Change point detection

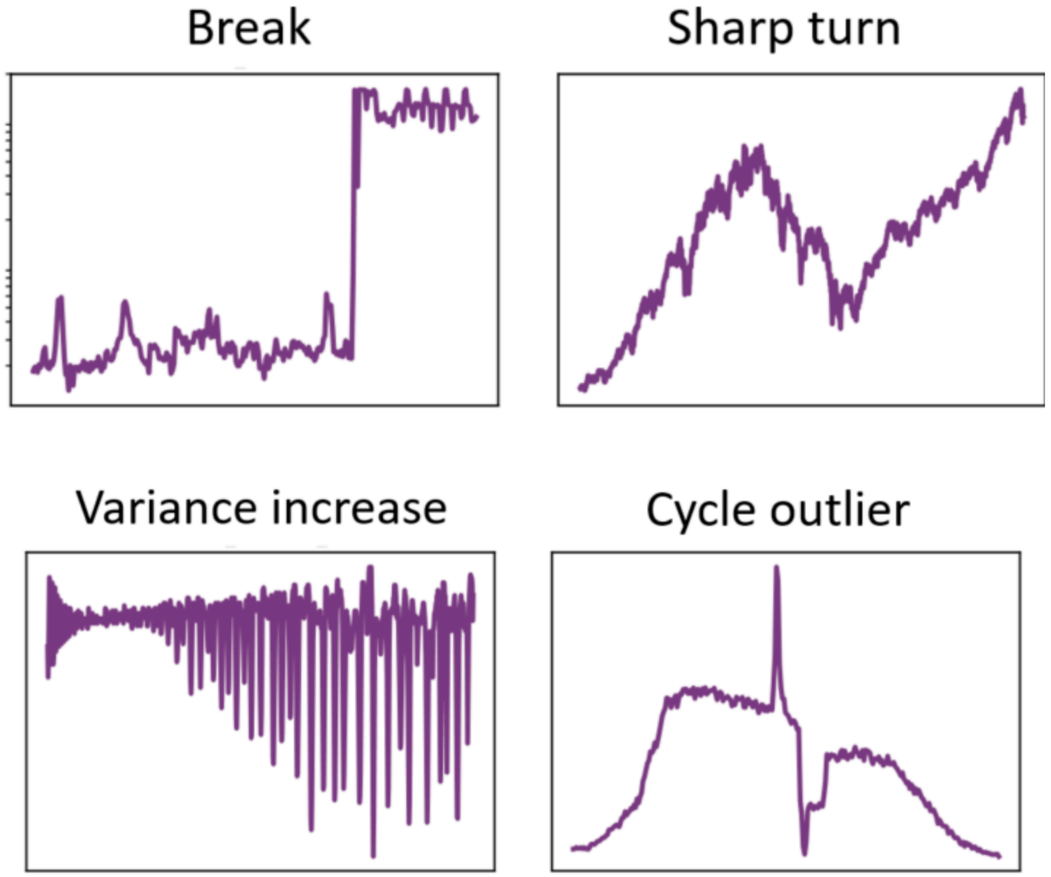
Skoltech

What is a change point?

Skoltech

What is a change point?

