Inference for multiple change-points in time series via scan statistics

Chun Yip Yau
Chinese University of Hong Kong

Joint with Zifeng Zhao (Univ. of Wisconsin-Madison)
Research supported in part by HKSAR-RGC-GRF
1 Motivation

2 Likelihood Ratio Scan for Change-Points Detection
   • Change-point Detection by Scan Statistic
   • Consistent Change-point Estimation by Model Selection
   • Construction of Confidence Intervals

3 Implementation Issue and Computational Complexity

4 Simulation Studies

5 Applications

6 Conclusions
Motivation

- Where are the change-points?

![Graphs showing time series data with change-points identified.](image-url)
Motivation

- Where are the change-points?

- Literatures:

\[ |\hat{\tau} - \tau_0| = O_p(1). \]

⇒

To locate a change-point,
We don’t really need *too many* data.
Change-point Estimation in Local Windows

Toy example:

\[ X_t = \begin{cases} 
    e_t & \text{if } t < 100 \\
    1 + e_t & \text{if } t \geq 100 
\end{cases}, \quad e_t \sim i.i.d. \ N(0, 1). \]

- Observations: \( \{X_1, X_2, \ldots, X_{200}\} \).
- Estimations:
  For \( Y = X_{91:110}; X_{81:120}; X_{71:140}; \ldots; X_{11:190}; X_{1:200} \)
  Use CUSUM statistic

\[ \hat{\tau} = \arg \max_k \left( \sum_{j=1}^{k} y_j - k \bar{y} \right). \]
Findings:

- In 80/100 replications, the 10 $\hat{\tau}$s are exactly the same.
- For the other 20 replications, the average s.d. of the 10 $\hat{\tau}$s is 3.43.
Main Objectives:

Literatures:
- Scan Statistics: Naus (1965), Glaz (2001) …
- Testing for Change-point: Bauer & Hackl (78, 80), Chan & Walther (13) …
- Location estimation: Fryzlewicz (14), Dette (14)

Objectives: By using a Likelihood Ratio Scan Statistic,
- Consistent estimation of multiple change-points in time series.
- Construction of confidence intervals.
- Fast and easy to implement
Content

1 Motivation

2 Likelihood Ratio Scan for Change-Points Detection
   • Change-point Detection by Scan Statistic
   • Consistent Change-point Estimation by Model Selection
   • Construction of Confidence Intervals

3 Implementation Issue and Computational Complexity

4 Simulation Studies

5 Applications

6 Conclusions
1. Motivation

2. Likelihood Ratio Scan for Change-Points Detection
   - Change-point Detection by Scan Statistic
   - Consistent Change-point Estimation by Model Selection
   - Construction of Confidence Intervals

3. Implementation Issue and Computational Complexity

4. Simulation Studies

5. Applications

6. Conclusions
Data: Piecewise stationary autoregressive time series \( \{X_t\}_{t=1}^n \)

The \( j \)-th stationary segment: \( Y_{t,j}, \quad \tau_{j-1} < t \leq \tau_j \)

\[
Y_{t,j} = \phi_{j0} + \phi_{j1}Y_{t-1,j} + \cdots + \phi_{jp_j}Y_{t-p_j,j} + \sigma_j \epsilon_t, \quad \epsilon_t \sim \text{i.i.d. } N(0,1).
\]

Parameters:

- Break-points: \( 0 = \tau_0, \tau_1, \tau_2, \ldots, \tau_m, \tau_{m+1} = n \)
- Relative location of breaks: \( \lambda_j := \frac{\tau_j}{n} \). Assume \( |\lambda_{j+1} - \lambda_j| > \epsilon_\lambda \).
- AR parameter vectors: \( \theta_j := (\phi_{j,0}, \phi_{j,1}, \ldots, \phi_{j,p_j}, \sigma_j^2) \).
1 Motivation

2 Likelihood Ratio Scan for Change-Points Detection
   - Change-point Detection by Scan Statistic
   - Consistent Change-point Estimation by Model Selection
   - Construction of Confidence Intervals

