Score-Based Change Point Detection and Region Localization for Spatio-Temporal Point Processes

Wenbin Zhou[1], Liyan Xie[2], Shixiang (Woody) Zhu[1]

1 Carnegie Mellon University
2 University of Minnesota

Link to paper

I. Motivating Application

Earthquakes are killing lives and damaging properties

Q: When changes happen? Where changes happen?

II. Preliminaries

(1) Spatio-Temporal Point Process

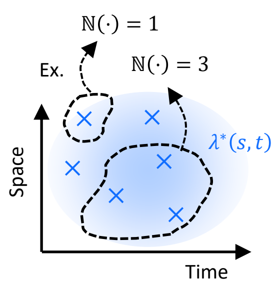
Counting measure:

$\mathbb{N}(s,t)$ = number of events at location s and time t

Conditional intensity:

$$\lambda^*(s,t) = \frac{d\mathbb{N}(s,t)}{ds\,dt}$$

*Random variables



(2) Change Point Detection

Define τ to be the (unknown) change point

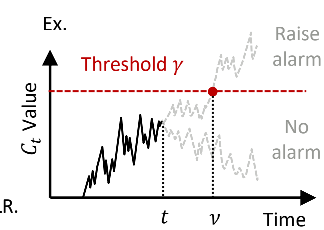
$$H_0: x_1, \ldots, x_n \sim P_0 \quad \text{versus} \quad H_1: \begin{array}{l} x_1, \ldots, x_\tau \sim P_0 \\ x_\tau, \ldots, x_n \sim P_1 \end{array}$$

Common tool: cumulative sum control chart (CUSUM) statistics

$$C_t = \max_{\tau \le t} \sum_{\tau \le i \le t} \Delta_i$$

Cumulative sum of anomalies across time

Traditionally, use $\Delta_i \approx$ LLR.



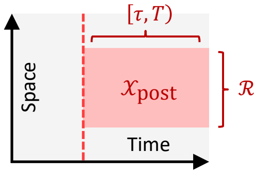
III. Problem Setting

Let $\mathcal{X}$ denote the joint spatial-temporal domain.

$$H_0: \lambda^*(x) = \lambda_0^*(x), \quad \forall x \in \mathcal{X}$$

$$H_1: \lambda^*(x) = \lambda_0^*(x), \quad \forall x \in \mathcal{X}_{\text{pre}}$$

$$\lambda^*(x) = \lambda_1^*(x), \quad \forall x \in \mathcal{X}_{\text{post}}$$



Goal: estimate change point τ and change region $\mathcal{R}$ and raise an alarm as quickly as possible after τ.

IV. Proposed Algorithm

(1) Score-Based Spatio-Temporal Statistics

$$C_t = \max_{\tau \le t} \max_{\Omega \subseteq 2^\delta} \int_{[\tau,t)\times\Omega} \Delta(s,t')d\mathbb{N}(s,t')$$

We configure $\Delta(s,t)$ as the **difference of Hyvarinen scores**

Pros: (*i*) avoids integral calculations; (*ii*) scales to high-D settings.

(2) Alternating Optimization

Idea: Decompose into subproblems + leverage discreteness of $\mathbb{N}$

For each time step t, repeat for K times:

1. Solve for inner problem (**I-step**):

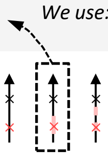
$$\max_{\tau \le t} \max_{\Omega \subseteq 2^\delta} \int_{[\tau,t)\times\Omega} \Delta(s,t')d\mathbb{N}(s,t')$$

We use:

Ex. Suppose we observe two events

× : Δ < 0
× : Δ > 0

These three Ω are equally optimal:



2. Solve for outer problem (**O-step**):

$$\max_{\tau \le t} \max_{\Omega \subseteq 2^\delta} \int_{[\tau,t)\times\Omega} \Delta(s,t')d\mathbb{N}(s,t')$$

Can be simplified to a finite optimization problem:

$$\max_{\tau \in \{t\} \cup \{t_j\}_{j=1}^{N(t)}} \sum_{x_i : [\tau,t)\times\Omega} \Delta(x_i)$$

V. Some Theoretical Results

$$\text{Expected detection delay} \le \mathcal{O}\left(\frac{1}{\overline{\lambda_{\text{post}}} \cdot D_F(f_{\text{pre}} \| f_{\text{post}})}\right)$$

$\overline{\lambda_{\text{post}}}$: average intensity $\quad D_F(f_{\text{pre}} \| f_{\text{post}})$: Fisher divergence

If assuming the underlying process is a self-exciting point process:

$$D_F(f_{\text{pre}} \| f_{\text{post}}) = \mathcal{O}\left((\mu_{\text{pre}} - \mu_{\text{post}})^2(1-\alpha)\right)$$

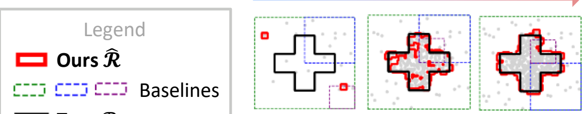
μ : base rate of pre/post-change process $\quad \alpha$: branching ratio

Takeaway: more "localized" differences → quicker detection

VI. Some Experiment Results

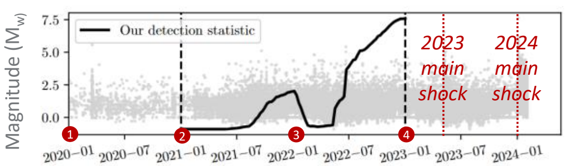
(1) Synthetic Data

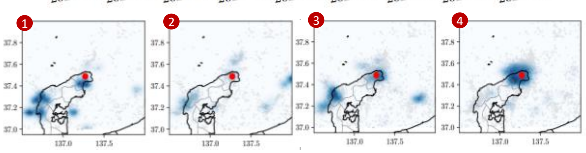
After a change occurs, spatially:



Also achieves a low detection delay

(2) Real Data: Noto Earthquake Catalog





● Seismic records *(Smoothed) detected change region*

Time series ver. **Zhou, W.**, Xie, L., Peng, Z., & Zhu, S. (2025). Sequential Change Point Detection via Denoising Score Matching. *arXiv preprint arXiv:2501.12667*.