- ***Change point***: change of statistical properties of a series



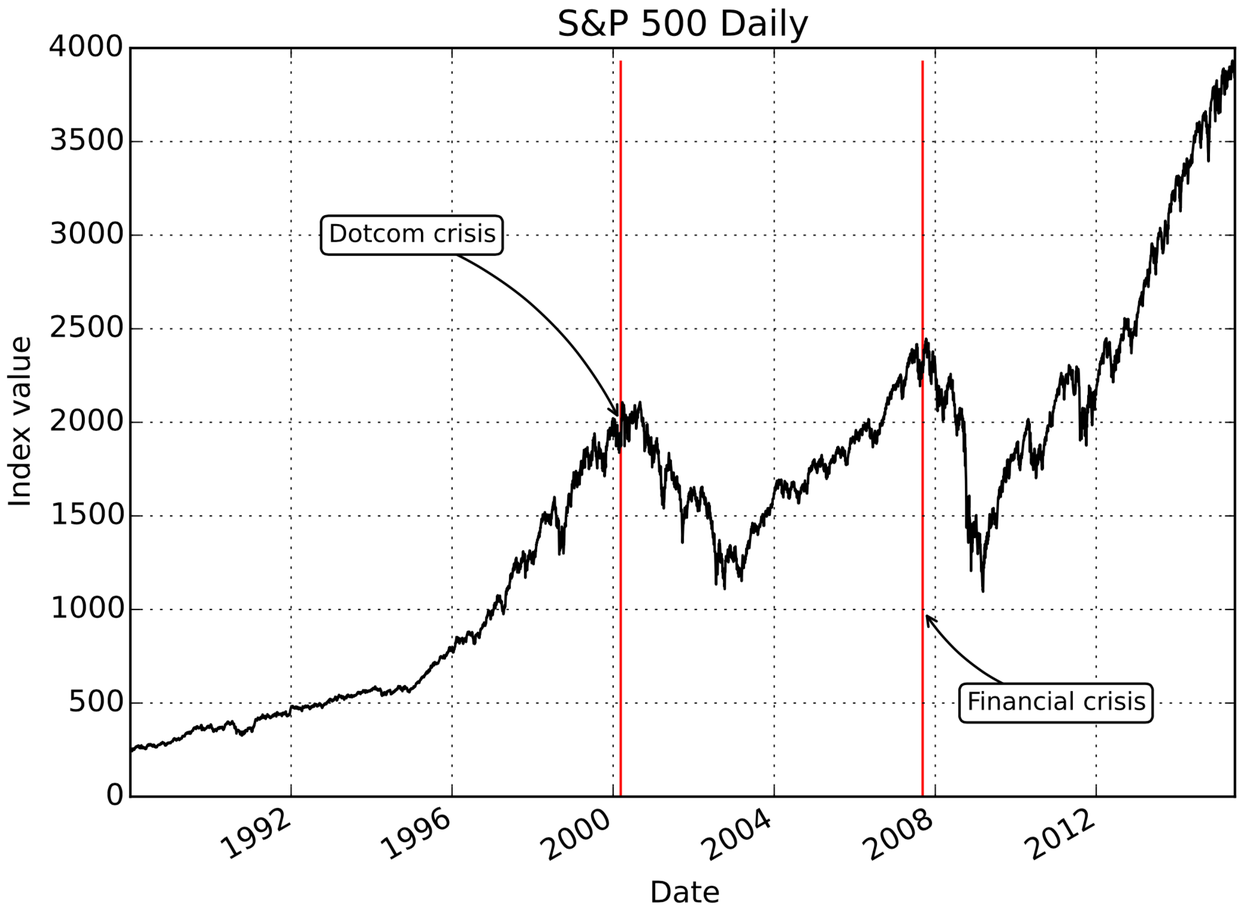
«Change point» problem

identify the change moment

Skoltech

Real world example: financial data

