{aligned} y_t &= (\boldsymbol{\alpha} + \boldsymbol{\delta}_{\mathrm{x},t})' \boldsymbol{x}_t + (\boldsymbol{\beta} + \boldsymbol{\delta}_{\mathrm{s},t})' \boldsymbol{s}_t + (\boldsymbol{\gamma} + \boldsymbol{\delta}_{\mathrm{v},t})' \boldsymbol{v}_t \\ &\quad + (\boldsymbol{\alpha}^*)' \boldsymbol{x}_t^* + (\boldsymbol{\beta}^*)' \boldsymbol{s}_t^* + (\boldsymbol{\gamma}^*)' \boldsymbol{v}_t^* + \epsilon_t, \end{aligned}$$

where

$$\boldsymbol{\delta}_t = (\boldsymbol{\delta}_{\mathrm{x},t}', \boldsymbol{\delta}_{\mathrm{s},t}', \boldsymbol{\delta}_{\mathrm{v},t}')' = \begin{cases} \mathbf{0}, & 1 \le t \le c, \\ \boldsymbol{\Delta}, & c < t \le n, \end{cases}$$

for some *unknown* changepoint time c, where $1 \le c < n$.

That is,

- Coefs. corresponding to $\boldsymbol{x}_t$, $\boldsymbol{s}_t$, and $\boldsymbol{v}_t$ are allowed to change;
- Coefs. corresponding to $\boldsymbol{x}_t^*$, $\boldsymbol{s}_t^*$, and $\boldsymbol{v}_t^*$ may *not* change;
- Designed to optimize flexibility when modeling a changepoint.
- $\boldsymbol{\Delta}$ (a vector) is the amount of the change.

A General Regression *Changepoint* Model

Let $\widetilde{\boldsymbol{x}}_t = (\boldsymbol{x}'_t, (\boldsymbol{x}^*_t)')'$, $\widetilde{\boldsymbol{s}}_t = (\boldsymbol{s}'_t, (\boldsymbol{s}^*_t)')'$ and $\widetilde{\boldsymbol{x}}_t = (\boldsymbol{x}'_t, (\boldsymbol{x}^*_t)')'$.

Update the model in (1) to allow for a changepoint in the trend:

$$\begin{aligned} y_t &= (\boldsymbol{\alpha} + \boldsymbol{\delta}_{\mathrm{x},t})'\boldsymbol{x}_t + (\boldsymbol{\beta} + \boldsymbol{\delta}_{\mathrm{s},t})'\boldsymbol{s}_t + (\boldsymbol{\gamma} + \boldsymbol{\delta}_{\mathrm{v},t})'\boldsymbol{v}_t \\ &\quad + (\boldsymbol{\alpha}^*)'\boldsymbol{x}^*_t + (\boldsymbol{\beta}^*)'\boldsymbol{s}^*_t + (\boldsymbol{\gamma}^*)'\boldsymbol{v}^*_t + \epsilon_t, \end{aligned}$$

where

$$\boldsymbol{\delta}_t = (\boldsymbol{\delta}'_{\mathrm{x},t}, \boldsymbol{\delta}'_{\mathrm{s},t}, \boldsymbol{\delta}'_{\mathrm{v},t})' = \begin{cases} \mathbf{0}, & 1 \le t \le c, \\ \boldsymbol{\Delta}, & c < t \le n, \end{cases}$$

for some *unknown* changepoint time c, where $1 \le c < n$.

That is,

- Coefs. corresponding to $\boldsymbol{x}_t$, $\boldsymbol{s}_t$, and $\boldsymbol{v}_t$ are allowed to change;
- Coefs. corresponding to $\boldsymbol{x}^*_t$, $\boldsymbol{s}^*_t$, and $\boldsymbol{v}^*_t$ may *not* change;
- Designed to optimize flexibility when modeling a changepoint.
- $\boldsymbol{\Delta}$ (a vector) is the amount of the change.

We test $H_0 : \boldsymbol{\Delta} = \mathbf{0}$ vs. $H_1 : \boldsymbol{\Delta} \neq \mathbf{0}$.

Test Statistics for Changepoint Detection

For now we assume that the ϵ_t are i.i.d. with $\text{Var}(\epsilon_t) = \tau^2$.

Test Statistics for Changepoint Detection

For now we assume that the ϵ_t are i.i.d. with $\text{Var}(\epsilon_t) = \tau^2$.

When allowing a changepoint at time k,

Test Statistics for Changepoint Detection

For now we assume that the ϵ_t are i.i.d. with $\text{Var}(\epsilon_t) = \tau^2$.

When allowing a changepoint at time k, let

- $\widehat{\boldsymbol{\Delta}}_k$ — The OLS estimator of $\boldsymbol{\Delta}$

Test Statistics for Changepoint Detection

For now we assume that the ϵ_t are i.i.d. with $\text{Var}(\epsilon_t) = \tau^2$.

When allowing a changepoint at time k, let

- $\widehat{\boldsymbol{\Delta}}_k$ — The OLS estimator of $\boldsymbol{\Delta}$
- $\widehat{\textbf{Var}}(\widehat{\boldsymbol{\Delta}}_k) = \hat{\tau}^2 \boldsymbol{C}_k^{-1}$ — The estimated covariance matrix of $\widehat{\boldsymbol{\Delta}}_k$

Test Statistics for Changepoint Detection

For now we assume that the ϵ_t are i.i.d. with $\text{Var}(\epsilon_t) = \tau^2$.

When allowing a changepoint at time k, let

- $\widehat{\boldsymbol{\Delta}}_k$ — The OLS estimator of $\boldsymbol{\Delta}$
- $\widehat{\textbf{Var}}(\widehat{\boldsymbol{\Delta}}_k) = \hat{\tau}^2 \boldsymbol{C}_k^{-1}$ — The estimated covariance matrix of $\widehat{\boldsymbol{\Delta}}_k$

We derive a Wald statistic via

$$\widehat{F}_k = \widehat{\boldsymbol{\Delta}}_k' [\widehat{\textbf{Var}}(\widehat{\boldsymbol{\Delta}}_k)]^{-1} \widehat{\boldsymbol{\Delta}}_k = \widehat{\boldsymbol{\Delta}}_k' \boldsymbol{C}_k \widehat{\boldsymbol{\Delta}}_k / \hat{\tau}^2,$$

which equivalent to the likelihood ratio and F-tests.

Test Statistics for Changepoint Detection

For now we assume that the ϵ_t are i.i.d. with $\text{Var}(\epsilon_t) = \tau^2$.

When allowing a changepoint at time k, let

- $\widehat{\boldsymbol{\Delta}}_k$ — The OLS estimator of $\boldsymbol{\Delta}$
- $\widehat{\textbf{Var}}(\widehat{\boldsymbol{\Delta}}_k) = \hat{\tau}^2 \boldsymbol{C}_k^{-1}$ — The estimated covariance matrix of $\widehat{\boldsymbol{\Delta}}_k$

We derive a Wald statistic via

$$\widehat{F}_k = \widehat{\boldsymbol{\Delta}}_k' [\widehat{\textbf{Var}}(\widehat{\boldsymbol{\Delta}}_k)]^{-1} \widehat{\boldsymbol{\Delta}}_k = \widehat{\boldsymbol{\Delta}}_k' \boldsymbol{C}_k \widehat{\boldsymbol{\Delta}}_k / \hat{\tau}^2,$$

which equivalent to the likelihood ratio and F-tests.

To detect a change at an *unknown* time, we consider

$$\widehat{F} = \max_{\ell \leq \frac{k}{n} \leq h} \widehat{F}_k$$

for truncation values ℓ and h that satisfy $0 < \ell < h < 1$.

Test Statistics for Changepoint Detection

Let $\hat{\mu}_t = \hat{\boldsymbol{\alpha}}'\widetilde{\boldsymbol{x}}_t + \hat{\boldsymbol{\beta}}'\widetilde{\boldsymbol{w}}_t + \hat{\boldsymbol{\gamma}}'\widetilde{\boldsymbol{s}}_t$.

Test Statistics for Changepoint Detection

Let $\hat{\mu}_t = \hat{\boldsymbol{\alpha}}' \widetilde{\boldsymbol{x}}_t + \hat{\boldsymbol{\beta}}' \widetilde{\boldsymbol{w}}_t + \hat{\boldsymbol{\gamma}}' \widetilde{\boldsymbol{s}}_t$. Next, write

$$\hat{\epsilon}_t = y_t - \hat{\mu}_t.$$

The sequence $\{\hat{\epsilon}_t\}$ is referred to as the **OLS residuals**.

Test Statistics for Changepoint Detection

Let $\hat{\mu}_t = \hat{\boldsymbol{\alpha}}'\widetilde{\boldsymbol{x}}_t + \hat{\boldsymbol{\beta}}'\widetilde{\boldsymbol{w}}_t + \hat{\boldsymbol{\gamma}}'\widetilde{\boldsymbol{s}}_t$. Next, write

$$\hat{\epsilon}_t = y_t - \hat{\mu}_t.$$

The sequence $\{\hat{\epsilon}_t\}$ is referred to as the **OLS residuals**.

We seek to express the $\widehat{F}_k$ statistic in terms of OLS residuals for two reasons:

- Helps when determining the sampling distribution (needed for critical values/p-values)
- Helps when extending the test to situations involving autocorrelation

Test Statistics for Changepoint Detection

Let $\hat{\mu}_t = \hat{\boldsymbol{\alpha}}' \widetilde{\boldsymbol{x}}_t + \hat{\boldsymbol{\beta}}' \widetilde{\boldsymbol{w}}_t + \hat{\boldsymbol{\gamma}}' \widetilde{\boldsymbol{s}}_t$. Next, write

$$\hat{\epsilon}_t = y_t - \hat{\mu}_t.$$

The sequence $\{\hat{\epsilon}_t\}$ is referred to as the **OLS residuals**.

