Integrating Causal Inference and Deep Learning for Spatiotemporal Decision-Making

Miruna Oprescu

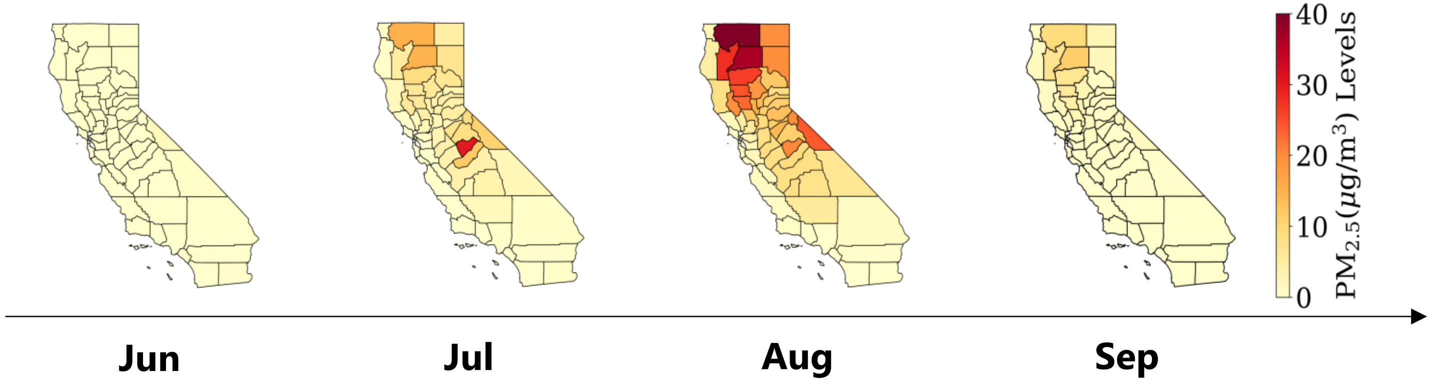
Joint work with David K. Park[2], Xihaier Luo[2], Shinjae Yoo[2], and Nathan Kallus[1]

[1] Cornell University, Cornell Tech

[2] Computing and Data Sciences, Brookhaven National Laboratory

Decision-Making in Spatiotemporal Contexts

- **Spatiotemporal Data**
  - Observations that vary across both spatial and temporal dimensions. E.g.: $PM_{2.5}$ levels during the 2018 California wildfires.



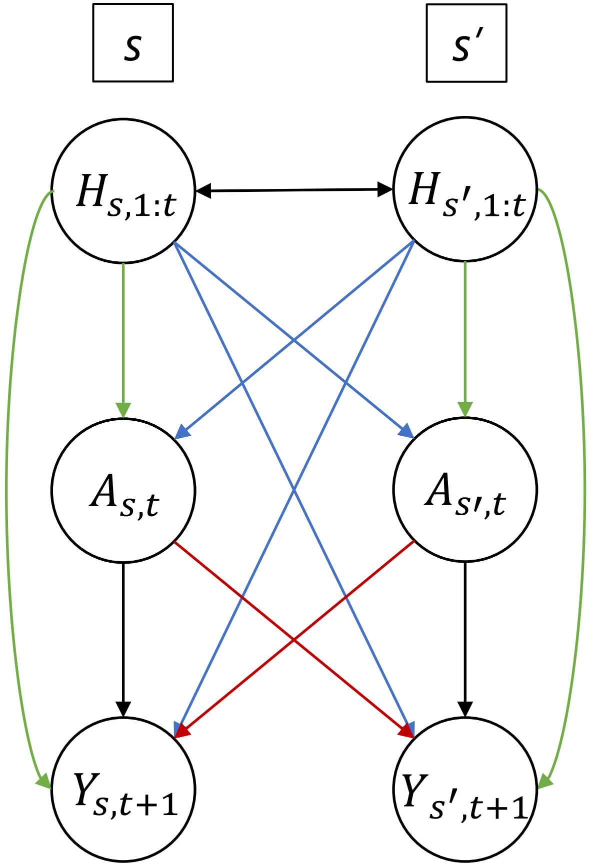
  - Often sourced from satellites, ground sensors, and weather stations, capturing how conditions evolve day by day and region by region.
- **Spatiotemporal Interventions**
  - Real-world actions or policies applied across space and time—such as wildfire prevention or pollution control measures—that shape local and regional outcomes (e.g., $PM_{2.5}$ levels, public health).

Decision-Making in Spatiotemporal Contexts

- **Notation**
  - Time $t \in \{1, \ldots, T\}$, spatial index $s \in \mathbb{G}$.
  - **Features (Covariates):** $X_{s,1}, X_{s,2}, \ldots, X_{s,T}$.
  - **Interventions (Treatments):** $A_{s,1}, A_{s,2}, \ldots, A_{s,T} \in \{0,1\}$.
  - **Outcomes:** $Y_{s,1}, Y_{s,2}, \ldots, Y_{s,T}$.
  - **History:** $H_{s,1:t} = (X_{s,1:t}, Y_{s,1:t}, A_{s,1:t-1})$.
  - Shorthand:

    $W_{s,1:t} = \{W_{s,1}, W_{s,2}, \ldots, W_{s,t}\}$, $\boldsymbol{W}_{1:t} = \{W_{s,1:t}: \forall s \in \mathbb{G}\}$

    for any $W \in \{X, A, Y, H\}$.



