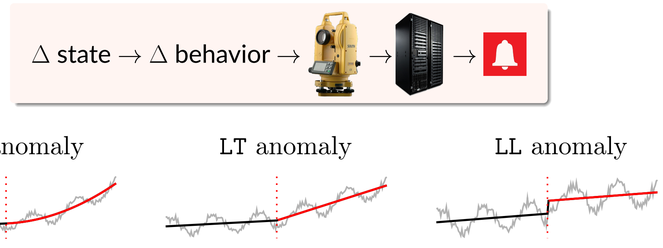
Kalman-based Anomaly Detection and Intervention via Imitation Learning

Zhanwen Xin, James-A. Goulet

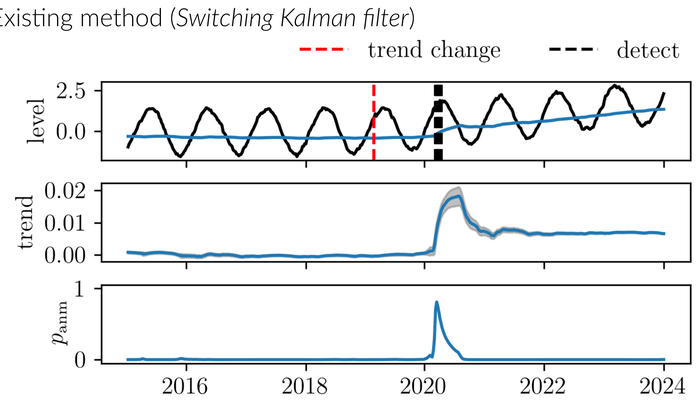
Department of Civil, Geological and Mining Engineering, Polytechnique Montreal, Montreal, Canada

Anomalies in time series

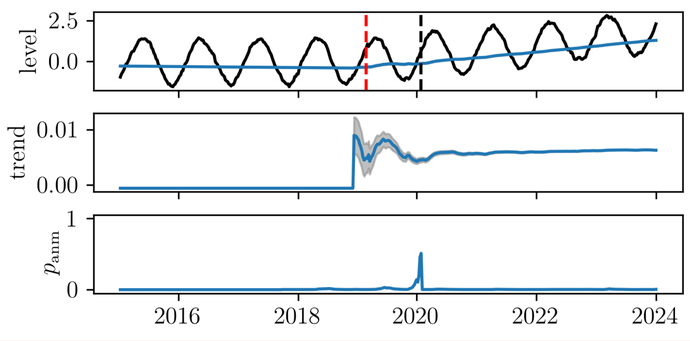


Too long; didn't read

Existing method (*Switching Kalman filter*)



Our new method



- 🔔 × 6 V.S. 1
- Time to search the new trend: 2 years faster

Online probabilistic Kalman methods

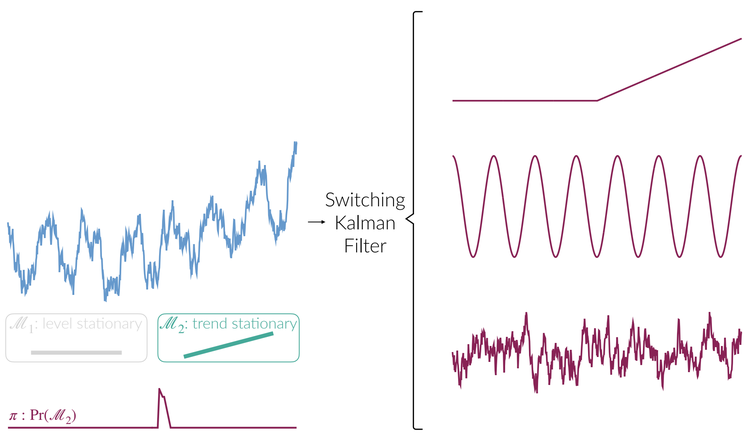
Real-time state estimation using Bayesian inference as new data arrives

- ***Kalman filter*** use observation until current time step
- ***Kalman smoother*** use both past and future observations