3 Implementation Issue and Computational Complexity

4 Simulation Studies

5 Applications

6 Conclusions
The Scan Statistic: \( h < \min(\lambda_i - \lambda_{i-1})/4 \)

\[
S_h(t) = \log \frac{\hat{L}(t - h + 1 : t)\hat{L}(t + 1 : t + h)}{\hat{L}(t - h + 1 : t + h)}
\]
Likelihood Ratio Scan Statistic (LRSS)

- Conditional log-Likelihood of AR model for data $y = (y_1, \ldots, y_n)$:

  $$
  l(\theta, y) = -\sum \frac{(y_t - \phi_0 - \phi_1 y_{t-1} - \cdots - \phi_p y_{t-p})^2}{2\sigma^2} - \frac{1}{2} \log \sigma^2
  $$

  $$
  l(\hat{\theta}, y) = -\frac{n}{2} - \frac{1}{2} \log \hat{\sigma}^2,
  $$

  where $\hat{\sigma}^2 = \hat{\gamma}(0) - \hat{\gamma}^T \hat{\phi}$, $\hat{\phi} = \hat{\Gamma}^{-1} \hat{\gamma}$.

- Scan Statistics

  $$
  S_h(t) = \log \hat{\sigma}_t^2 - \frac{1}{2} \log \hat{\sigma}_{1,t}^2 - \frac{1}{2} \log \hat{\sigma}_{2,t}^2.
  $$
The set of Local Change-Point Estimators:

$$\hat{L}^{(1)} = \left\{ t : S_h(t) = \max_{u \in [t-h, t+h]} S_h(u) \right\}$$
Theorem 1

Let

- True CP: $\mathcal{L}_o = (\tau^o_1, \ldots, \tau^o_{m_o})$
- Local CP estimates: $\hat{\mathcal{L}}^{(1)} = \{\hat{\tau}^{(1)}_1, \hat{\tau}^{(1)}_2, \ldots, \hat{\tau}^{(1)}_{\hat{m}(1)}\}$
- $\min_{j=1,\ldots,m_o}(\lambda^o_{j+1} - \lambda^o_j) > \epsilon\lambda$

There exists some $d > 0$ such that for $h \geq d \log n$ and $h < n\epsilon\lambda/4$,

$$\mathbb{P} \left( \max_{\tau^o_j \in \mathcal{L}_0} \min_{\hat{\tau}^{(1)}_k \in \hat{\mathcal{L}}^{(1)}} |\tau^o_j - \hat{\tau}^{(1)}_k| < h \right) \to 1.$$

- Possibly overestimation: $\hat{m}(1) > m_o$
- Each of the true $\tau^o_j$ is surrounded by a $\hat{\tau}^{(1)}_k$ in a $h$-neighborhood.
Outline

1 Motivation

2 Likelihood Ratio Scan for Change-Points Detection
   • Change-point Detection by Scan Statistic
   • Consistent Change-point Estimation by Model Selection
   • Construction of Confidence Intervals

3 Implementation Issue and Computational Complexity

4 Simulation Studies

5 Applications

6 Conclusions
Estimation of Change-points

From the scanning step, we have $\hat{L}^{(1)} = \{\hat{\tau}_1^{(1)}, \hat{\tau}_2^{(1)}, \ldots, \hat{\tau}_{\hat{m}}^{(1)}\}$ with

$$P \left( \max_{\tau_o^j \in L_0} \min_{\hat{\tau}_k^{(1)} \in \hat{L}^{(1)}} |\tau_o^j - \hat{\tau}_k^{(1)}| < h \right) \to 1.$$  