Let

$$\boldsymbol{N}_{\mathrm{x},k} = \sum_{t=1}^{k} \hat{\epsilon}_t \boldsymbol{x}_t, \quad \boldsymbol{N}_{\mathrm{s},k} = \sum_{t=1}^{k} \hat{\epsilon}_t \boldsymbol{s}_t, \quad \text{and} \quad \boldsymbol{N}_{\mathrm{v},k} = \sum_{t=1}^{k} \hat{\epsilon}_t \boldsymbol{v}_t.$$

Test Statistics for Changepoint Detection

Let $\hat{\mu}_t = \hat{\boldsymbol{\alpha}}'\widetilde{\boldsymbol{x}}_t + \hat{\boldsymbol{\beta}}'\widetilde{\boldsymbol{w}}_t + \hat{\boldsymbol{\gamma}}'\widetilde{\boldsymbol{s}}_t$. Next, write

$$\hat{\epsilon}_t = y_t - \hat{\mu}_t.$$

The sequence $\{\hat{\epsilon}_t\}$ is referred to as the **OLS residuals**.

Let

$$\boldsymbol{N}_{\mathrm{x},k} = \sum_{t=1}^{k} \hat{\epsilon}_t \boldsymbol{x}_t, \quad \boldsymbol{N}_{\mathrm{s},k} = \sum_{t=1}^{k} \hat{\epsilon}_t \boldsymbol{s}_t, \quad \text{and} \quad \boldsymbol{N}_{\mathrm{v},k} = \sum_{t=1}^{k} \hat{\epsilon}_t \boldsymbol{v}_t.$$

One can show that

$$\widehat{F}_k = \widehat{F}_{\mathrm{x},k} + \widehat{F}_{\mathrm{s},k} + \widehat{F}_{\mathrm{v},k},$$

Test Statistics for Changepoint Detection

Let $\hat{\mu}_t = \hat{\boldsymbol{\alpha}}'\widetilde{\boldsymbol{x}}_t + \hat{\boldsymbol{\beta}}'\widetilde{\boldsymbol{w}}_t + \hat{\boldsymbol{\gamma}}'\widetilde{\boldsymbol{s}}_t$. Next, write

$$\hat{\epsilon}_t = y_t - \hat{\mu}_t.$$

The sequence $\{\hat{\epsilon}_t\}$ is referred to as the **OLS residuals**.

Let

$$\boldsymbol{N}_{\mathrm{x},k} = \sum_{t=1}^{k} \hat{\epsilon}_t \boldsymbol{x}_t, \quad \boldsymbol{N}_{\mathrm{s},k} = \sum_{t=1}^{k} \hat{\epsilon}_t \boldsymbol{s}_t, \quad \text{and} \quad \boldsymbol{N}_{\mathrm{v},k} = \sum_{t=1}^{k} \hat{\epsilon}_t \boldsymbol{v}_t.$$

One can show that

$$\widehat{F}_k = \widehat{F}_{\mathrm{x},k} + \widehat{F}_{\mathrm{s},k} + \widehat{F}_{\mathrm{v},k},$$

where

$$\widehat{F}_{\mathrm{x},k} = \boldsymbol{N}'_{\mathrm{x},k}[\mathbf{Var}(\boldsymbol{N}_{\mathrm{x},k})]^{-1}\boldsymbol{N}_{\mathrm{x},k} = \frac{\boldsymbol{N}'_{\mathrm{x},k}\boldsymbol{C}^{-1}_{\mathrm{x},k}\boldsymbol{N}_{\mathrm{x},k}}{\hat{\tau}^2},$$

- $\widehat{F}_{\mathrm{x},k}$ is a statistic to test for only a change in trend

Test Statistics for Changepoint Detection

Let $\hat{\mu}_t = \hat{\boldsymbol{\alpha}}'\widetilde{\boldsymbol{x}}_t + \hat{\boldsymbol{\beta}}'\widetilde{\boldsymbol{w}}_t + \hat{\boldsymbol{\gamma}}'\widetilde{\boldsymbol{s}}_t$. Next, write

$$\hat{\epsilon}_t = y_t - \hat{\mu}_t.$$

The sequence $\{\hat{\epsilon}_t\}$ is referred to as the **OLS residuals**.

Let

$$\boldsymbol{N}_{\mathrm{x},k} = \sum_{t=1}^{k} \hat{\epsilon}_t \boldsymbol{x}_t, \quad \boldsymbol{N}_{\mathrm{s},k} = \sum_{t=1}^{k} \hat{\epsilon}_t \boldsymbol{s}_t, \quad \text{and} \quad \boldsymbol{N}_{\mathrm{v},k} = \sum_{t=1}^{k} \hat{\epsilon}_t \boldsymbol{v}_t.$$

One can show that

$$\widehat{F}_k = \widehat{F}_{\mathrm{x},k} + \widehat{F}_{\mathrm{s},k} + \widehat{F}_{\mathrm{v},k},$$

where

$$\widehat{F}_{\mathrm{s},k} = \boldsymbol{N}'_{\mathrm{s},k}[\mathbf{Var}(\boldsymbol{N}_{\mathrm{s},k})]^{-1}\boldsymbol{N}_{\mathrm{s},k} = \frac{\boldsymbol{N}'_{\mathrm{s},k}\boldsymbol{C}^{-1}_{\mathrm{s},k}\boldsymbol{N}_{\mathrm{s},k}}{\hat{\tau}^2},$$

- $\widehat{F}_{\mathrm{s},k}$ is a statistic to test for only a change in seasonal coefs

Test Statistics for Changepoint Detection

Let $\hat{\mu}_t = \hat{\boldsymbol{\alpha}}'\widetilde{\boldsymbol{x}}_t + \hat{\boldsymbol{\beta}}'\widetilde{\boldsymbol{w}}_t + \hat{\boldsymbol{\gamma}}'\widetilde{\boldsymbol{s}}_t$. Next, write

$$\hat{\epsilon}_t = y_t - \hat{\mu}_t.$$

The sequence $\{\hat{\epsilon}_t\}$ is referred to as the **OLS residuals**.

Let

$$\boldsymbol{N}_{\mathrm{x},k} = \sum_{t=1}^{k} \hat{\epsilon}_t \boldsymbol{x}_t, \quad \boldsymbol{N}_{\mathrm{s},k} = \sum_{t=1}^{k} \hat{\epsilon}_t \boldsymbol{s}_t, \quad \text{and} \quad \boldsymbol{N}_{\mathrm{v},k} = \sum_{t=1}^{k} \hat{\epsilon}_t \boldsymbol{v}_t.$$

One can show that

$$\widehat{F}_k = \widehat{F}_{\mathrm{x},k} + \widehat{F}_{\mathrm{s},k} + \widehat{F}_{\mathrm{v},k},$$

where

$$\widehat{F}_{\mathrm{v},k} = \boldsymbol{N}'_{\mathrm{v},k}[\mathbf{Var}(\boldsymbol{N}_{\mathrm{v},k})]^{-1}\boldsymbol{N}_{\mathrm{v},k} = \frac{\boldsymbol{N}'_{\mathrm{v},k}\boldsymbol{C}^{-1}_{\mathrm{v},k}\boldsymbol{N}_{\mathrm{v},k}}{\hat{\tau}^2},$$

- $\widehat{F}_{\mathrm{v},k}$ is a statistic to test for only a change in covariate coefs

Test Statistics for Changepoint Detection

Let $\hat{\mu}_t = \hat{\boldsymbol{\alpha}}'\widetilde{\boldsymbol{x}}_t + \hat{\boldsymbol{\beta}}'\widetilde{\boldsymbol{w}}_t + \hat{\boldsymbol{\gamma}}'\widetilde{\boldsymbol{s}}_t$. Next, write

$$\hat{\epsilon}_t = y_t - \hat{\mu}_t.$$

The sequence $\{\hat{\epsilon}_t\}$ is referred to as the **OLS residuals**.

Let

$$\boldsymbol{N}_{\mathrm{x},k} = \sum_{t=1}^{k} \hat{\epsilon}_t \boldsymbol{x}_t, \quad \boldsymbol{N}_{\mathrm{s},k} = \sum_{t=1}^{k} \hat{\epsilon}_t \boldsymbol{s}_t, \quad \text{and} \quad \boldsymbol{N}_{\mathrm{v},k} = \sum_{t=1}^{k} \hat{\epsilon}_t \boldsymbol{v}_t.$$

One can show that

$$\widehat{F}_k = \widehat{F}_{\mathrm{x},k} + \widehat{F}_{\mathrm{s},k} + \widehat{F}_{\mathrm{v},k},$$

where

$$\widehat{F}_{\mathrm{v},k} = \boldsymbol{N}'_{\mathrm{v},k}[\mathbf{Var}(\boldsymbol{N}_{\mathrm{v},k})]^{-1}\boldsymbol{N}_{\mathrm{v},k} = \frac{\boldsymbol{N}'_{\mathrm{v},k}\boldsymbol{C}^{-1}_{\mathrm{v},k}\boldsymbol{N}_{\mathrm{v},k}}{\hat{\tau}^2},$$

$\widehat{F}_{\mathrm{x},k}$, $\widehat{F}_{\mathrm{s},k}$, and $\widehat{F}_{\mathrm{v},k}$ are independent of one another.

Test Statistics for Changepoint Detection

How to derive the null hypothesis distribution of $\widehat{F}$, our *maximally selected* statistic for an unknown changepoint time...

Test Statistics for Changepoint Detection

How to derive the null hypothesis distribution of $\widehat{F}$, our *maximally selected* statistic for an unknown changepoint time...

