Denoising Score Matching for Online Change Point Detection: Spatial-Temporal Extension and Case Study on Noto Earthquakes

Wenbin Zhou[1](wenbinz2@andrew.cmu.edu), Xu Si[2], Liyan Xie[3], Zhigang Peng[2], Shixiang Zhu[1]

[1]*Carnegie Mellon University,* [2]*Georgia Institute of Technology,* [3]*University of Minnesota*

Abstract

1. Proposes a machine-learning framework for automated precursor detection in high dimensional setting
2. The framework provides not only the estimated precursor time but also the corresponding geographic region in a principled manner.
3. Studies precursors the 2018 - 2024 Noto JMA earthquake catalog using the developed framework

Introduction: Precursors

What are precursors: Measurable physical phenomena that occur before an earthquake and may signal that an earthquake is about to happen

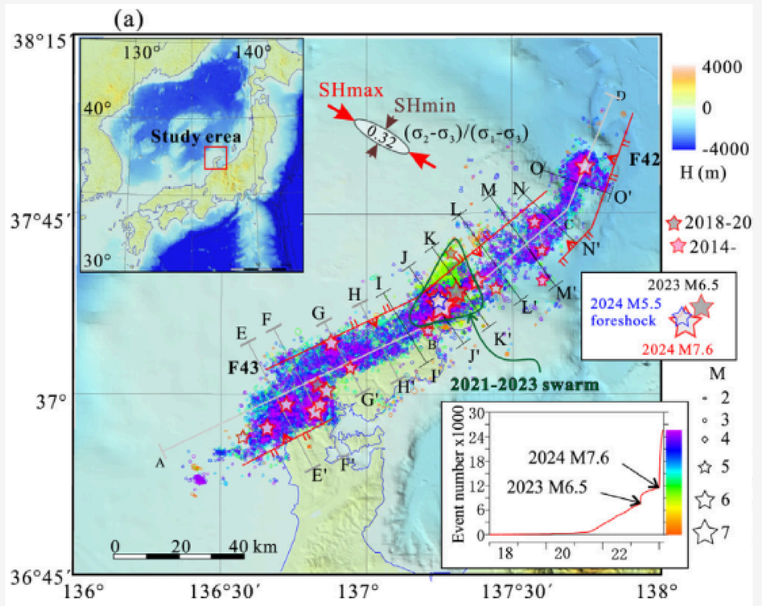


Figure 1: 2024 M7.5 Noto earthquake swarm region identified through inspecting relocated seismicity, inferred fluid migration [Peng et.al. 2025]

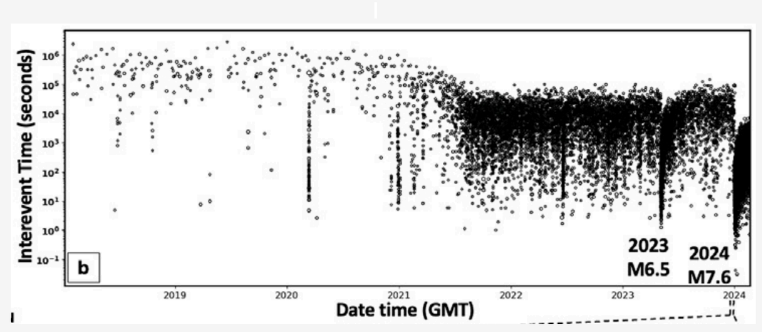


Figure 2: The inter-event recurrence time following the time of the relocated catalog in the whole region from January 2018 to February 2024

Research Questions

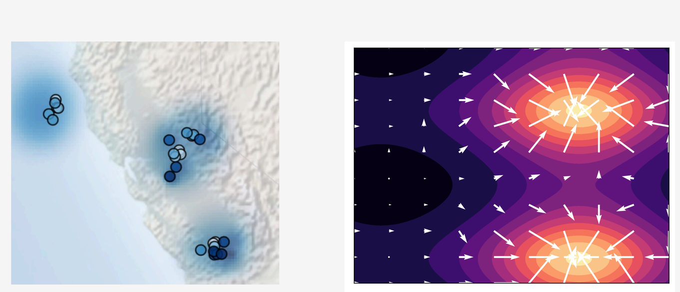
1. Can precursor detection be automated for high-dimensional setting?
2. Can precursor detection identify both alarm time [2] and region-at-risk?
3. Other concerns: (i) Computation efficiency, (ii) Detection effectiveness