- **Infeasible:** $(\hat{m}, \hat{L}, \hat{p}) = \arg \min_{L \subset \{1,2,\ldots,n\}} IC(m, L, p)$
- **Feasible:** $(\hat{m}^{(2)}, \hat{L}^{(2)}, \hat{p}^{(2)}) = \arg \min_{L \subset \hat{L}^{(1)}} IC(m, L, p)$
- **Examples:**
  - $AIC(m, L, p) = \sum_{j=1}^{m+1} \frac{n_j}{2} \log(\hat{\sigma}_j^2) + 2(m + \sum_{j=1}^{m+1} (p_j + 2))$
  - $BIC(m, L, p) = \sum_{j=1}^{m+1} \frac{n_j}{2} \log(\hat{\sigma}_j^2) + (m + \sum_{j=1}^{m+1} (p_j + 2)) \log n$
  - $MDL(m, L, p) = \sum_{j=1}^{m+1} \frac{n_j}{2} \log(\hat{\sigma}_j^2) + \sum_{j=1}^{m+1} \frac{p_j + 2}{2} \log(n_j) + (m + 1) \log(n)$
Theorem 2

**Under the setting of Theorem 1, if**

\[ IC(m, L, p) = \sum_{j=1}^{m+1} \frac{n_j}{2} \log(\hat{\sigma}_j^2) + \omega_n(m + \sum_{j=1}^{m+1} p_j) \]

**with** \( \omega_n = o(n) \) **and** \( \frac{\omega_n}{\log \log n} \to \infty \), **then**

\[ \hat{m}^{(2)}_p \to m_o \), \quad P \left( \max_{j=1, \ldots, m_o} |\hat{\tau}_j^{(2)} - \tau^o_j| < h \right) \to 1, \quad \max_{j=1, \ldots, m_o} |\hat{p}_j^{(2)} - p^o_j| \xrightarrow{p} 0. \]

- \( h > d \log n \to \infty \Rightarrow \) no guarantee that \( |\hat{\tau}_j^{(2)} - \tau^o_j| = O_p(1) \).
- \( \tau^o_j \in \left( \hat{\tau}_j^{(2)} - h, \hat{\tau}_j^{(2)} + h \right] \Rightarrow \) improvement by extended local window.
Outline

1. Motivation

2. Likelihood Ratio Scan for Change-Points Detection
   - Change-point Detection by Scan Statistic
   - Consistent Change-point Estimation by Model Selection
   - Construction of Confidence Intervals

3. Implementation Issue and Computational Complexity

4. Simulation Studies

5. Applications

6. Conclusions
Final Estimates and Confidence Intervals

- From the model selection step we have

\[ P \left( \tau^o_j \in \left( \hat{\tau}^{(2)}_j - h, \hat{\tau}^{(2)}_j + h \right) \right) \to 1 \]

- Define the extended local window

\[ E_j = \left( \hat{\tau}^{(2)}_j - 2h, \hat{\tau}^{(2)}_j + 2h \right) \]

⇒ \( \tau^o_j \) is within \( \left( \frac{1}{4}, \frac{3}{4} \right) \) of \( E_j \).

- \( E_j \) contains one \( \tau^o_j \)
- Number of observation before/after \( \tau^o_j \to \infty \)
⇒ Final estimate and C.I. can be obtained from \( E_j \)
Final estimates and Confidence Intervals

Final Estimates of Change Points

$$\hat{\tau}^{(3)}_j = \arg \max_{\tau \in (\hat{\tau}^{(2)}_j - h, \hat{\tau}^{(2)}_j + h]} \log L(\hat{\tau}^{(2)}_j - 2h + 1 : \tau) L(\tau + 1 : \hat{\tau}^{(2)}_j + 2h)$$

Theorem 3

Under the setting of Theorem 1 and 2,

$$\max_{j=1,\ldots,m_o} |\hat{\tau}^{(3)}_j - \tau^o_j| = O_p(1).$$
Theorem 4

Asymptotic distribution of Change Points (Ling (2013))

\[ \hat{\tau}_j^{(3)} - \tau_j^o \xrightarrow{d} \arg \max_k W_k, \]

where

\[ W_k = \begin{cases} 
\sum_{t=\tau_j^o}^{\tau_j^o+k} (l_t(\theta_1^o) - l_t(\theta_2^o)) & k > 0 \\
0 & k = 0 \\
\sum_{t=\tau_j^o-1}^{\tau_j^o+k} (l_t(\theta_2^o) - l_t(\theta_1^o)) & k < 0 
\end{cases} \]

is the double-sided random walk.