- Focus on the large sample (asymptotic) distribution;

Test Statistics for Changepoint Detection

How to derive the null hypothesis distribution of $\widehat{F}$, our *maximally selected* statistic for an unknown changepoint time...

- Focus on the large sample (asymptotic) distribution;

It follows that under H_0

$$\widehat{F} \xrightarrow{\mathcal{D}} \sup_{\ell < x < h} \{B_1(x) + B_2(x)\}$$

Test Statistics for Changepoint Detection

How to derive the null hypothesis distribution of $\widehat{F}$, our *maximally selected* statistic for an unknown changepoint time...

- Focus on the large sample (asymptotic) distribution;

It follows that under H_0

$$\widehat{F} \xrightarrow{\mathcal{D}} \sup_{\ell < x < h} \{B_1(x) + B_2(x)\}$$

Note the following about the limit distribution

- $B_1(x)$ is an ugly Gaussian process
- $B_2(x) = \mathbf{B}_{p_s+p_v}(z)'\mathbf{B}_{p_s+p_v}(z)/[z(1-z)]$
  - $\mathbf{B}_d(z)$ is a d-dimensional set of independent Brownian bridges
- The limit distribution depends only on the form of $\boldsymbol{x}_t$ and $\boldsymbol{x}_t^*$ (and the dimensionality of $\boldsymbol{s}_t$ and $\boldsymbol{v}_t$).

Autocorrelated Errors

Relax the assumption of independent errors (ϵ_t).

Autocorrelated Errors

Relax the assumption of independent errors (ϵ_t). Let

$$\tau^2 = \lim_{n\to\infty} \frac{1}{n}\text{Var}\left(\sum_{t=1}^{n} \epsilon_t\right),$$

Autocorrelated Errors

Relax the assumption of independent errors (ϵ_t). Let

$$\tau^2 = \lim_{n\to\infty} \frac{1}{n}\mathrm{Var}\left(\sum_{t=1}^{n} \epsilon_t\right),$$

The Bartlett estimator of τ^2 (requires q_n, a bandwidth parameter):

$$\hat{\tau}_B^2 = \frac{1}{n}\sum_{t=1}^{n} \hat{\epsilon}_t^2 + 2\sum_{s=1}^{q_n}\left(1 - \frac{s}{q_n + 1}\right)\frac{1}{n-s}\sum_{t=1}^{n-s} \hat{\epsilon}_t \hat{\epsilon}_{t+s},$$

Autocorrelated Errors

Relax the assumption of independent errors (ϵ_t). Let

$$\tau^2 = \lim_{n\to\infty} \frac{1}{n}\mathrm{Var}\left(\sum_{t=1}^{n} \epsilon_t\right),$$

The Bartlett estimator of τ^2 (requires q_n, a bandwidth parameter):

$$\hat{\tau}_B^2 = \frac{1}{n}\sum_{t=1}^{n} \hat{\epsilon}_t^2 + 2\sum_{s=1}^{q_n}\left(1 - \frac{s}{q_n+1}\right)\frac{1}{n-s}\sum_{t=1}^{n-s} \hat{\epsilon}_t\hat{\epsilon}_{t+s},$$

It is possible to construct versions of $\widehat{F}_{\mathrm{x},k}$, $\widehat{F}_{\mathrm{s},k}$, and $\widehat{F}_{\mathrm{v},k}$ that still use OLS residuals but encapsulate autocorrelation through a Bartlett-type variance estimator.

Autocorrelated Errors

Relax the assumption of independent errors (ϵ_t). Let

$$\tau^2 = \lim_{n\to\infty} \frac{1}{n}\text{Var}\left(\sum_{t=1}^{n} \epsilon_t\right),$$

The Bartlett estimator of τ^2 (requires q_n, a bandwidth parameter):

$$\hat{\tau}_B^2 = \frac{1}{n}\sum_{t=1}^{n} \hat{\epsilon}_t^2 + 2\sum_{s=1}^{q_n}\left(1 - \frac{s}{q_n+1}\right)\frac{1}{n-s}\sum_{t=1}^{n-s} \hat{\epsilon}_t \hat{\epsilon}_{t+s},$$

It is possible to construct versions of $\widehat{F}_{\mathrm{x},k}$, $\widehat{F}_{\mathrm{s},k}$, and $\widehat{F}_{\mathrm{v},k}$ that still use OLS residuals but encapsulate autocorrelation through a Bartlett-type variance estimator.

- In this case, $\widehat{F} = \max_{\ell<\frac{k}{n}<h}\{\widehat{F}_{\mathrm{x},k} + \widehat{F}_{\mathrm{s},k} + \widehat{F}_{\mathrm{v},k}\}$ has the same limit distribution as in the iid case.

Autocorrelated Errors

Relax the assumption of independent errors (ϵ_t). Let

$$\tau^2 = \lim_{n\to\infty} \frac{1}{n}\text{Var}\left(\sum_{t=1}^{n} \epsilon_t\right),$$

The Bartlett estimator of τ^2 (requires q_n, a bandwidth parameter):

$$\hat{\tau}_B^2 = \frac{1}{n}\sum_{t=1}^{n}\hat{\epsilon}_t^2 + 2\sum_{s=1}^{q_n}\left(1 - \frac{s}{q_n+1}\right)\frac{1}{n-s}\sum_{t=1}^{n-s}\hat{\epsilon}_t\hat{\epsilon}_{t+s},$$

It is possible to construct versions of $\widehat{F}_{\mathrm{x},k}$, $\widehat{F}_{\mathrm{s},k}$, and $\widehat{F}_{\mathrm{v},k}$ that still use OLS residuals but encapsulate autocorrelation through a Bartlett-type variance estimator.

- In this case, $\widehat{F} = \max_{\ell<\frac{k}{n}<h}\{\widehat{F}_{\mathrm{x},k} + \widehat{F}_{\mathrm{s},k} + \widehat{F}_{\mathrm{v},k}\}$ has the same limit distribution as in the iid case.
- Convergence is slow.

ARMA Models

Can assume that ϵ_t obeys an ARMA($p_{\mathrm{ar}}, q_{\mathrm{ma}}$) model:

$$\epsilon_t - \phi_1 \epsilon_{t-1} - \cdots - \phi_{p_{\mathrm{ar}}} \epsilon_{t-p_{\mathrm{ar}}} = Z_t + \theta_1 Z_{t-1} + \cdots + \theta_{q^*} Z_{t-q_{\mathrm{ma}}},$$

where $Z_t \overset{i.i.d.}{\sim} N(0, \sigma^2)$ are the ARMA errors.

ARMA Models

Can assume that ϵ_t obeys an ARMA($p_{\text{ar}}, q_{\text{ma}}$) model:

$$\epsilon_t - \phi_1 \epsilon_{t-1} - \cdots - \phi_{p_{\text{ar}}} \epsilon_{t-p_{\text{ar}}} = Z_t + \theta_1 Z_{t-1} + \cdots + \theta_{q^*} Z_{t-q_{\text{ma}}},$$

where $Z_t \stackrel{i.i.d.}{\sim} N(0, \sigma^2)$ are the ARMA errors.

We may calculate $\{\hat{Z}_t\}$, a sequence of **ARMA residuals**, via

$$\hat{Z}_t = \hat{\epsilon}_t - \hat{\phi}_1 \hat{\epsilon}_{t-1} - \cdots - \hat{\phi}_{p_{\text{ar}}} \hat{\epsilon}_{t-p_{\text{ar}}} - \hat{\theta}_1 \hat{Z}_{t-1} - \cdots - \hat{\theta}_{q_{\text{ma}}} \hat{Z}_{t-q_{\text{ma}}}.$$

ARMA Models

Can assume that ϵ_t obeys an ARMA($p_{\mathrm{ar}}, q_{\mathrm{ma}}$) model:

$$\epsilon_t - \phi_1 \epsilon_{t-1} - \cdots - \phi_{p_{\mathrm{ar}}} \epsilon_{t-p_{\mathrm{ar}}} = Z_t + \theta_1 Z_{t-1} + \cdots + \theta_{q^*} Z_{t-q_{\mathrm{ma}}},$$

where $Z_t \overset{i.i.d.}{\sim} N(0, \sigma^2)$ are the ARMA errors.

We may calculate $\{\hat{Z}_t\}$, a sequence of **ARMA residuals**, via

$$\hat{Z}_t = \hat{\epsilon}_t - \hat{\phi}_1 \hat{\epsilon}_{t-1} - \cdots - \hat{\phi}_{p_{\mathrm{ar}}} \hat{\epsilon}_{t-p_{\mathrm{ar}}} - \hat{\theta}_1 \hat{Z}_{t-1} - \cdots - \hat{\theta}_{q_{\mathrm{ma}}} \hat{Z}_{t-q_{\mathrm{ma}}}.$$

If we replace OLS residuals with ARMA residuals when calculating the test statistics, we can obtain better finite sample performance.

ARMA Models

Can assume that ϵ_t obeys an ARMA($p_{\mathrm{ar}}, q_{\mathrm{ma}}$) model:

$$\epsilon_t - \phi_1 \epsilon_{t-1} - \cdots - \phi_{p_{\mathrm{ar}}} \epsilon_{t-p_{\mathrm{ar}}} = Z_t + \theta_1 Z_{t-1} + \cdots + \theta_{q^*} Z_{t-q_{\mathrm{ma}}},$$

where $Z_t \overset{i.i.d.}{\sim} N(0, \sigma^2)$ are the ARMA errors.