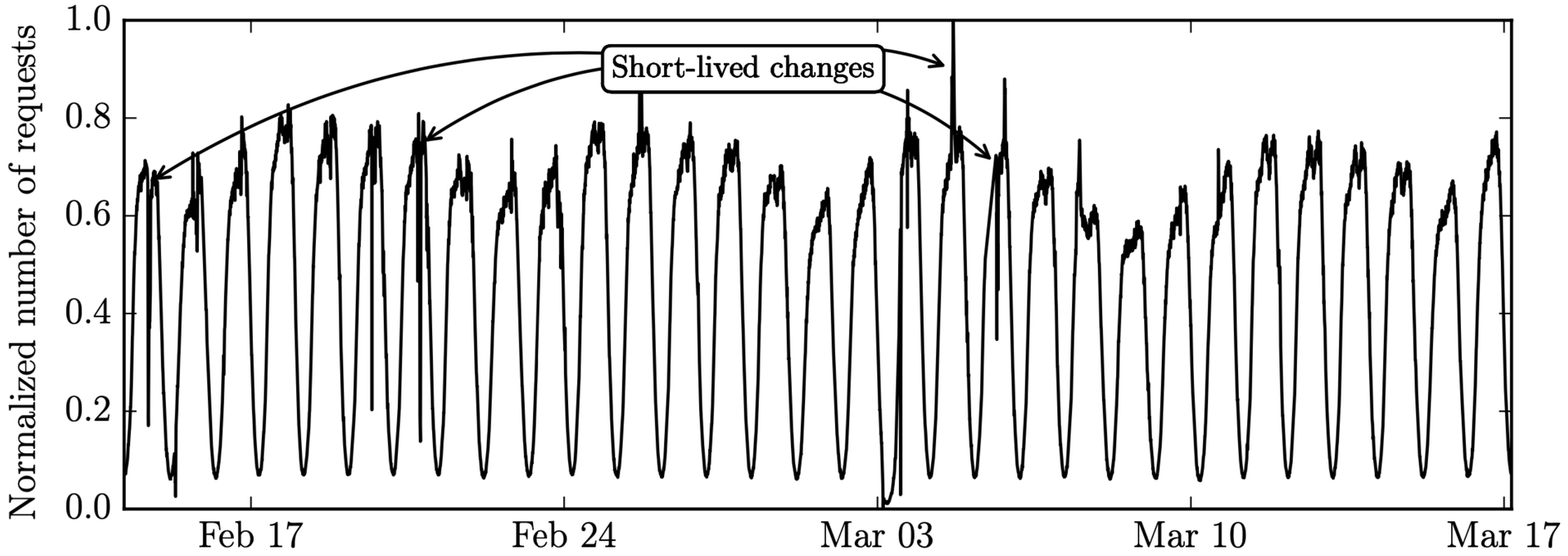
- S&P 500 index during 27 years



Skoltech

Real world example: internet data

- Internet service audience



Skoltech

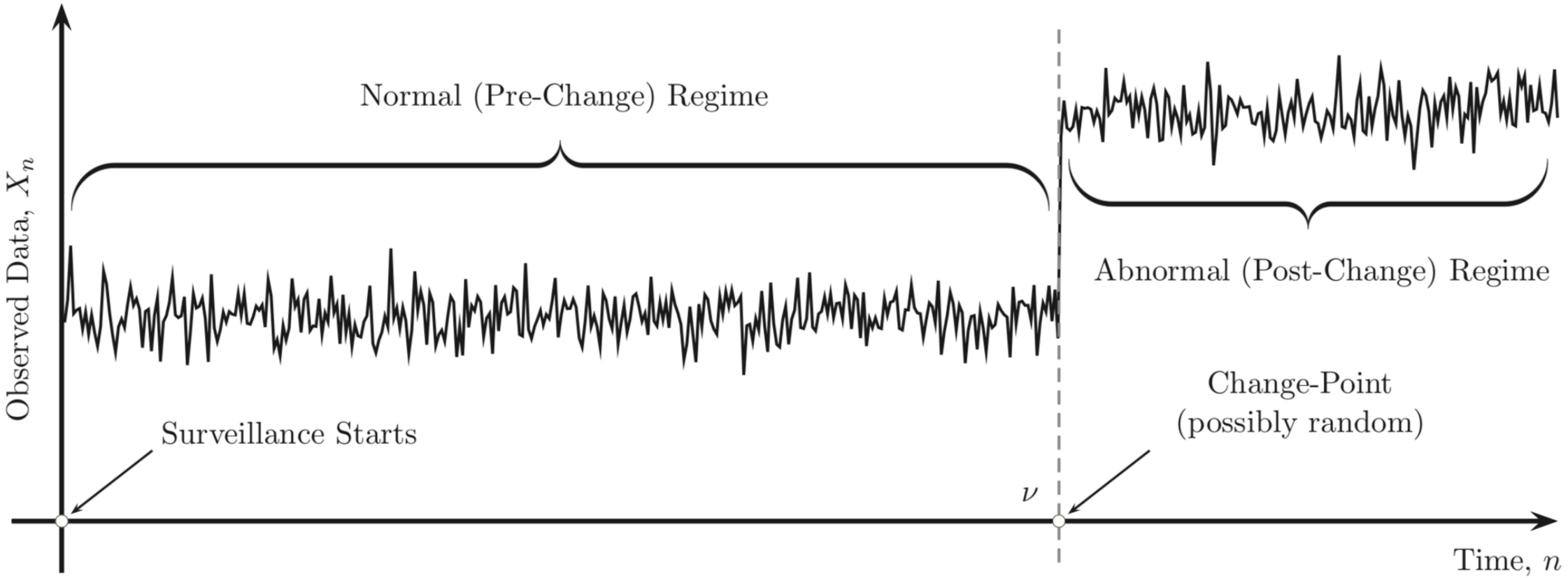
Is it important? Examples of applications