Methodology Overview

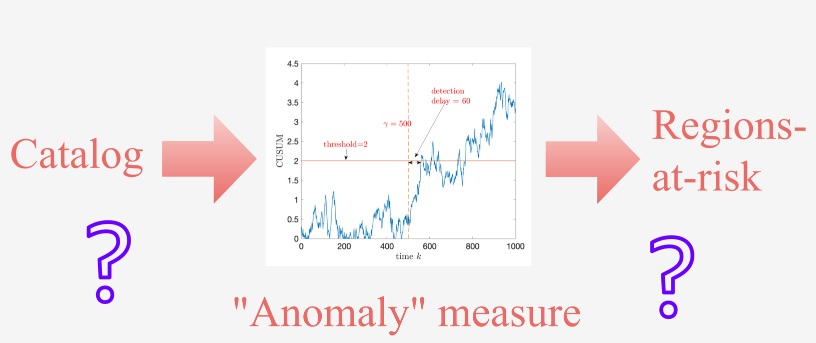
Key idea: Detect distribution discrepancy for catalog events across different spatio-temporal domains

Novelties

(a) Instead of estimating intensity, we estimate the Stein score function (gradient of log density). (Zhou et. al. 2025)



(b) Extension from traditional change-point detection to the spatio-temporal domain



Proposed Algorithm

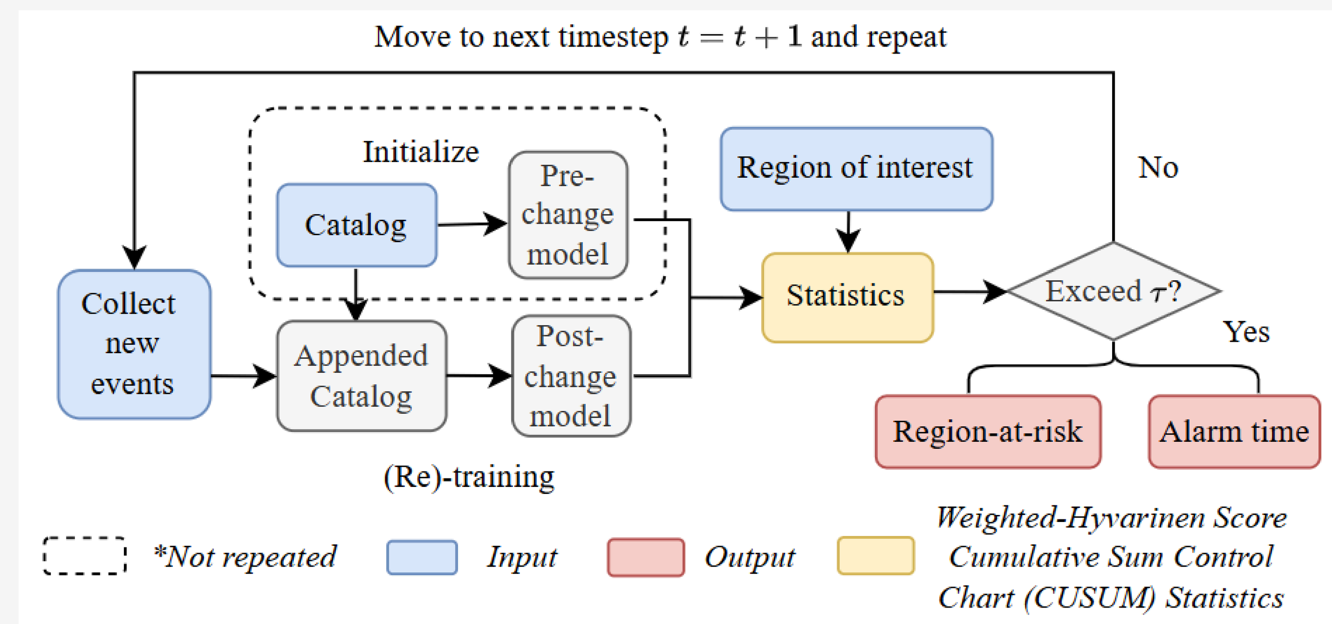


Figure 3: The algorithmic flowchart of the proposed framework

Key features:

1. Unsupervised learning on events
2. Recurrent algorithmic structure
3. End-to-end pipeline

Benefits:

1. Abundance of data in practice
2. Real-time detection and fast response
3. Fully automated detection with less human intervention

Numerical Experiments and Results

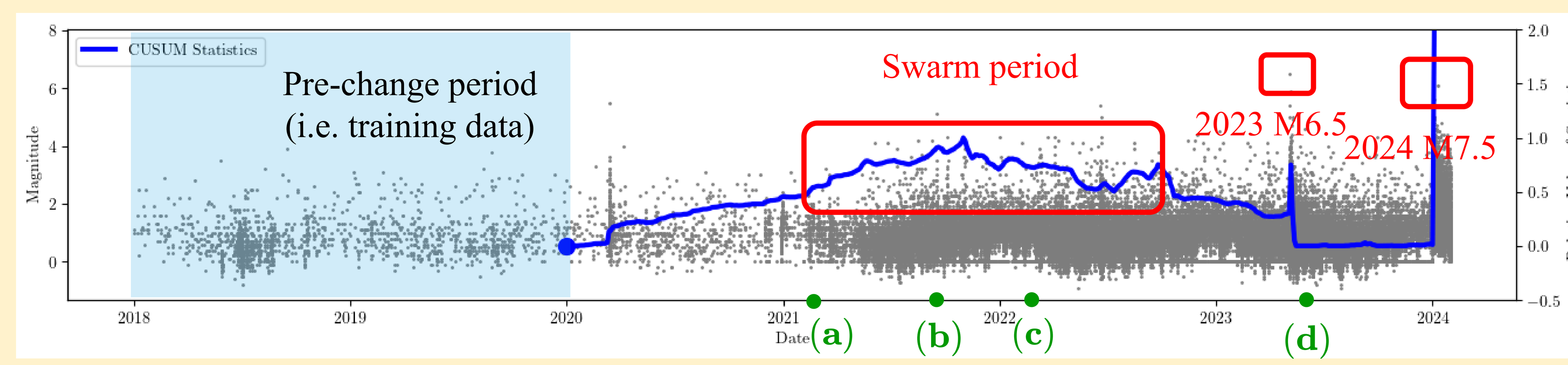


Figure 4: Magnitude versus date of Noto JMA Catalog (2018 - 2024). The blue curve is the estimated (whole region) CUSUM statistics by date produced by our algorithm