We may calculate $\{\hat{Z}_t\}$, a sequence of **ARMA residuals**, via

$$\hat{Z}_t = \hat{\epsilon}_t - \hat{\phi}_1 \hat{\epsilon}_{t-1} - \cdots - \hat{\phi}_{p_{\mathrm{ar}}} \hat{\epsilon}_{t-p_{\mathrm{ar}}} - \hat{\theta}_1 \hat{Z}_{t-1} - \cdots - \hat{\theta}_{q_{\mathrm{ma}}} \hat{Z}_{t-q_{\mathrm{ma}}}.$$

If we replace OLS residuals with ARMA residuals when calculating the test statistics, we can obtain better finite sample performance.

- This is a so-called L statistic (e.g., $\widehat{L}_{\mathrm{x},k}$, $\widehat{L}_{\mathrm{s},k}$, $\widehat{L}_{\mathrm{v},k}$).

ARMA Models

Can assume that ϵ_t obeys an ARMA($p_{\mathrm{ar}}, q_{\mathrm{ma}}$) model:

$$\epsilon_t - \phi_1 \epsilon_{t-1} - \cdots - \phi_{p_{\mathrm{ar}}} \epsilon_{t-p_{\mathrm{ar}}} = Z_t + \theta_1 Z_{t-1} + \cdots + \theta_{q^*} Z_{t-q_{\mathrm{ma}}},$$

where $Z_t \overset{i.i.d.}{\sim} N(0, \sigma^2)$ are the ARMA errors.

We may calculate $\{\hat{Z}_t\}$, a sequence of **ARMA residuals**, via

$$\hat{Z}_t = \hat{\epsilon}_t - \hat{\phi}_1 \hat{\epsilon}_{t-1} - \cdots - \hat{\phi}_{p_{\mathrm{ar}}} \hat{\epsilon}_{t-p_{\mathrm{ar}}} - \hat{\theta}_1 \hat{Z}_{t-1} - \cdots - \hat{\theta}_{q_{\mathrm{ma}}} \hat{Z}_{t-q_{\mathrm{ma}}}.$$

If we replace OLS residuals with ARMA residuals when calculating the test statistics, we can obtain better finite sample performance.

- This is a so-called L statistic (e.g., $\widehat{L}_{\mathrm{x},k}$, $\widehat{L}_{\mathrm{s},k}$, $\widehat{L}_{\mathrm{v},k}$).
- The resulting quantity is underpinned by white noise components.

ARMA Models

Can assume that ϵ_t obeys an ARMA($p_{\mathrm{ar}}, q_{\mathrm{ma}}$) model:

$$\epsilon_t - \phi_1 \epsilon_{t-1} - \cdots - \phi_{p_{\mathrm{ar}}} \epsilon_{t-p_{\mathrm{ar}}} = Z_t + \theta_1 Z_{t-1} + \cdots + \theta_{q^*} Z_{t-q_{\mathrm{ma}}},$$

where $Z_t \overset{i.i.d.}{\sim} N(0, \sigma^2)$ are the ARMA errors.

We may calculate $\{\hat{Z}_t\}$, a sequence of **ARMA residuals**, via

$$\hat{Z}_t = \hat{\epsilon}_t - \hat{\phi}_1 \hat{\epsilon}_{t-1} - \cdots - \hat{\phi}_{p_{\mathrm{ar}}} \hat{\epsilon}_{t-p_{\mathrm{ar}}} - \hat{\theta}_1 \hat{Z}_{t-1} - \cdots - \hat{\theta}_{q_{\mathrm{ma}}} \hat{Z}_{t-q_{\mathrm{ma}}}.$$

If we replace OLS residuals with ARMA residuals when calculating the test statistics, we can obtain better finite sample performance.

- This is a so-called L statistic (e.g., $\widehat{L}_{\mathrm{x},k}$, $\widehat{L}_{\mathrm{s},k}$, $\widehat{L}_{\mathrm{v},k}$).
- The resulting quantity is underpinned by white noise components.
- Faster convergence

Procedures Based on ARMA residuals

Let

$$\boldsymbol{R}_{\mathrm{x},k} = \sum_{t=1}^{k} \boldsymbol{x}_t \hat{Z}_t, \qquad \left(\text{OLS:} \quad \boldsymbol{N}_{\mathrm{x},k} = \sum_{t=1}^{k} \boldsymbol{x}_t \hat{\epsilon}_t\right)$$

Procedures Based on ARMA residuals

Let

$$\boldsymbol{R}_{\mathrm{x},k} = \sum_{t=1}^{k} \boldsymbol{x}_t \hat{Z}_t, \qquad \left(\text{OLS:} \quad \boldsymbol{N}_{\mathrm{x},k} = \sum_{t=1}^{k} \boldsymbol{x}_t \hat{\epsilon}_t\right)$$

A lemma gives an asymptotic equivalence b/w $\boldsymbol{N}_{\mathrm{x},k}/\tau$ and $\boldsymbol{R}_{\mathrm{x},k}/\sigma$.

Let

$$\boldsymbol{R}_{\mathrm{x},k} = \sum_{t=1}^{k} \boldsymbol{x}_t \hat{Z}_t, \qquad \left(\text{OLS:} \;\; \boldsymbol{N}_{\mathrm{x},k} = \sum_{t=1}^{k} \boldsymbol{x}_t \hat{\epsilon}_t\right)$$

A lemma gives an asymptotic equivalence b/w $\boldsymbol{N}_{\mathrm{x},k}/\tau$ and $\boldsymbol{R}_{\mathrm{x},k}/\sigma$.

Also let

$$\widehat{L}_{\mathrm{x},k} = \frac{\boldsymbol{R}'_{\mathrm{x},k}\boldsymbol{C}^{-1}_{\mathrm{x},k}\boldsymbol{R}_{\mathrm{x},k}}{\hat{\sigma}^2}, \qquad \left(\text{OLS:} \;\; \widehat{F}_{\mathrm{x},k} = \frac{\boldsymbol{N}'_{\mathrm{x},k}\boldsymbol{C}^{-1}_{\mathrm{x},k}\boldsymbol{N}_{\mathrm{x},k}}{\hat{\tau}^2}\right)$$

Procedures Based on ARMA residuals

Let

$$\boldsymbol{R}_{\mathrm{x},k} = \sum_{t=1}^{k} \boldsymbol{x}_t \hat{Z}_t, \qquad \left(\text{OLS:} \quad \boldsymbol{N}_{\mathrm{x},k} = \sum_{t=1}^{k} \boldsymbol{x}_t \hat{\epsilon}_t\right)$$

A lemma gives an asymptotic equivalence b/w $\boldsymbol{N}_{\mathrm{x},k}/\tau$ and $\boldsymbol{R}_{\mathrm{x},k}/\sigma$.

Also let

$$\widehat{L}_{\mathrm{x},k} = \frac{\boldsymbol{R}'_{\mathrm{x},k} \boldsymbol{C}^{-1}_{\mathrm{x},k} \boldsymbol{R}_{\mathrm{x},k}}{\hat{\sigma}^2}, \qquad \left(\text{OLS:} \quad \widehat{F}_{\mathrm{x},k} = \frac{\boldsymbol{N}'_{\mathrm{x},k} \boldsymbol{C}^{-1}_{\mathrm{x},k} \boldsymbol{N}_{\mathrm{x},k}}{\hat{\tau}^2}\right)$$

Thus, $\widehat{F}_{\mathrm{x},k}$ and $\widehat{L}_{\mathrm{x},k}$ are asymptotically equivalent.

Let

$$\boldsymbol{R}_{\mathrm{s},k} = \sum_{t=1}^{k} \boldsymbol{s}_t \hat{Z}_t - \frac{k}{n} \sum_{t=1}^{n} \boldsymbol{s}_t \hat{Z}_t, \qquad \left(\text{OLS:} \quad \boldsymbol{N}_{\mathrm{s},k} = \sum_{t=1}^{k} \boldsymbol{s}_t \hat{\epsilon}_t \right)$$

Let

$$\boldsymbol{R}_{\mathrm{s},k} = \sum_{t=1}^{k} \boldsymbol{s}_t \hat{Z}_t - \frac{k}{n} \sum_{t=1}^{n} \boldsymbol{s}_t \hat{Z}_t, \qquad \left(\text{OLS:} \;\; \boldsymbol{N}_{\mathrm{s},k} = \sum_{t=1}^{k} \boldsymbol{s}_t \hat{\epsilon}_t \right)$$

We can show that $\boldsymbol{N}_{\mathrm{s},k}$ and $\boldsymbol{R}_{\mathrm{s},k}$ have the same asymptotic distribution when scaled properly.

Procedures Based on ARMA residuals

Let

$$\boldsymbol{R}_{\mathrm{s},k} = \sum_{t=1}^{k} \boldsymbol{s}_t \hat{Z}_t - \frac{k}{n} \sum_{t=1}^{n} \boldsymbol{s}_t \hat{Z}_t, \qquad \left(\text{OLS:} \quad \boldsymbol{N}_{\mathrm{s},k} = \sum_{t=1}^{k} \boldsymbol{s}_t \hat{\epsilon}_t \right)$$

We can show that $\boldsymbol{N}_{\mathrm{s},k}$ and $\boldsymbol{R}_{\mathrm{s},k}$ have the same asymptotic distribution when scaled properly.

Also let

$$\widehat{L}_{\mathrm{s},k} = \frac{\boldsymbol{R}_{\mathrm{s},k}' \mathbf{D}_T^{-1} \boldsymbol{R}_{\mathrm{s},k}}{\hat{\sigma}^2 k (1 - \frac{k}{n})}, \qquad \left(\text{OLS:} \quad \widehat{F}_{\mathrm{s},k} = \boldsymbol{N}_{\mathrm{s},k}' \widehat{\boldsymbol{\tau}}_s^{-1} \boldsymbol{N}_{\mathrm{s},k} \right)$$

Let

$$\boldsymbol{R}_{\mathrm{s},k} = \sum_{t=1}^{k} \boldsymbol{s}_t \hat{Z}_t - \frac{k}{n} \sum_{t=1}^{n} \boldsymbol{s}_t \hat{Z}_t, \qquad \left(\text{OLS:} \quad \boldsymbol{N}_{\mathrm{s},k} = \sum_{t=1}^{k} \boldsymbol{s}_t \hat{\epsilon}_t \right)$$

We can show that $\boldsymbol{N}_{\mathrm{s},k}$ and $\boldsymbol{R}_{\mathrm{s},k}$ have the same asymptotic distribution when scaled properly.