- Intrusion into networks detection (attack on change of transferred traffic)
- Detection of anomalies in data transmission networks (video streams in video surveillance systems, network traffic, etc.)
- Detection and isolation of system node failures for vehicle management
- Disease outbreak detection
- Automatic detection of abnormal human behaviour via video surveillance
- Monitoring and analysis of mortality and incidence of lung cancer

Change point detection (CPD)

Skoltech

Typical scenario for observations with a change point



Source: Polunchenko, Aleksey S., and Alexander G. Tartakovsky. "State-of-the-art in sequential change-point detection."Methodology and computing in applied probability 14.3 (2012): 649-684.

Skoltech

Change point detection: alarm

- We have observations up to time n

$$\mathbf{X}_n = (X_1, \ldots, X_n)$$

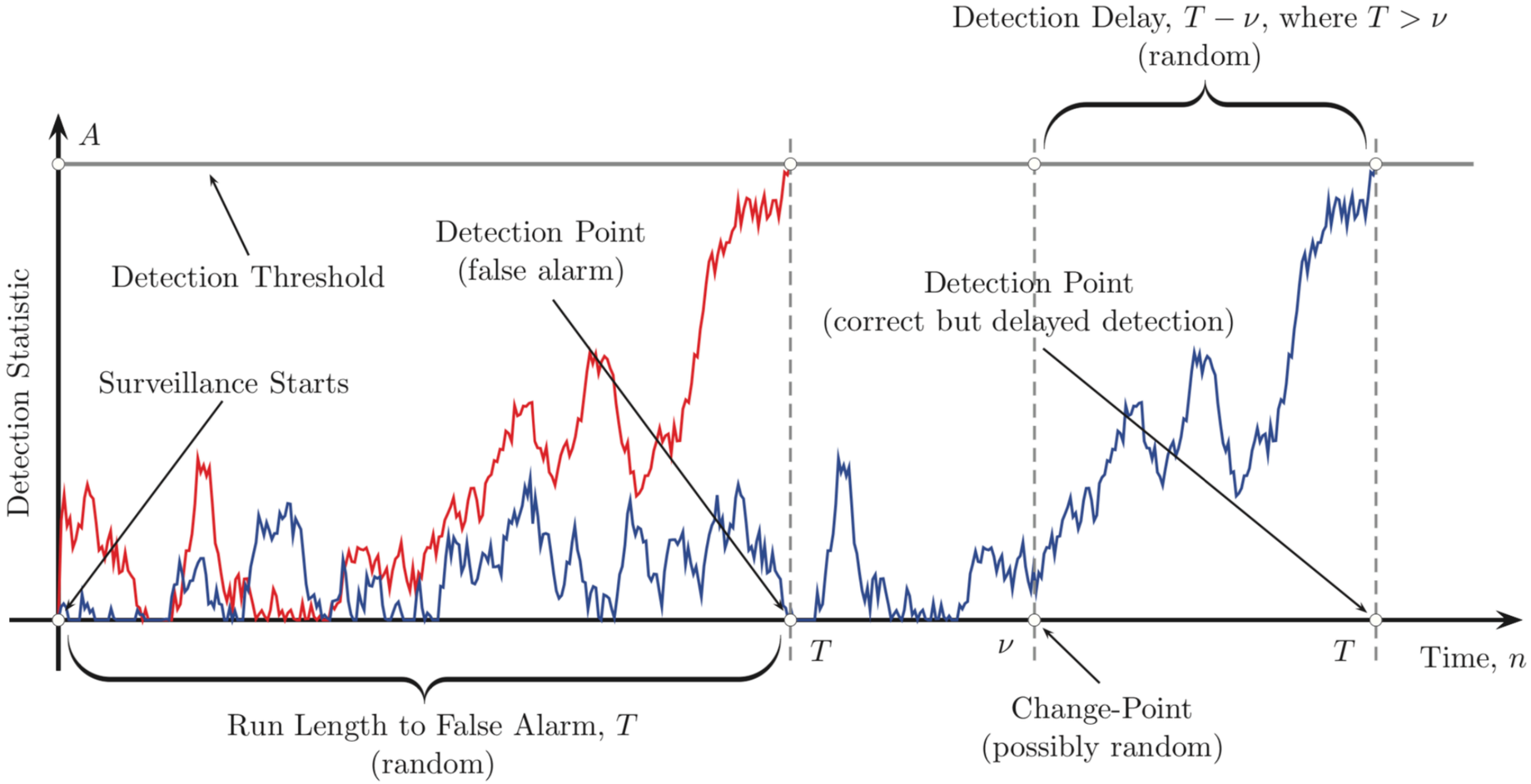
- We want to give alarm at $\tau = n$, if there is a change point
- **Stopping time**: statistic $\tau = \tau(\mathbf{X}_n)$

$$\tau \in \{0, 1 \ldots, \infty\}, \qquad \{\mathbf{X}_n : \tau(\mathbf{X}_n) = n\} \in \sigma(\mathbf{X}_n)$$

- Note, that we can't use information from future
- Stopping time τ is based on the statistic of observations

$$\gamma = (\gamma_t)_{t \geqslant 0}, \gamma_n = \gamma_n(\mathbf{X}_n)$$

Typical scenario for observations with a change point



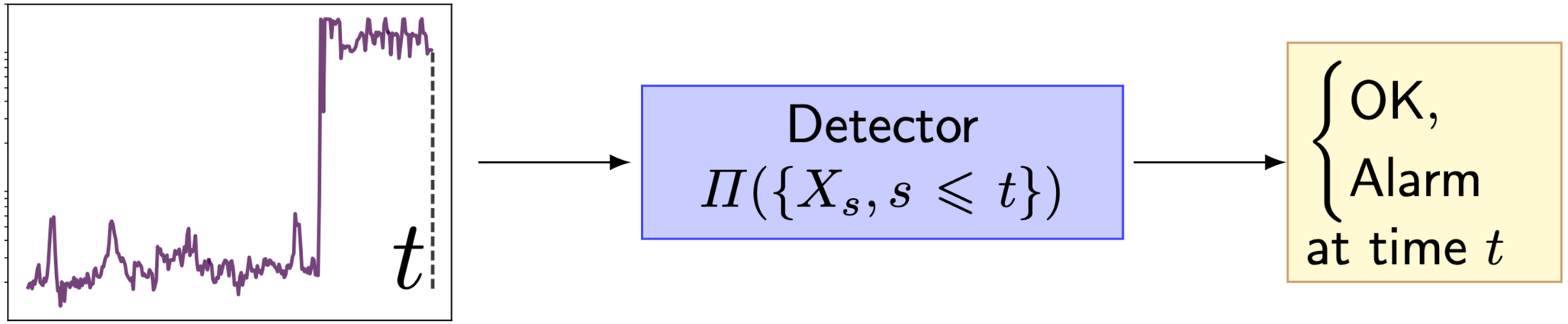
Source: Polunchenko, Aleksey S., and Alexander G. Tartakovsky. "State-of-the-art in sequential change-point detection."Methodology and computing in applied probability 14.3 (2012): 649-684.