- Unknown closed form expression for the c.d.f. of \( W_k \).
Theorem 5

**Approximation (Ling 2013):**

\[
(d\hat{\Sigma}d)^2(d\hat{\Omega}d)^{-1}(\hat{r}_j^{(3)} - \tau_j^o) \xrightarrow{d} \arg\max_{r \in \mathbb{R}} B(r) - \frac{1}{2}|r|,
\]

where \(d = \hat{\theta}_1 - \hat{\theta}_2\), \(D_t(\theta) = \partial l_t(\hat{\theta})/\partial \theta\), \(\bar{D} = \sum_{t \in E_j} D_t(\hat{\theta})/4h\),

\[
\hat{\Sigma} = \frac{1}{4h} \sum_{t \in E_j} \left. \frac{\partial^2 l_t(\theta)}{\partial \theta \partial \theta'} \right|_{\hat{\theta}_2}, \quad \hat{\Omega} = \frac{1}{4h} \sum_{t \in E_j} (D_t(\hat{\theta}_2) - \bar{D})(D_t(\hat{\theta}_2) - \bar{D})'.
\]

- **Confidence Interval for \(\tau_j^o\):**

\[
\hat{r}_j^{(3)} \pm \left(1 + \left[\hat{\Delta} F_{\alpha/2}\right]\right)
\]

where \(\hat{\Delta} = (d\hat{\Sigma}d)^{-2}(d\hat{\Omega}d)\), \(P\left(\left|\arg\max_{r \in \mathbb{R}} B(r) - \frac{1}{2}|r|\right| < F_{\alpha/2}\right) = 1 - \alpha\).
Content

1 Motivation

2 Likelihood Ratio Scan for Change-Points Detection
   - Change-point Detection by Scan Statistic
   - Consistent Change-point Estimation by Model Selection
   - Construction of Confidence Intervals

3 Implementation Issue and Computational Complexity

4 Simulation Studies

5 Applications

6 Conclusions
Three-step procedure for change-point estimation

Summary

- Step 1: Detection by Likelihood Ratio Scan Statistics.
- Step 2: Model Selection by Information Criterion.
- Step 3: Construction of Confidence Intervals.
Three-step procedure for change-point estimation

Step 1: Detection by Likelihood Ratio Scan Statistics.

Computational Complexity

- \( O(h) \) for computing \( S_h(t) \).
- \( n - 2h \) replicates of \( S_h(t) \), \( t = h + 1, \ldots, n - h \).
- Sorting for Local Change-point estimator: Negligible.

⇒ Total: \( O(nh) \).
Three-step procedure for change-point estimation

Step 2: Model Selection Approach for consistent change-point estimation.

$$(\hat{m}^{(2)}, \hat{L}^{(2)}, \hat{p}^{(2)}) = \arg\min_{\mathcal{L} \subseteq \hat{L}^{(1)}} IC(m, \mathcal{L}, p)$$

Computational Complexity

- Exact Maximization: $O(\hat{m}^{(1)})^2 \times O(n)$ by dynamic programming.
- Short cut:
  - Sort $\{S_h(t)\}_{t \in \hat{L}^{(1)}}$ and consider $\arg\min \{B$ largest $S_h(\cdot)\}$
  - $O(B^2) \times O(n)$

$\Rightarrow$ Total: $O(B^2n)$. 
Three-step procedure for change-point estimation

Step 3: Final Estimates and Confidence Intervals.

\[ \hat{\tau}^{(3)}_j = \arg \max_{\tau \in (\hat{\tau}^{(2)}_j - h, \hat{\tau}^{(2)}_j + h]} \log L(\hat{\tau}^{(2)}_j - 2h + 1 : \tau)L(\tau + 1 : \hat{\tau}^{(2)}_j + 2h) \]

C.I. = \[ \hat{\tau}^{(3)}_j \pm \left( 1 + \left[ \hat{\Delta} F_{\alpha/2} \right] \right) \]

Computational Complexity

- Find \( \hat{\tau}^{(3)}_j \): \( 2h \times O(h) \).
- Construct C.I.: \( O(h) \).