Also let

$$\widehat{L}_{\mathrm{s},k} = \frac{\boldsymbol{R}'_{\mathrm{s},k} \mathbf{D}_T^{-1} \boldsymbol{R}_{\mathrm{s},k}}{\hat{\sigma}^2 k (1 - \frac{k}{n})}, \qquad \left(\text{OLS:} \quad \widehat{F}_{\mathrm{s},k} = \boldsymbol{N}'_{\mathrm{s},k} \widehat{\boldsymbol{\tau}}_s^{-1} \boldsymbol{N}_{\mathrm{s},k} \right)$$

It follows that $\widehat{F}_{\mathrm{s},k}$ and $\widehat{L}_{\mathrm{s},k}$ have the same asymptotic distribution.

Procedures Based on ARMA residuals

Let

$$\boldsymbol{R}_{\mathrm{v},k} = \sum_{t=1}^{k} \boldsymbol{v}_t \hat{Z}_t - \frac{k}{n} \sum_{t=1}^{n} \boldsymbol{v}_t \hat{Z}_t, \qquad \left(\text{OLS:} \quad \boldsymbol{N}_{\mathrm{v},k} = \sum_{t=1}^{k} \boldsymbol{v}_t \hat{\epsilon}_t \right)$$

Procedures Based on ARMA residuals

Let

$$\boldsymbol{R}_{\mathrm{v},k} = \sum_{t=1}^{k} \boldsymbol{v}_t \hat{Z}_t - \frac{k}{n} \sum_{t=1}^{n} \boldsymbol{v}_t \hat{Z}_t, \qquad \left(\text{OLS:} \quad \boldsymbol{N}_{\mathrm{v},k} = \sum_{t=1}^{k} \boldsymbol{v}_t \hat{\epsilon}_t \right)$$

We can show that $\boldsymbol{N}_{\mathrm{v},k}$ and $\boldsymbol{R}_{\mathrm{v},k}$ have the same asymptotic distribution when scaled properly.

Procedures Based on ARMA residuals

Let

$$\boldsymbol{R}_{\mathrm{v},k} = \sum_{t=1}^{k} \boldsymbol{v}_t \hat{Z}_t - \frac{k}{n} \sum_{t=1}^{n} \boldsymbol{v}_t \hat{Z}_t, \qquad \left(\text{OLS:} \quad \boldsymbol{N}_{\mathrm{v},k} = \sum_{t=1}^{k} \boldsymbol{v}_t \hat{\epsilon}_t \right)$$

We can show that $\boldsymbol{N}_{\mathrm{v},k}$ and $\boldsymbol{R}_{\mathrm{v},k}$ have the same asymptotic distribution when scaled properly.

Also let

$$\widehat{L}_{\mathrm{v},k} = \frac{\boldsymbol{R}'_{\mathrm{v},k} \widehat{\boldsymbol{\Sigma}}_v^{-1} \boldsymbol{R}_{\mathrm{v},k}}{\hat{\sigma}^2 k(1 - \frac{k}{n})}, \qquad \left(\text{OLS:} \quad \widehat{F}_{\mathrm{v},k} = \boldsymbol{N}'_{\mathrm{v},k} \widehat{\boldsymbol{\tau}}_v^{-1} \boldsymbol{N}_{\mathrm{v},k} \right)$$

Let

$$\boldsymbol{R}_{\mathrm{v},k} = \sum_{t=1}^{k} \boldsymbol{v}_t \hat{Z}_t - \frac{k}{n} \sum_{t=1}^{n} \boldsymbol{v}_t \hat{Z}_t, \qquad \left(\text{OLS:} \quad \boldsymbol{N}_{\mathrm{v},k} = \sum_{t=1}^{k} \boldsymbol{v}_t \hat{\epsilon}_t \right)$$

We can show that $\boldsymbol{N}_{\mathrm{v},k}$ and $\boldsymbol{R}_{\mathrm{v},k}$ have the same asymptotic distribution when scaled properly.

Also let

$$\widehat{L}_{\mathrm{v},k} = \frac{\boldsymbol{R}'_{\mathrm{v},k} \widehat{\boldsymbol{\Sigma}}_v^{-1} \boldsymbol{R}_{\mathrm{v},k}}{\hat{\sigma}^2 k(1 - \frac{k}{n})}, \qquad \left(\text{OLS:} \quad \widehat{F}_{\mathrm{v},k} = \boldsymbol{N}'_{\mathrm{v},k} \widehat{\boldsymbol{\tau}}_v^{-1} \boldsymbol{N}_{\mathrm{v},k} \right)$$

It follows that $\widehat{F}_{\mathrm{v},k}$ and $\widehat{L}_{\mathrm{v},k}$ have the same asymptotic distribution.

Procedures Based on ARMA residuals

An omnibus test based on ARMA residuals for a changepoint at time k is

$$\widehat{L}_k = \widehat{L}_{\mathrm{x},k} + \widehat{L}_{\mathrm{s},k} + \widehat{L}_{\mathrm{v},k}$$

Procedures Based on ARMA residuals

An omnibus test based on ARMA residuals for a changepoint at time k is

$$\widehat{L}_k = \widehat{L}_{\mathrm{x},k} + \widehat{L}_{\mathrm{s},k} + \widehat{L}_{\mathrm{v},k}$$

To detect a change at an *unknown* time, we consider

$$\widehat{L} = \max_{\ell \leq \frac{k}{n} \leq h} \widehat{L}_k$$

Procedures Based on ARMA residuals

An omnibus test based on ARMA residuals for a changepoint at time k is

$$\widehat{L}_k = \widehat{L}_{\mathrm{x},k} + \widehat{L}_{\mathrm{s},k} + \widehat{L}_{\mathrm{v},k}$$

To detect a change at an *unknown* time, we consider

$$\widehat{L} = \max_{\ell \le \frac{k}{n} \le h} \widehat{L}_k$$

It follows that under H_0

$$\widehat{L} \xrightarrow{\mathcal{D}} \sup_{\ell < x < h} \{B_1(x) + B_2(x)\},$$

which is the limit process that was observed by the statistic $\widehat{F}$.

Simulations

The efficacy of the proposed methodology is studied via simulation.

Simulations

The efficacy of the proposed methodology is studied via simulation.

Here, we isolate to the following predictor model:

$$y_t = \alpha_0 + \alpha_1 \left(\frac{t}{n}\right) + \alpha_2 \left(\frac{t}{n}\right)^2 + \gamma_1 \cos\left(\frac{2\pi t}{12}\right) + \gamma_2 \sin\left(\frac{2\pi t}{12}\right) + \zeta C_t + \epsilon_t$$

Simulations

The efficacy of the proposed methodology is studied via simulation.

Here, we isolate to the following predictor model:

$$y_t = \alpha_0 + \alpha_1 \left(\frac{t}{n}\right) + \alpha_2 \left(\frac{t}{n}\right)^2 + \gamma_1 \cos\left(\frac{2\pi t}{12}\right) + \gamma_2 \sin\left(\frac{2\pi t}{12}\right) + \zeta C_t + \epsilon_t$$

Values of parameters are chosen to coincide with those estimated under H_0 for the CO_2 data example.

The efficacy of the proposed methodology is studied via simulation.

Here, we isolate to the following predictor model:

$$y_t = \alpha_0 + \alpha_1 \left(\frac{t}{n}\right) + \alpha_2 \left(\frac{t}{n}\right)^2 + \gamma_1 \cos\left(\frac{2\pi t}{12}\right) + \gamma_2 \sin\left(\frac{2\pi t}{12}\right) + \zeta C_t + \epsilon_t$$

Values of parameters are chosen to coincide with those estimated under H_0 for the CO_2 data example.

The error sequence $\{\epsilon_t\}$ is generated via an AR(1):

$$\epsilon_t = \phi_1 \epsilon_{t-1} + Z_t$$

Simulations

Four settings are examined (fix $n = 1000$ with $c = 500$ under H_1):

- Setting 1: **All** regression coefficients may change under H_1.
- Setting 2: Only coefficients governing trend may change.
- Setting 3: Only coefficients governing seasonality may change.
- Setting 4: Only coefficients governing covariates may change.

Simulations

Four settings are examined (fix $n = 1000$ with $c = 500$ under H_1):

- Setting 1: **All** regression coefficients may change under H_1.
- Setting 2: Only coefficients governing trend may change.
- Setting 3: Only coefficients governing seasonality may change.
- Setting 4: Only coefficients governing covariates may change.

We examine type I error as a function of ϕ, the AR(1) parameter.

Simulations

Four settings are examined (fix $n = 1000$ with $c = 500$ under H_1):

- Setting 1: **All** regression coefficients may change under H_1.
- Setting 2: Only coefficients governing trend may change.
- Setting 3: Only coefficients governing seasonality may change.
- Setting 4: Only coefficients governing covariates may change.

We examine type I error as a function of ϕ, the AR(1) parameter.

We examine power as a function of δ, where $\boldsymbol{\Delta} = \delta \boldsymbol{d}$, where $\boldsymbol{d}$ is an arbitrary vector (we fix $\phi = 0.5$).