Skoltech

High quality of the stopping time $\tau(X_n)$

- The stopping time: statistic $\tau = \tau(\mathbf{X}_n)$
- $\mathrm{E}_\infty \tau$: (mean) false detection delay, $\mathrm{FDD}(\tau)$
- **Good:** $\mathrm{E}_\infty \tau \to \infty$ (rare false alarms)
- $\mathrm{E}_0 \tau$ or $\mathrm{E}_\theta[\tau - \theta | \tau > \theta]$: average detection delay, $\mathrm{ADD}(\tau)$
- **Good:** $\mathrm{E}_0 \tau \to 0$ (fast detection)

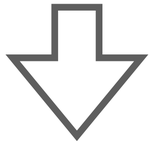
General Procedure for CPD



- Π : detector, procedure for change point detection (takes into account assumptions about the signal and about the change point)
- Π : alarm time $\tau = \inf\{t \geqslant 0 : \gamma_t \geqslant h\}$

Pruned Exact Linear Time (PELT)

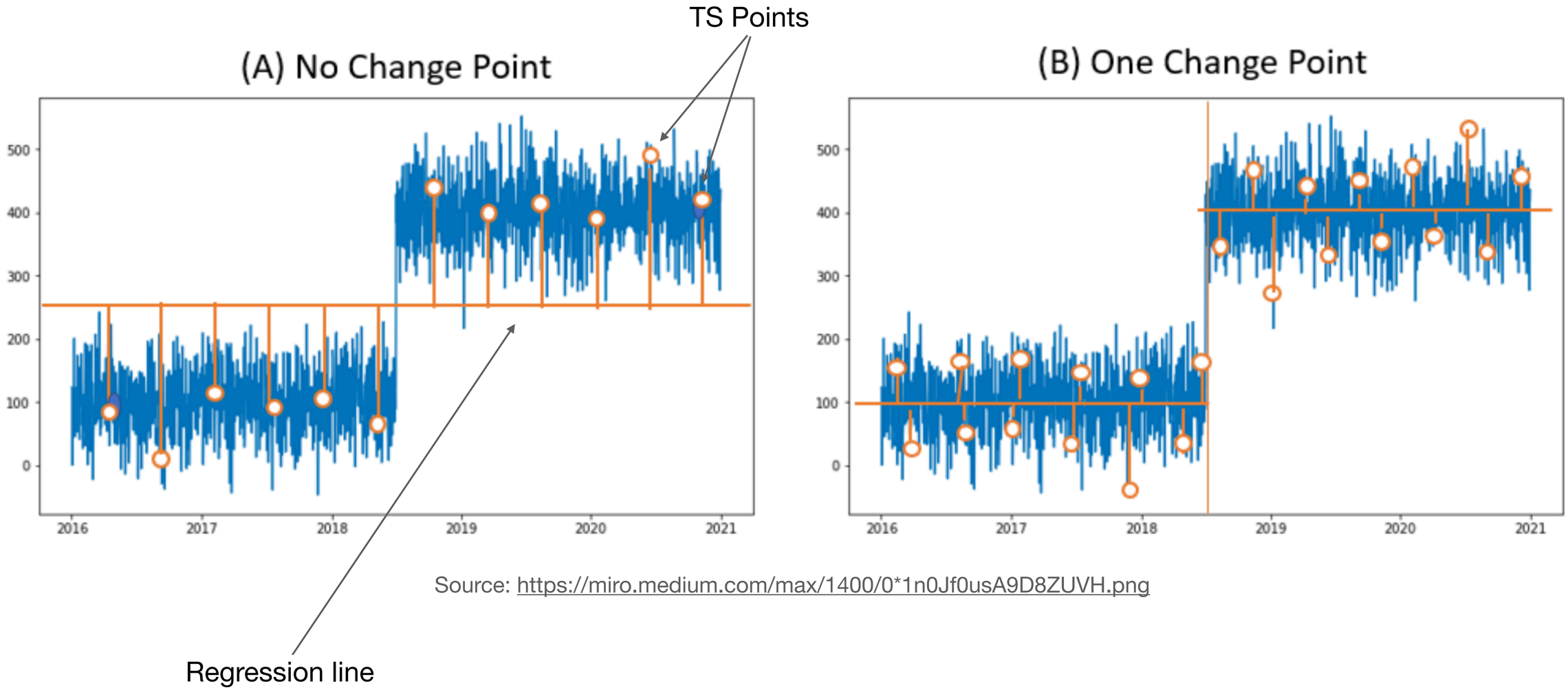
- We try to find the underlying pattern such as a regression line to forecast the future
- Regression line is not straight if there are change points