\( \Rightarrow \) Total: \( O(h^2) \).

Overall Computational Complexity:

\[ O(nh) + O(B^2 n) + O(h^2) = O(nh) = O(n \log n) \]
1 Motivation

2 Likelihood Ratio Scan for Change-Points Detection
   - Change-point Detection by Scan Statistic
   - Consistent Change-point Estimation by Model Selection
   - Construction of Confidence Intervals

3 Implementation Issue and Computational Complexity

4 Simulation Studies

5 Applications

6 Conclusions
In the 3-step procedure, there is only one tuning parameter:

\[ h = d \log n \]

Sensitivity analysis for \( d \)
- How does \( d \) affect \( \hat{m}^{(1)} = |\hat{L}^{(1)}| \)?
- How does \( d \) affect the estimation results?

Model we use:

\[
X_t = \begin{cases} 
0.9X_{t-1} + \epsilon_t, & \text{if } 1 \leq t \leq 0.5n \\
1.69X_{t-1} - 0.81X_{t-2} + \epsilon_t, & \text{if } 0.5n + 1 \leq t \leq 0.75n \\
1.32X_{t-1} - 0.81X_{t-2} + \epsilon_t, & \text{if } 0.75n + 1 \leq t \leq n
\end{cases}
\]

where \( \epsilon_t \sim \text{i.i.d. } N(0, 1) \).
Once $h$ is not too large, $m_o$ can be estimated consistently.

Sudden increase in MDL when $h > \min(\tau^o_{j+1} - \tau^o_j)$.

$h \in (50, 100)$ works well in most cases.
Simulation Experiment 1

Piecewise AR model

\[
X_t = \begin{cases} 
0.9X_{t-1} + \epsilon_t, & \text{if } 1 \leq t \leq 0.5n \\
1.69X_{t-1} - 0.81X_{t-2} + \epsilon_t, & \text{if } 0.5n + 1 \leq t \leq 0.75n \\
1.32X_{t-1} - 0.81X_{t-2} + \epsilon_t, & \text{if } 0.75n + 1 \leq t \leq n
\end{cases}
\]

where \( \epsilon_t \overset{i.i.d.}{\sim} N(0, 1) \).
### Simulation Experiment 1

100 replications, $\tau_1=0.5$, $\tau_2=0.75$

<table>
<thead>
<tr>
<th></th>
<th>$\hat{\tau}_1$</th>
<th>$\hat{\tau}_2$</th>
<th>Coverage</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>LRSM2</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N=1024</td>
<td>0.500(0.009)</td>
<td>0.749(0.013)</td>
<td>98</td>
<td>2.24</td>
</tr>
<tr>
<td>N=2048</td>
<td>0.499(0.008)</td>
<td>0.750(0.003)</td>
<td>99</td>
<td>5.17</td>
</tr>
<tr>
<td>N=4096</td>
<td>0.499(0.003)</td>
<td>0.750(0.001)</td>
<td>98</td>
<td>11.23</td>
</tr>
<tr>
<td>N=8192</td>
<td>0.500(0.001)</td>
<td>0.750(0.001)</td>
<td>99</td>
<td>24.69</td>
</tr>
<tr>
<td>N=16384</td>
<td>0.500(0.001)</td>
<td>0.750(0.001)</td>
<td>99</td>
<td>58.72</td>
</tr>
<tr>
<td>N=32768</td>
<td>0.500(0.0002)</td>
<td>0.750(0.0002)</td>
<td>99</td>
<td>129.34</td>
</tr>
<tr>
<td><strong>Auto-PARM</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N=1024</td>
<td>0.496(0.013)</td>
<td>0.750(0.008)</td>
<td>100</td>
<td>172.05</td>
</tr>
<tr>
<td>N=2048</td>
<td>0.498(0.006)</td>
<td>0.751(0.005)</td>
<td>100</td>
<td>202.18</td>
</tr>
<tr>
<td>N=4096</td>
<td>0.498(0.004)</td>
<td>0.750(0.003)</td>
<td>100</td>
<td>371.41</td>
</tr>
<tr>
<td>N=8192</td>
<td>0.500(0.002)</td>
<td>0.750(0.002)</td>
<td>100</td>
<td>776.82</td>
</tr>
<tr>
<td>N=16384</td>
<td>0.499(0.002)</td>
<td>0.750(0.001)</td>
<td>100</td>
<td>1513.58</td>
</tr>
<tr>
<td>N=32768</td>
<td>0.500(0.001)</td>
<td>0.750(0.001)</td>
<td>98</td>
<td>2996.56</td>
</tr>
<tr>
<td><strong>PELT (Rep = 5)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N=1024</td>
<td>0.501(0.006)</td>
<td>0.750(0.005)</td>
<td>100</td>
<td>1141.92</td>
</tr>
<tr>
<td>N=2048</td>
<td>0.498(0.003)</td>
<td>0.751(0.005)</td>
<td>100</td>
<td>2911.11</td>
</tr>
</tbody>
</table>
Simulation Experiment 2