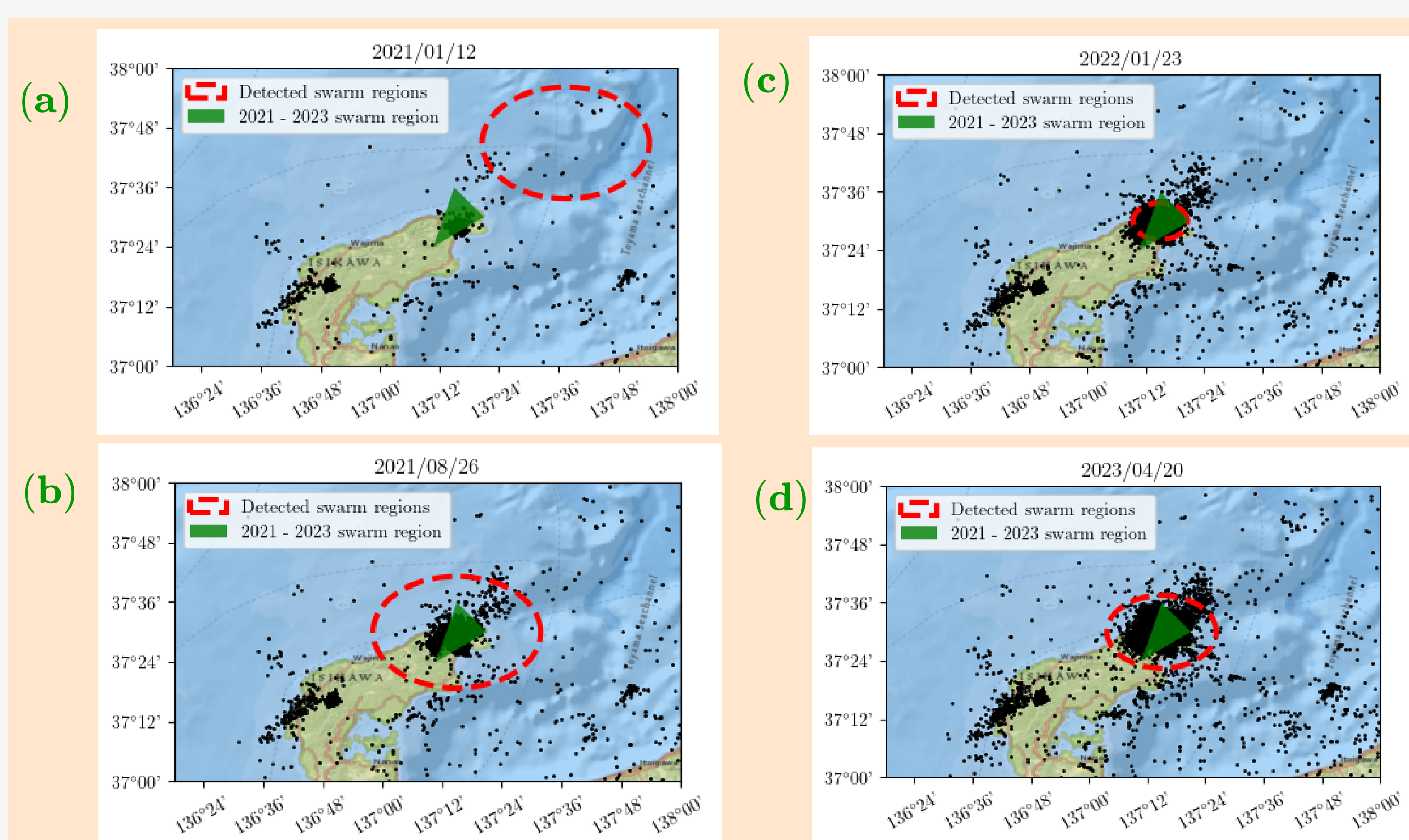


Figure 5: Geographical view of the spatial-temporal detected region-at-risk (red dashed circles) and the swarm region identified in Peng et. al. 2025.

Subplots (a) (b) (c) and (d) correspond to timestamps marked in Figure 4

The earliest correct detection time for the swarm region is around August 2021.