$$
\begin{array}{lccccccccc}
t & 0 & & 1 & & t & & \mathtt{T}-1 & & \mathtt{T} \\
\text{Filter:} & \underset{\boldsymbol{\Sigma}_0}{\boldsymbol{\mu}_0} & \rightarrow & \underset{\boldsymbol{\Sigma}_{1|1}}{\boldsymbol{\mu}_{1|1}} & \rightarrow \cdots \rightarrow & \underset{\boldsymbol{\Sigma}_{t|t}}{\boldsymbol{\mu}_{t|t}} & \rightarrow \cdots \rightarrow & \underset{\boldsymbol{\Sigma}_{\mathtt{T}-1|\mathtt{T}-1}}{\boldsymbol{\mu}_{\mathtt{T}-1|\mathtt{T}-1}} & \rightarrow & \underset{\boldsymbol{\Sigma}_{\mathtt{T}|\mathtt{T}}}{\boldsymbol{\mu}_{\mathtt{T}|\mathtt{T}}} \\
\text{Smooth:} & \underset{\boldsymbol{\Sigma}_{0|\mathtt{T}}}{\boldsymbol{\mu}_{0|\mathtt{T}}} & \leftarrow & \underset{\boldsymbol{\Sigma}_{1|\mathtt{T}}}{\boldsymbol{\mu}_{1|\mathtt{T}}} & \leftarrow \cdots \leftarrow & \underset{\boldsymbol{\Sigma}_{t|\mathtt{T}}}{\boldsymbol{\mu}_{t|\mathtt{T}}} & \leftarrow \cdots \leftarrow & \underset{\boldsymbol{\Sigma}_{\mathtt{T}-1|\mathtt{T}}}{\boldsymbol{\mu}_{\mathtt{T}-1|\mathtt{T}}} & &
\end{array}
$$

Regime change detection

- ***Switching Kalman filter*** (SKF) different dynamics over time



Generic sub-components in ***state-space models*** (SSM):

..., implemented in an open-source Python package **canari**.

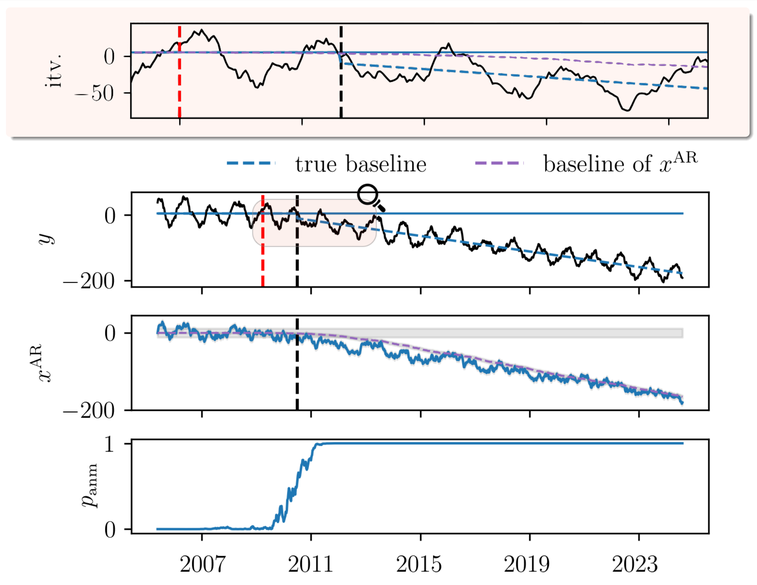
TAGI to quantify intervention uncertainty

Tractable Approximate Gaussian Inference (TAGI):

- Bayesian analytical inference { parameters / hidden states
- End-to-end treatment of uncertainty
- TAGI → CNN, LSTM, GAN, RL, ...
- No gradient backpropagation
- Open-source codebase (pip install pytagi)