Simulations

Four settings are examined (fix $n = 1000$ with $c = 500$ under H_1):

- Setting 1: **All** regression coefficients may change under H_1.
- Setting 2: Only coefficients governing trend may change.
- Setting 3: Only coefficients governing seasonality may change.
- Setting 4: Only coefficients governing covariates may change.

We examine type I error as a function of ϕ, the AR(1) parameter.

We examine power as a function of δ, where $\boldsymbol{\Delta} = \delta \boldsymbol{d}$, where $\boldsymbol{d}$ is an arbitrary vector (we fix $\phi = 0.5$).

We calculate two types of statistics:

- Procedures based on OLS residuals ($\widehat{F}$);
  - Autocorrelation is captured through Bartlett-based estimators.

Simulations

Four settings are examined (fix $n = 1000$ with $c = 500$ under H_1):

- Setting 1: **All** regression coefficients may change under H_1.
- Setting 2: Only coefficients governing trend may change.
- Setting 3: Only coefficients governing seasonality may change.
- Setting 4: Only coefficients governing covariates may change.

We examine type I error as a function of ϕ, the AR(1) parameter.

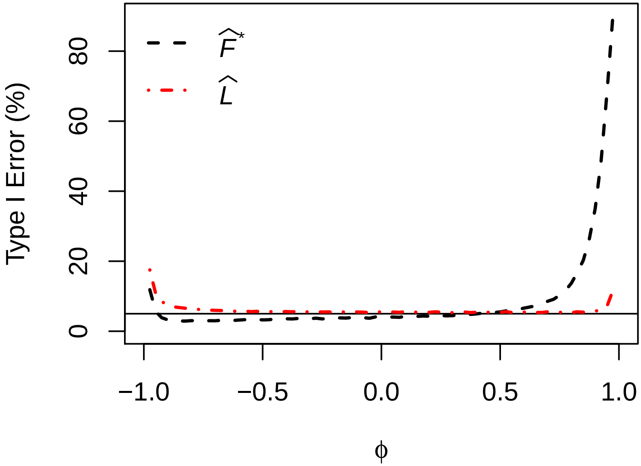
We examine power as a function of δ, where $\boldsymbol{\Delta} = \delta \boldsymbol{d}$, where $\boldsymbol{d}$ is an arbitrary vector (we fix $\phi = 0.5$).

We calculate two types of statistics:

- Procedures based on OLS residuals ($\widehat{F}$);
  - Autocorrelation is captured through Bartlett-based estimators.
- Procedures based on ARMA residuals ($\widehat{L}$).

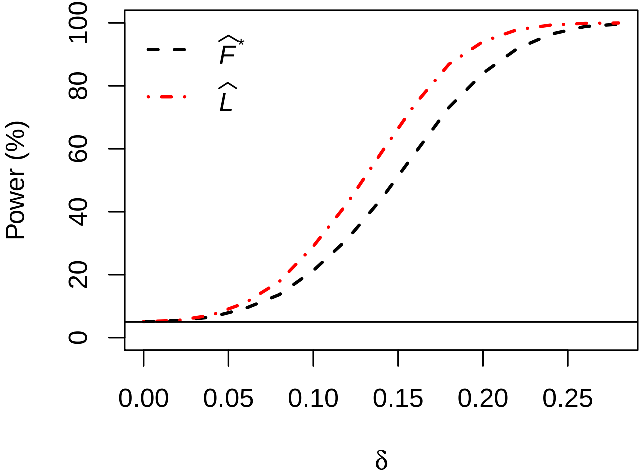
Simulations

Setting 1: **All** regression coefficients may change under H_1.

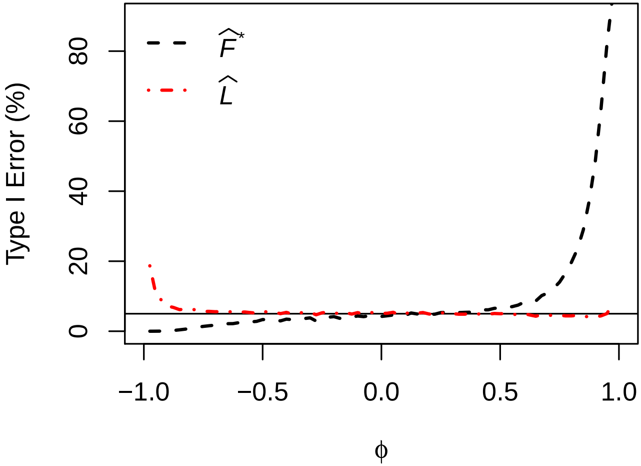


Simulations

Setting 1: **All** regression coefficients may change under H_1.

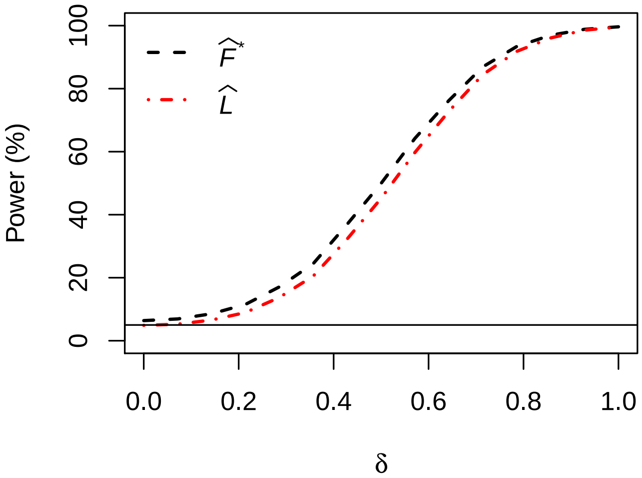


Setting 2: Only coefficients governing trend may change.

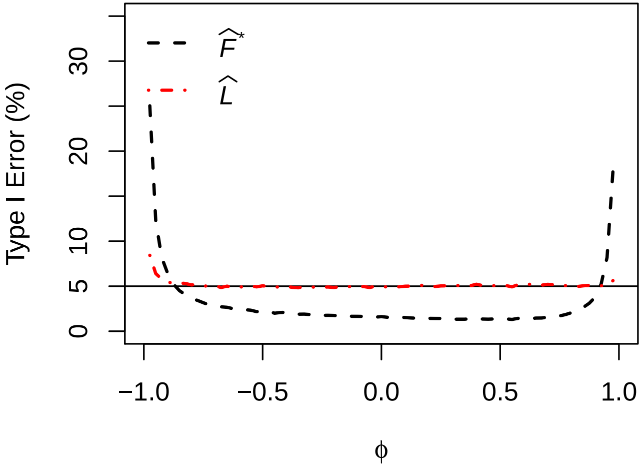


Simulations

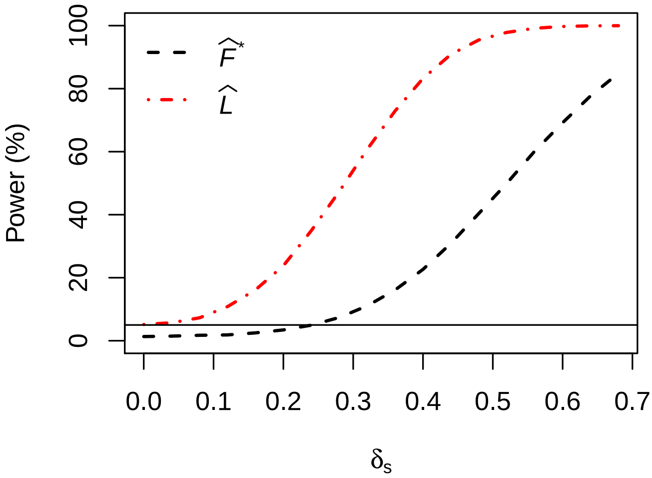
Setting 2: Only coefficients governing trend may change.



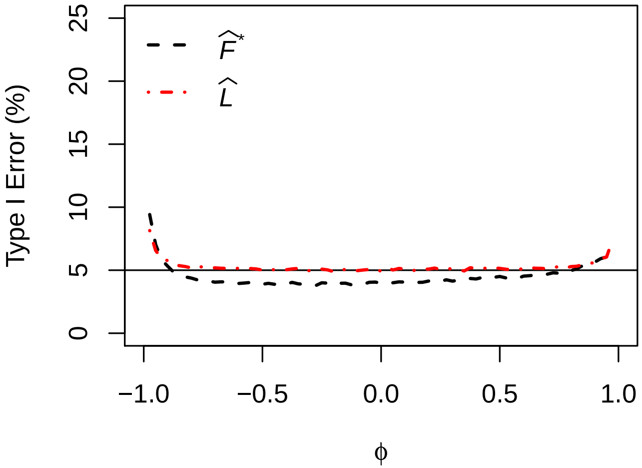
Setting 3: Only coefficients governing seasonality may change.



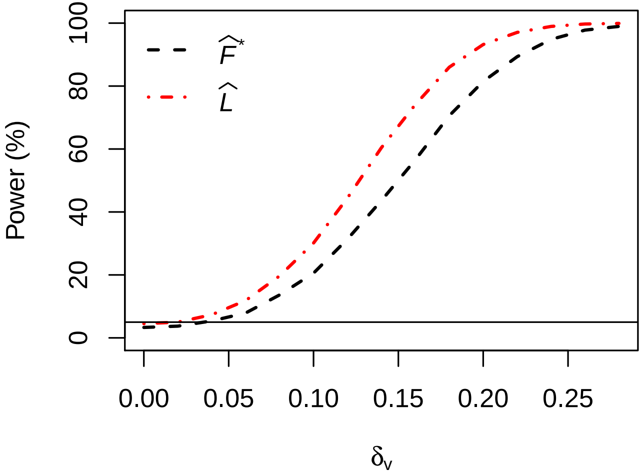
Setting 3: Only coefficients governing seasonality may change.