build segmented regression lines where the kinked points are the change points

Source: Gachomo Dorcas Wambui, Gichuhi Anthony Waititu, Jomo Kenyatta (2015). The Power of the Pruned Exact Linear Time(PELT) Test in Multiple Changepoint Detection American Journal of Theoretical and Applied Statistics.Volume 4, Issue 6, November 2015, Pages: 581–586 (arxiv_link)

PELT illustration

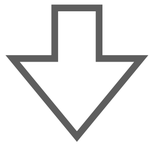


Source: https://miro.medium.com/max/1400/0*1n0Jf0usA9D8ZUVH.png

Skoltech

PELT algorithm

- Slide the cut point from left to right of the TS



- Find the appropriate change point for the time series that minimizes the sum of the distances or errors:

$$min \sum_{i=1}^{m+1} [C(y_{\tau_{i-1}+1}, \ldots, y_{\tau_i}) + \beta],$$

where $C(\,.\,)$ is the distance or the *cost function,*

β - number of segments as a penalty term to prevent the search from yielding too many segments

ruptures package

Pros:

- Many methods available
- Many other configurable options
- Python

Cons:

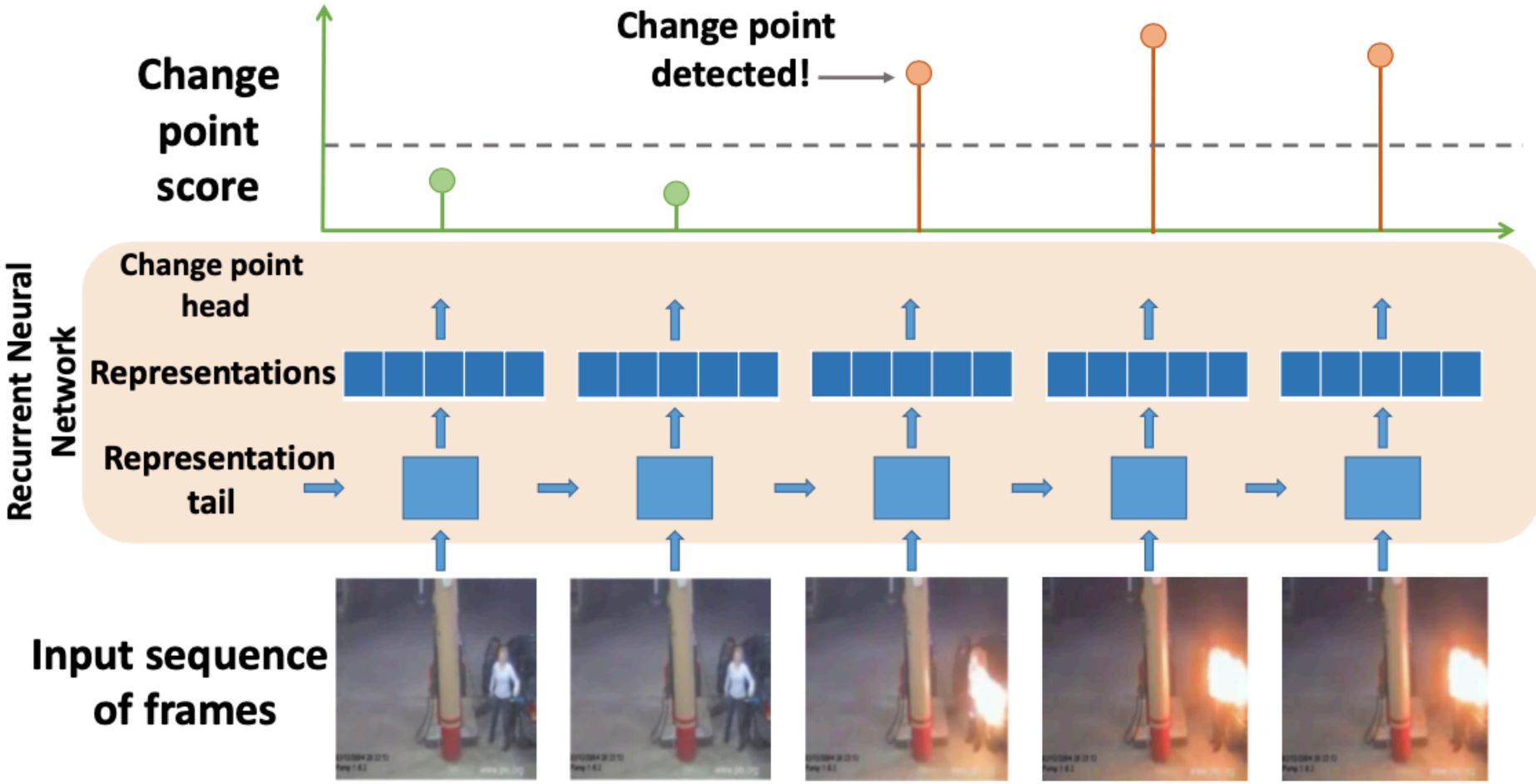
- Offline
- Requires careful selection of configuration

https://centre-borelli.github.io/ruptures-docs/