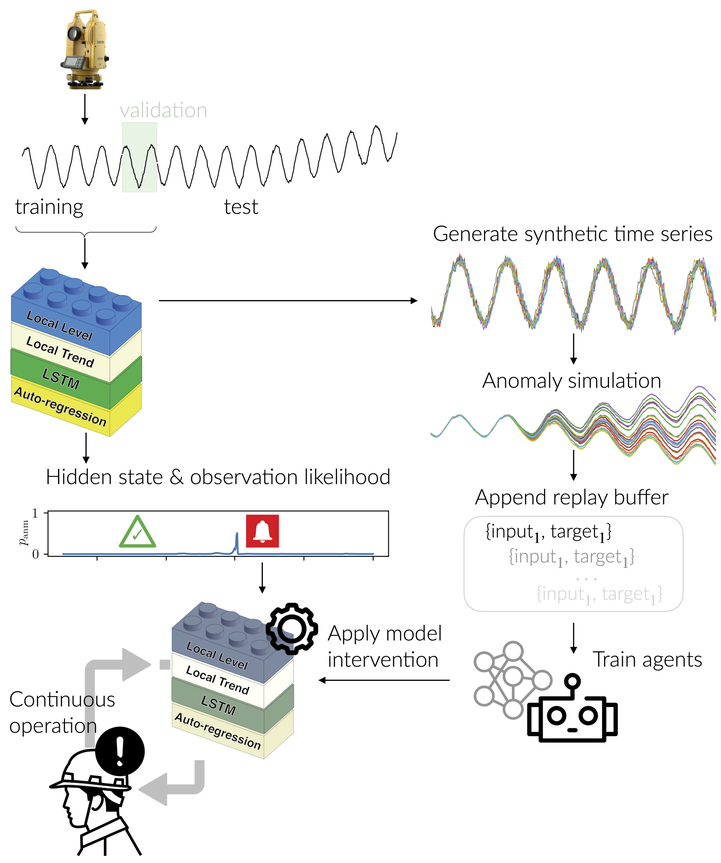
cuTAGI

A new detection method using hidden-state likelihood and imitation learning

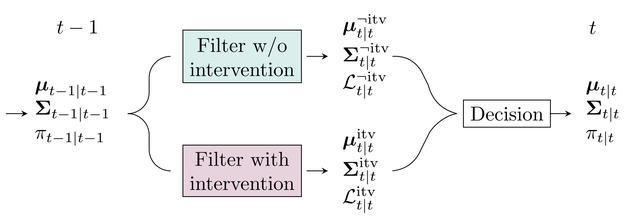


- 🔔 New proxy: $\mathcal{L}(x^{\text{LTd}}$ is stationary)
- ⚙ Intervention: correct the baseline by using the drift in x^{AR}
- ⌛ Drift is delayed
- ⚛ Imitation learning: delayed drift → intervention
- 🚀 Faster adaptation in new regime

Procedure



Hidden-states likelihood for detection

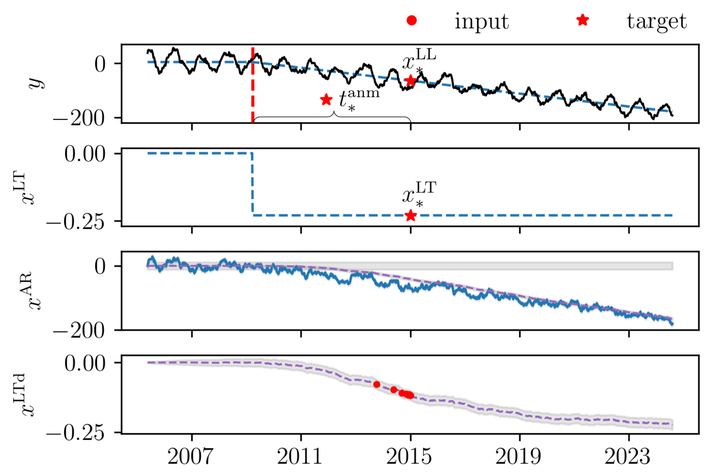


$$\pi_{t|t} = \Pr(\text{itv}|\boldsymbol{y}_{1:t}) = \frac{\mathcal{L}^{y,\text{itv}}_{t|t} \cdot \mathcal{L}^{x^{\text{LTd}},\text{itv}}_{t|t} \cdot p(a^{\text{itv}})}{\mathcal{L}^{y,\neg\text{itv}}_{t|t} \cdot \mathcal{L}^{x^{\text{LTd}},\neg\text{itv}}_{t|t} \cdot p(a^{\neg\text{itv}}) + \mathcal{L}^{y,\text{itv}}_{t|t} \cdot \mathcal{L}^{x^{\text{LTd}},\text{itv}}_{t|t} \cdot p(a^{\text{itv}})}$$

x^{LTd}: local trend of the drifting autoregressive hidden state x^{AR}

$\pi_{t|t}$: probability of anomaly, p_{anm}, at time t using all observation $\boldsymbol{y}_{1:t}$

Imitation learning agent for intervention



Input: $\boldsymbol{x}^{\text{LTd}}_{\mathcal{H}}$, $\mathcal{H} = \{t, t-1, t-2, t-4, ..., t-64\}$

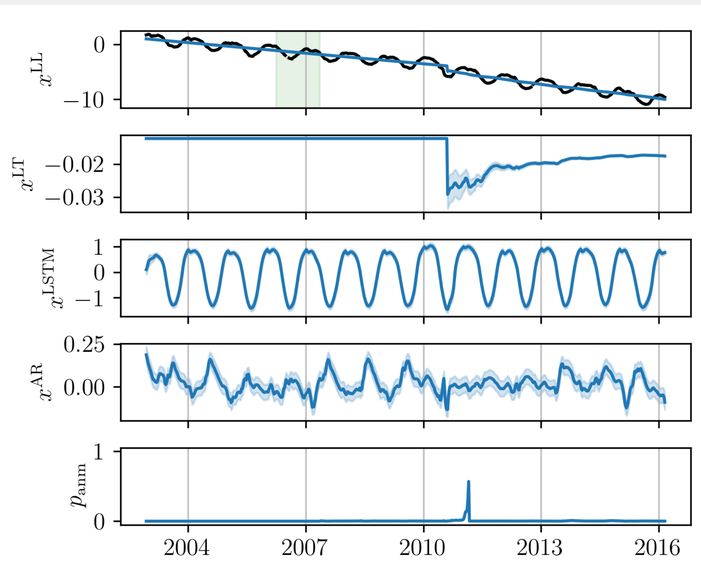
Target: $\{x^{\text{LL}}_*,\ x^{\text{LT}}_*,\ t^{\text{anm}}_*\}$