Piecewise AR/MA

\[
X_t = \begin{cases} 
-0.9X_{t-1} + \epsilon_t + 0.7\epsilon_{t-1}, & \text{if } 1 \leq t \leq 0.5n \\
0.9X_{t-1} + \epsilon_t, & \text{if } 0.5N + 1 \leq t \leq 0.75n \\
\epsilon_t - 0.7\epsilon_{t-1}, & \text{if } 0.75n + 1 \leq t \leq n
\end{cases}
\]

where \( \epsilon_t \overset{i.i.d.}{\sim} N(0, 1) \).
### Simulation Experiment 2

100 replications, $\tau_1=0.5$, $\tau_2=0.75$

<table>
<thead>
<tr>
<th>LRSS</th>
<th>$\hat{\tau}_1$</th>
<th>$\hat{\tau}_2$</th>
<th>Coverage</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>N=1024</td>
<td>0.498(0.006)</td>
<td>0.751(0.002)</td>
<td>100</td>
<td>2.21</td>
</tr>
<tr>
<td>N=2048</td>
<td>0.499(0.003)</td>
<td>0.751(0.001)</td>
<td>100</td>
<td>4.90</td>
</tr>
<tr>
<td>N=4096</td>
<td>0.499(0.001)</td>
<td>0.750(0.001)</td>
<td>100</td>
<td>11.03</td>
</tr>
<tr>
<td>N=8192</td>
<td>0.500(0.001)</td>
<td>0.750(0.0002)</td>
<td>100</td>
<td>23.78</td>
</tr>
<tr>
<td>N=16384</td>
<td>0.500(0.0004)</td>
<td>0.750(0.0001)</td>
<td>100</td>
<td>54.99</td>
</tr>
<tr>
<td>N=32768</td>
<td>0.500(0.0001)</td>
<td>0.750(0.0001)</td>
<td>100</td>
<td>117.45</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Auto-PARM</th>
<th>$\hat{\tau}_1$</th>
<th>$\hat{\tau}_2$</th>
<th>Coverage</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>N=1024</td>
<td>0.496(0.008)</td>
<td>0.752(0.004)</td>
<td>100</td>
<td>162.85</td>
</tr>
<tr>
<td>N=2048</td>
<td>0.499(0.005)</td>
<td>0.751(0.002)</td>
<td>100</td>
<td>305.88</td>
</tr>
<tr>
<td>N=4096</td>
<td>0.499(0.003)</td>
<td>0.751(0.001)</td>
<td>100</td>
<td>509.16</td>
</tr>
<tr>
<td>N=8192</td>
<td>0.499(0.001)</td>
<td>0.750(0.001)</td>
<td>99</td>
<td>1204.02</td>
</tr>
<tr>
<td>N=16384</td>
<td>0.500(0.001)</td>
<td>0.750(0.0004)</td>
<td>98</td>
<td>2105.29</td>
</tr>
<tr>
<td>N=32768</td>
<td>0.500(0.0005)</td>
<td>0.750(0.0003)</td>
<td>84</td>
<td>4924.47</td>
</tr>
</tbody>
</table>
Simulation Experiment 3