Change detection for semi-structured data

- CPD model signals about time of change in the data distribution
- Semi-structured data – sequences of semi-structured data (images, texts)
- *Goal*: minimize Detection Delay & minimize number of False Alarms
- *Problem*: Can’t apply classic method for semi-structured data

Change detection for semi-structured data



Skoltech

Change detection for semi-structured data

Method	Mean Time to FA ↑	Mean DD ↓	AUC ↓	F1 ↑	Covering ↑
Explosions					
BCE simple	11.20	**0.44**	8.388	0.3023	0.8484
CPD seq2seq	14.23	1.54	9.466	0.4048	0.8798
BCE seq2seq	**14.95**	1.81	8.546	0.4667	0.8836
InDiD (ours)	12.76	0.64	**7.470**	**0.5472**	**0.8955**
Road Accidents					
BCE simple	9.69	**0.48**	**12.486**	0.0417	0.7817
CPD seq2seq	14.64	2.18	13.257	0.1176	0.8299
BCE seq2seq	**15.23**	2.32	12.770	**0.1860**	**0.8440**
InDiD (ours)	15.20	2.31	12.896	0.1647	0.8418

Skoltech

Conclusion

- CPD play crucial role in various applications
- CPD algorithms have Python implementation
- We can do change point detection for semi-structured data

Next lecture: TS in discrete events

Skoltech