Results on real time series

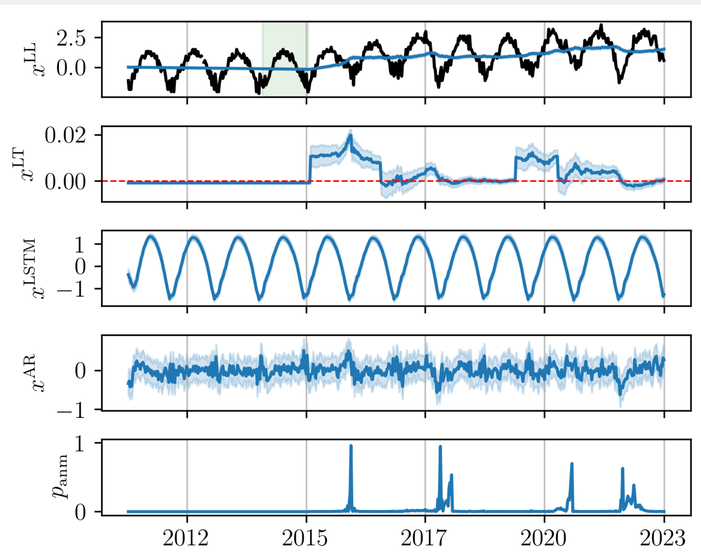
Features:

- Weekly frequency (can handle different frequencies)
- Missing data
- Baseline change recorded from a Canadian dam
- Unknown anomaly dynamics

TS #1: displacement

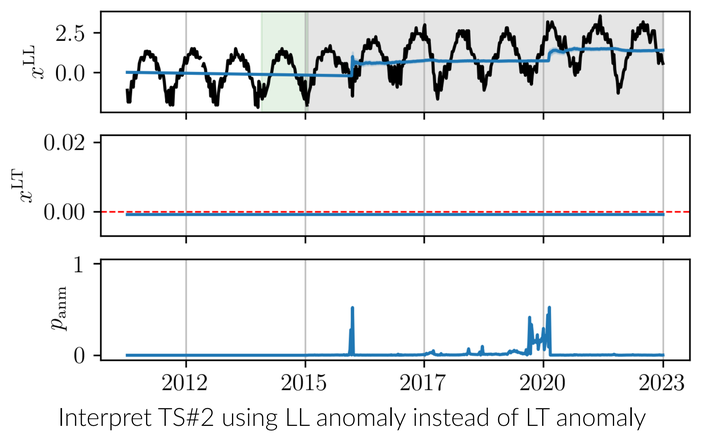


TS #2: crack opening



Ongoing developments

- **Stationary LSTM for stronger input signal**; all drifts → AR
- **Distinguish different anomaly type**; identify other anomalies and connect them to structural deterioration behaviours



Interpret TS#2 using LL anomaly instead of LT anomaly

- **Include explanatory variable**; SSM ⚠, anomaly detection layer ✗

Some methodology that we tried...

- **Constrain AR in SKF** ; improve the detectability and allow the tradeoff with false alarm rate
- **Classification** ; require anomaly label, typically hard to acquire
- **Reinforcement learning** ; DQN ⚠, PPO ⚠, but robustness ✗; challenge in environment design, especially in reward definition

References

- *Chapter 12: State-Space Models, Probabilistic Machine Learning for Civil Engineers* (Goulet, The MIT press, 2020)
- *Tractable approximate Gaussian inference for Bayesian neural networks* (Goulet, Nguyen, & Amiri, JMLR, 2021)
- *Analytically tractable hidden-states inference in Bayesian neural networks* (Nguyen and Goulet, Journal-to-conference track, ICLR 2024)
- *Analytically tractable heteroscedastic uncertainty quantification in Bayesian neural networks for regression tasks* (Deka, Nguyen & Goulet, Neurocomputing, 2024)
- Coupling LSTM Neural Networks and SSM through Analytically Tractable Inference, (Vuong, Nguyen & Goulet, International Journal of Forecasting, 2024)