Piecewise AR

\[
X_t = \begin{cases} 
0.9X_{t-1} + \epsilon_t, & \text{if } 1 \leq t \leq 0.5n \\
1.69X_{t-1} - 0.81X_{t-2} + \epsilon_t, & \text{if } 0.5n + 1 \leq t \leq 0.75n \\
1.32X_{t-1} - 0.81X_{t-2} + \epsilon_t, & \text{if } 0.75n + 1 \leq t \leq n 
\end{cases}
\]

where \( \epsilon_t \sim i.i.d. N(0, 1) \).
Coverage accuracy of confidence intervals:

<table>
<thead>
<tr>
<th>$\tau_j^o$</th>
<th>Mean of $\hat{\tau}^{(3)}$</th>
<th>Mean of 90% C.I.</th>
<th>Coverage Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>512</td>
<td>512</td>
<td>[488, 536]</td>
<td>91.7%</td>
</tr>
<tr>
<td>768</td>
<td>768</td>
<td>[760, 775]</td>
<td>90.3%</td>
</tr>
</tbody>
</table>

n=1024, 1000 replications
Content

1 Motivation

2 Likelihood Ratio Scan for Change-Points Detection
   • Change-point Detection by Scan Statistic
   • Consistent Change-point Estimation by Model Selection
   • Construction of Confidence Intervals

3 Implementation Issue and Computational Complexity

4 Simulation Studies

5 Applications

6 Conclusions
Electroencephalogram (EEG) Time Series

- Recorded from a patient diagnosed with left temporal lobe epilepsy.
- Data collection:
  - Sampling rate: 100Hz,
  - Recording period: 5 minutes and 28 seconds,
  - Sample size: $n=32,768$.
Electroencephalogram (EEG) Time Series

Results of EEG data analysis: $n = 32768$

<table>
<thead>
<tr>
<th></th>
<th>Three-step</th>
<th>Auto-PARM</th>
<th>PELT</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of $\hat{\tau}$</td>
<td>14</td>
<td>12</td>
<td>33</td>
</tr>
<tr>
<td>MDL</td>
<td>114195</td>
<td>114335</td>
<td>114074</td>
</tr>
<tr>
<td>Time</td>
<td>441s</td>
<td>5726s</td>
<td>$&gt; 1$ week (186.9 hours)</td>
</tr>
</tbody>
</table>
Electroencephalogram (EEG) Time Series

- PELT (Killick, Fearnhead, Eckley (2012))

- AUTO-PARM (Davis, Lee and Rodriguez-Yam (2005))

- Three-step procedure
IBM return data

- From Jan 1926 to Dec 2008.
- 996 data points.

- Time for computation: 4.2s
IBM return

Results:

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{\tau}$</td>
<td>41</td>
<td>174</td>
</tr>
<tr>
<td>C.I.</td>
<td>[19, 63]</td>
<td>[168, 180]</td>
</tr>
<tr>
<td>Length</td>
<td>44</td>
<td>12</td>
</tr>
</tbody>
</table>

Interpretations:

- July 1927 - Oct 1931: Great Depression.
- Dec 1939 - Dec 1940: World War II.
1 Motivation

2 Likelihood Ratio Scan for Change-Points Detection
   - Change-point Detection by Scan Statistic
   - Consistent Change-point Estimation by Model Selection
   - Construction of Confidence Intervals

3 Implementation Issue and Computational Complexity

4 Simulation Studies

5 Applications

6 Conclusions
Conclusions

- Fast three-step procedure for change-points inference.
  
  **Step 1:** Detection using Scan statistics
  **Step 2:** Estimation by model selection information criterion.
  **Step 3:** Extended window for CI

- Advantages:
  - Fast
  - Few and non-sensitive tuning parameters
  - Especially suitable for large $n$ small $m$ case.
Thank You!