Setting 4: Only coefficients governing covariates may change.



Setting 4: Only coefficients governing covariates may change.



Fit an AR(p_{ar}) model to the OLS residuals of the CO_2 data.

Setting 2: Only coefficients governing trend change.

$\widehat{F}$				$\widehat{L}$			
q_n	$\hat{c}$	Stat.	p-value	p_{ar}	$\hat{c}$	Stat.	p-value
2	1991	155.1	0.0000	2	1991	24.1	0.0027
4	1991	97.5	0.0000	4	1991	20.3	0.0117
8	1991	57.5	0.0000	8	1991	18.9	0.0223
12	1991	41.2	0.0000	12	1991	20.5	0.0110
16	1991	32.3	0.0001	16	1991	16.3	0.0587
24	1991	23.3	0.0037	24	1991	17.7	0.0325

Table: Results of the tests for values of the autoregressive order (p_{ar}) and the lag cut-off for the Bartlett estimator (q_n).

Application to the CO_2 Data

Fit an AR(p_{ar}) model to the OLS residuals of the CO_2 data.

Setting 3: Only coefficients governing seasonality change.

$\widehat{F}$				$\widehat{L}$			
q_n	$\hat{c}$	Stat.	p-value	p_{ar}	$\hat{c}$	Stat.	p-value
2	1973	24.7	0.0520	2	1976	60.2	0.0000
4	1973	20.8	0.1706	4	1976	62.2	0.0000
8	1973	17.9	0.3610	8	1976	61.3	0.0000
12	2011	22.2	0.1132	12	1976	54.7	0.0000
16	2011	41.7	0.0001	16	1976	41.0	0.0001
24	2012	33.7	0.0018	24	1976	39.8	0.0002

Table: Results of the tests for values of the autoregressive order (p_{ar}) and the lag cut-off for the Bartlett estimator (q_n).

Application to the CO_2 Data

Fit an AR(p_{ar}) model to the OLS residuals of the CO_2 data.

Setting 3: Only coefficients governing seasonality change.

$\widehat{F}$				$\widehat{L}$			
q_n	$\hat{c}$	Stat.	p-value	p_{ar}	$\hat{c}$	Stat.	p-value
2	1973	24.7	0.0520	2	1976	60.2	0.0000
4	1973	20.8	0.1706	4	1976	62.2	0.0000
8	1973	17.9	0.3610	8	1976	61.3	0.0000
12	2011	22.2	0.1132	12	1976	54.7	0.0000
16	2011	41.7	0.0001	16	1976	41.0	0.0001
24	2012	33.7	0.0018	24	1976	39.8	0.0002

Table: Results of the tests for values of the autoregressive order (p_{ar}) and the lag cut-off for the Bartlett estimator (q_n).

*Note: We assume a change in trend occurred in 1991

Application to the CO_2 Data

Fit an AR(p_{ar}) model to the OLS residuals of the CO_2 data.

Setting 4: Only coefficients governing covariates change.

$\widehat{F}$				$\widehat{L}$			
q_n	$\hat{c}$	Stat.	p-value	p_{ar}	$\hat{c}$	Stat.	p-value
2	1996	4.0	0.5016	2	2006	3.7	0.5529
4	1996	3.0	0.6911	4	2006	3.7	0.5462
8	1996	2.4	0.8193	8	2006	4.2	0.4696
12	1996	2.4	0.8360	12	2006	4.7	0.3860
16	1996	2.4	0.8332	16	2010	4.5	0.4191
24	1996	2.2	0.8681	24	2010	4.9	0.3509

Table: Results of the tests for values of the autoregressive order (p_{ar}) and the lag cut-off for the Bartlett estimator (q_n).

*Note: We assume a change in trend occurred in 1991

Application to the CO_2 Data

Fit an AR(p_{ar}) model to the OLS residuals of the CO_2 data.

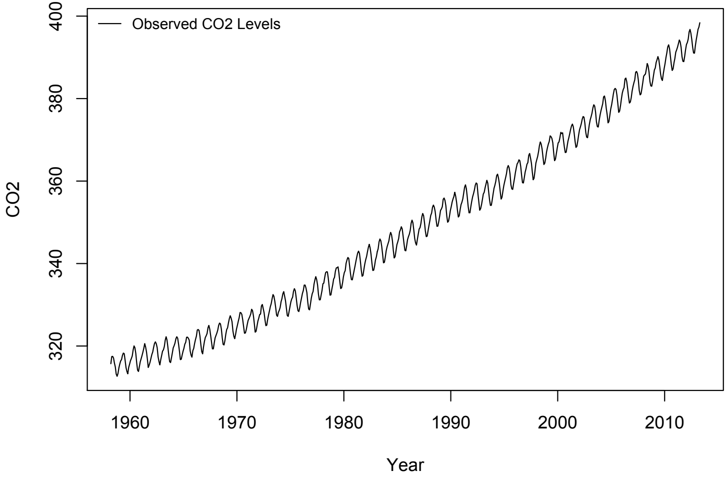
Setting 1: All regression coefficients change under H_1

	$\widehat{F}$				$\widehat{L}$		
q_n	$\hat{c}$	Stat.	p-value	p_{ar}	$\hat{c}$	Stat.	p-value
2	1991	174.6	0.0000	2	1977	77.8	0.0000
4	1988	114.8	0.0000	4	1976	84.6	0.0000
8	1988	77.0	0.0000	8	1977	78.2	0.0000
12	1988	60.2	0.0000	12	1977	61.9	0.0000
16	1989	54.5	0.0000	16	1976	47.9	0.0004
24	1988	56.5	0.0000	24	1977	46.0	0.0006

Table: Results of the tests for values of the autoregressive order (p_{ar}) and the lag cut-off for the Bartlett estimator (q_n).

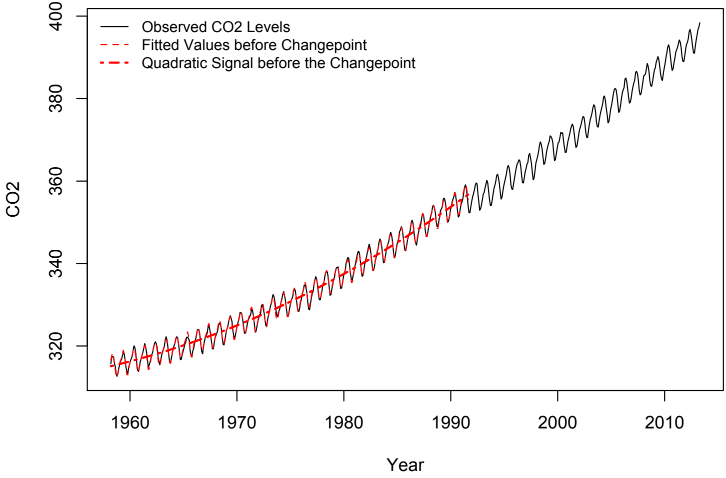
Application to the CO_2 Data

The Mauna Loa CO_2 data:



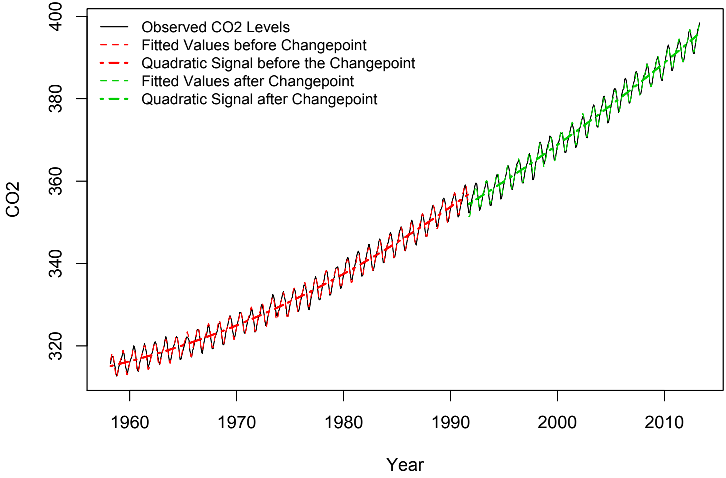
Application to the CO_2 Data

The Mauna Loa CO_2 data:



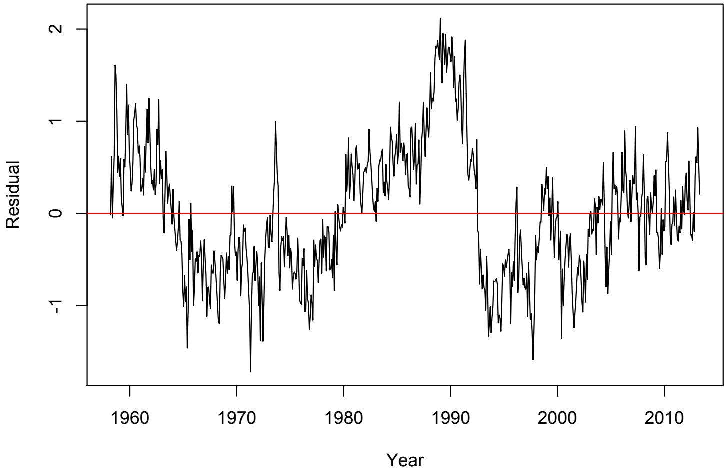
Application to the CO_2 Data

The Mauna Loa CO_2 data:



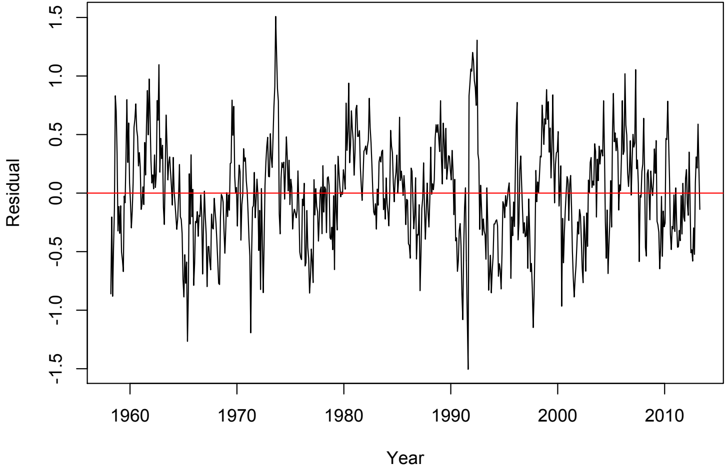
Application to the CO_2 Data

The OLS residuals without accounting for changepoint:



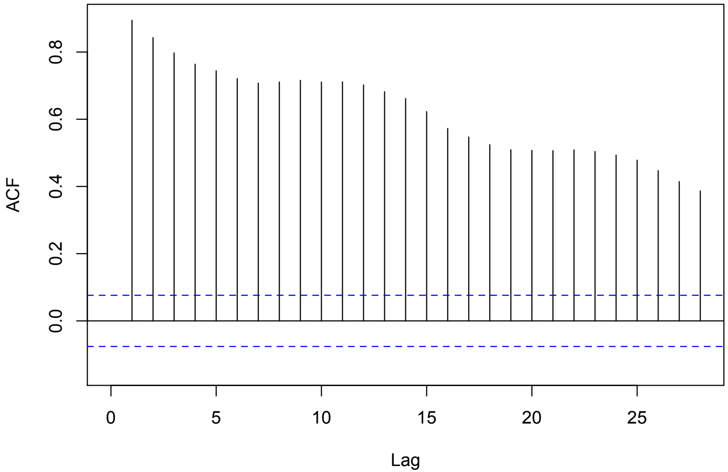
Application to the CO_2 Data

The OLS residuals while accounting for changepoint:



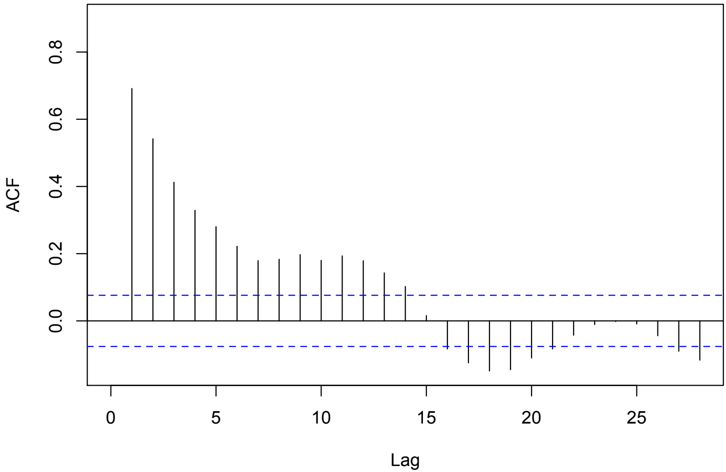
Application to the CO_2 Data

The ACF of the OLS residuals without accounting for changepoint:



Application to the CO_2 Data

The ACF of the OLS residuals while accounting for changepoint:



Why the change in trend???

On June 15, 1991, Mount Pinatubo (in the Philippines) erupted:

Why the change in trend???

On June 15, 1991, Mount Pinatubo (in the Philippines) erupted:

- Pumped tons of sulfur (SO_2) into the stratosphere;
- Less sunlight reached the ocean surface;
- Cooler sea surface temperature meant more CO_2 is absorbed by the ocean.

Why the change in trend???

On June 15, 1991, Mount Pinatubo (in the Philippines) erupted:

- Pumped tons of sulfur (SO_2) into the stratosphere;
- Less sunlight reached the ocean surface;
- Cooler sea surface temperature meant more CO_2 is absorbed by the ocean.

Mauna Loa is a volcano itself:

Why the change in trend???

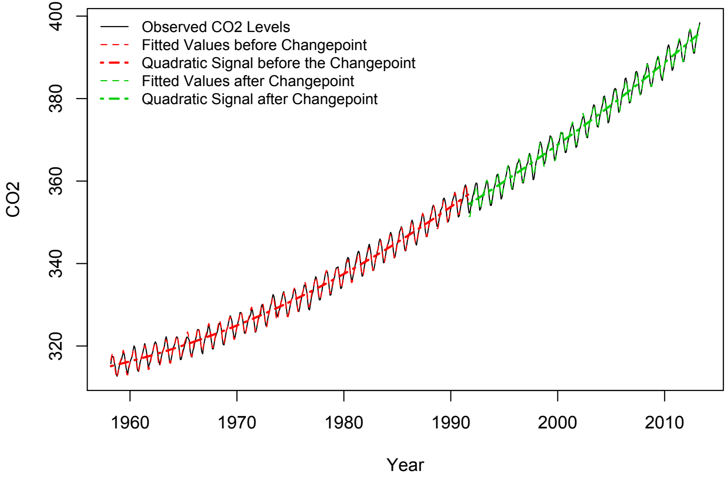
On June 15, 1991, Mount Pinatubo (in the Philippines) erupted:

- Pumped tons of sulfur (SO_2) into the stratosphere;
- Less sunlight reached the ocean surface;
- Cooler sea surface temperature meant more CO_2 is absorbed by the ocean.

Mauna Loa is a volcano itself:

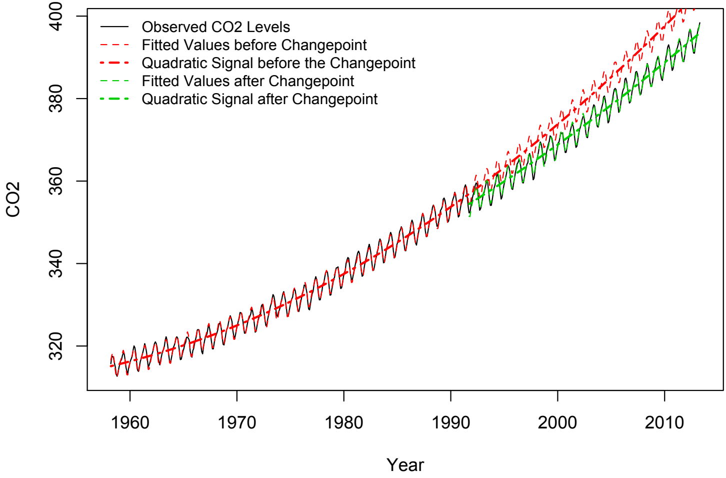
- Weak explosions
- Particles do not reach the stratosphere;
- No major effect on CO_2.

The Mauna Loa CO_2 data:



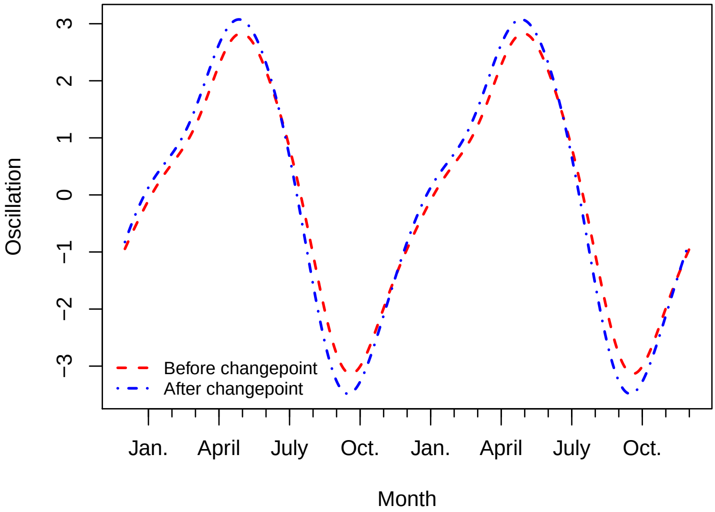
Application to the CO_2 Data

The Mauna Loa CO_2 data:



The 1976 changepoint in seasonality

The seasonal pattern before and after the 1976 changepoint:



Why the change in seasonality???

Seasonality in CO_2 is due to vegetation growth:

Why the change in seasonality???

Seasonality in CO_2 is due to vegetation growth:

Why are the oscillations changing?

- Increased prominence of droughts (due to global warming)
- Increased levels of vegetation (due to higher CO_2 levels)
- Increased prominence of agriculture

Why the change in seasonality???

Seasonality in CO_2 is due to vegetation growth:

Why are the oscillations changing?

- Increased prominence of droughts (due to global warming)
- Increased levels of vegetation (due to higher CO_2 levels)
- Increased prominence of agriculture

Seasonal amplitude is likely steadily increasing (no true changepoint)

Why the change in seasonality???

Seasonality in CO_2 is due to vegetation growth:

Why are the oscillations changing?

- Increased prominence of droughts (due to global warming)
- Increased levels of vegetation (due to higher CO_2 levels)
- Increased prominence of agriculture

Seasonal amplitude is likely steadily increasing (no true changepoint)

Changes in trend and seasonality caused by separate environmental mechanisms

- Need to test for the different types of changes separately

References

Changes in trend only:

- Robbins, M. W., C. M. Gallagher and R. B. Lund (2016) "A General Regression Changepoint Test for Time Series Data." *Journal of the American Statistical Association*. Forthcoming in the June issue

Changes in all regression coefficients (trend, seasonality, covariates):

- Robbins, M. W. (2016) "A Fully Flexible Changepoint Test for Regression Models with Stationary Errors." Submitted.

`RegCpt`: A R package that implements the method

- In development

E-mail: `mrobbins@rand.org`