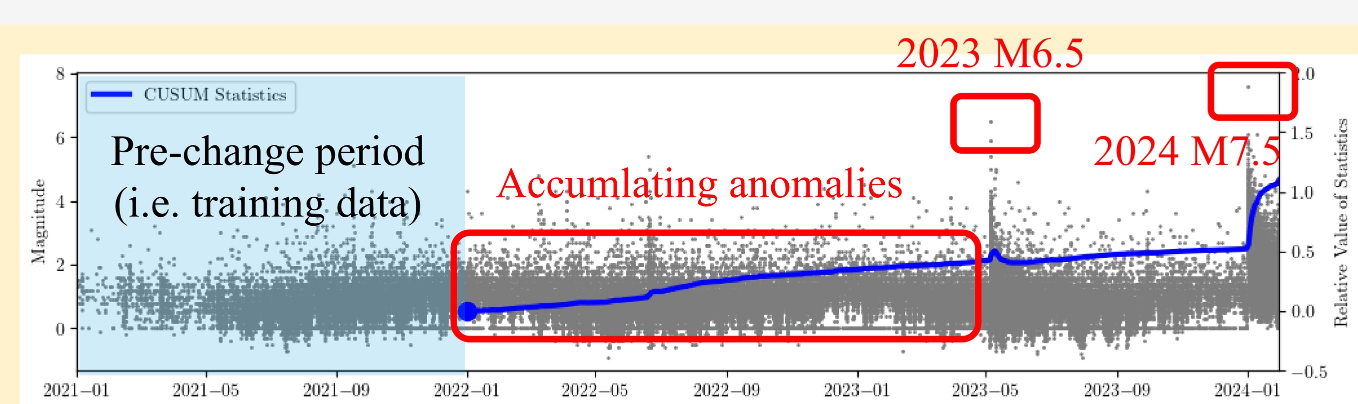


Figure 6: Similar plot to Figure 4, but the pre-change model is trained with fine-grained catalog data within 2021-2022.

Conclusion

1. Proposed a principled spatio-temporal change point detection framework, facilitating earthquake early warning
2. Captures anomalous events in the 2021 - 2023 swarm period
3. Pin-points the exact location of the swarm region

Future directions: more complete evaluation of the method to make it trustworthy.

References

[1] Peng, Z., Lei, X., Wang, Q. Y., Wang, D., Mach, P., Yao, D., ... & Campillo, M. (2025). The evolution process between the earthquake swarm beneath the noto peninsula, central japan and the 2024 m 7.6 noto hanto earthquake sequence. *Earthquake Research Advances, 5(1), 100332.*

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

Appendix: Mathematical Details

Addressing Challenge 3(i): Approximation via truncated likelihood

$$\overline{\ell}(\mathcal{H}_t, \nu, t) \approx \int_{\underbrace{[\nu,t)\times\mathcal{M}}_{\text{Mark space}}} \log p_i(x|\mathcal{H}_{t(x)}) \underbrace{\mathrm{d}\mathbb{N}(x)}_{\text{Counting measure}}$$

Approximation error:

$$\ell - \overline{\ell} = \int_{[t_n, t(x))} \lambda_i(u|\mathcal{H}_{t(u)})\mathrm{d}u$$

Drastically reduce computation load without sacrificing too much modeling accraucy

Solution to Challenge 3(ii): Use Stein score function instead of intensity function

$$\widehat{\lambda}_\theta^{(i)}(x|\mathcal{H}_{t(x)}) \qquad \hat{s}_\theta^{(i)}(x|\mathcal{H}_{t(x)}) := \nabla_x \log p_i(x|\mathcal{H}_{t(x)})$$

Benefits: (a) Flexible parameterization; (b) Accurate estimation

Proposed scoring rule: Weighted Hyvarinen score (WHS)

$$H_{\mathbf{w}}^{(i)}(x|\mathcal{H}_{t(x)}) = \left\| \underbrace{\mathbf{w}(x)}_{\text{Weighting function}} \odot \underbrace{\hat{s}_\theta^{(i)}(x|\mathcal{H}_{t(x)})}_{\text{Parametrized score model}} \right\|_2^2 + \mathrm{Tr}\left(\nabla_x \left[\mathbf{w}(x) \odot \hat{s}_\theta^{(i)}(x|\mathcal{H}_{t(x)})\right]\right)$$

Why weighting? "Correct" for boundary conditions to become a proper scoring rule

Weighted Hyvarinen Score CUSUM statistics:

$$\hat{z}(S,t) \approx \left\{ \underbrace{\max_{i=1,\dots,N(t)} \sum_{j: x_j \in [t(x_i), t)\times S}}_{\text{Maximum of sum of WHS over all events in the spatio-temporal domain}} \left[H_{\mathbf{w}}(x_j|\mathcal{H}_{t(x_j)}) - H_{\mathbf{w}}(x_j|\mathcal{H}_{t(x_j)})\right] \right\}^+ .$$

Hypothesis class of regions-at-risk

Exceed τ:

$$T_\tau = \inf\{t : \exists S \in \mathcal{R} \text{ s.t. } z(S,t) \geq \tau\},$$

$$\mathcal{S}_\tau = \{S : z(S,T) \geq \tau\}.$$