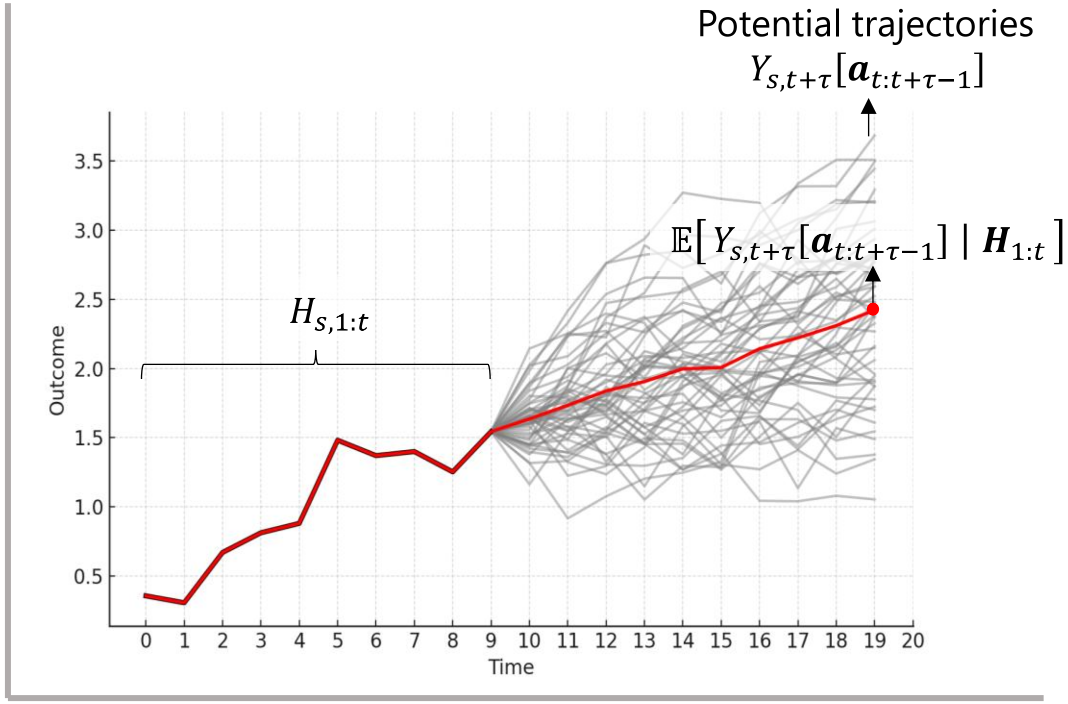
Schematic of the spatiotemporal data (X, A, Y, H) across time t and location s.

Decision-Making in Spatiotemporal Contexts

- **Counterfactuals:**

$$\mathbb{E}[\,\boldsymbol{Y}_{t+\tau}[\boldsymbol{a}_{t:t+\tau-1}] \mid \boldsymbol{H}_{1:t} = \boldsymbol{h}_{1:t}\,]$$

  - Average potential outcome after τ time steps under a series of fixed τ interventions, $\boldsymbol{a}_{t:t+\tau-1}$, given an observed history $\boldsymbol{h}_{1:t}$.
  - “What if stricter wildfire prevention measures had been implemented 2 weeks earlier—how would PM2.5 and health outcomes change over τ time steps?”



- **Treatment Effects:**

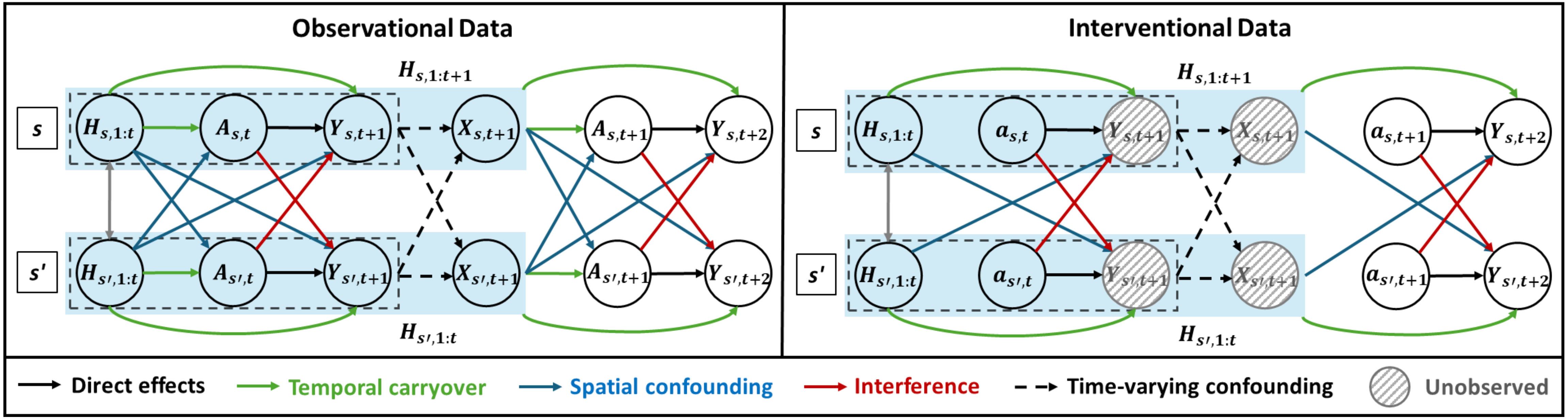
$$\mathbb{E}[\,\boldsymbol{Y}_{t+\tau}[\boldsymbol{a}_{t:t+\tau-1}] \mid \boldsymbol{H}_{1:t} = \boldsymbol{h}_{1:t}\,] - \mathbb{E}[\,\boldsymbol{Y}_{t+\tau}[\boldsymbol{a}'_{t:t+\tau-1}] \mid \boldsymbol{H}_{1:t} = \boldsymbol{h}_{1:t}\,]$$

  - “What was the effect of wildfire smoke on health outcomes over τ time steps?”

Challenges in Spatiotemporal Causal Inference

1. Single Spatiotemporal Chain

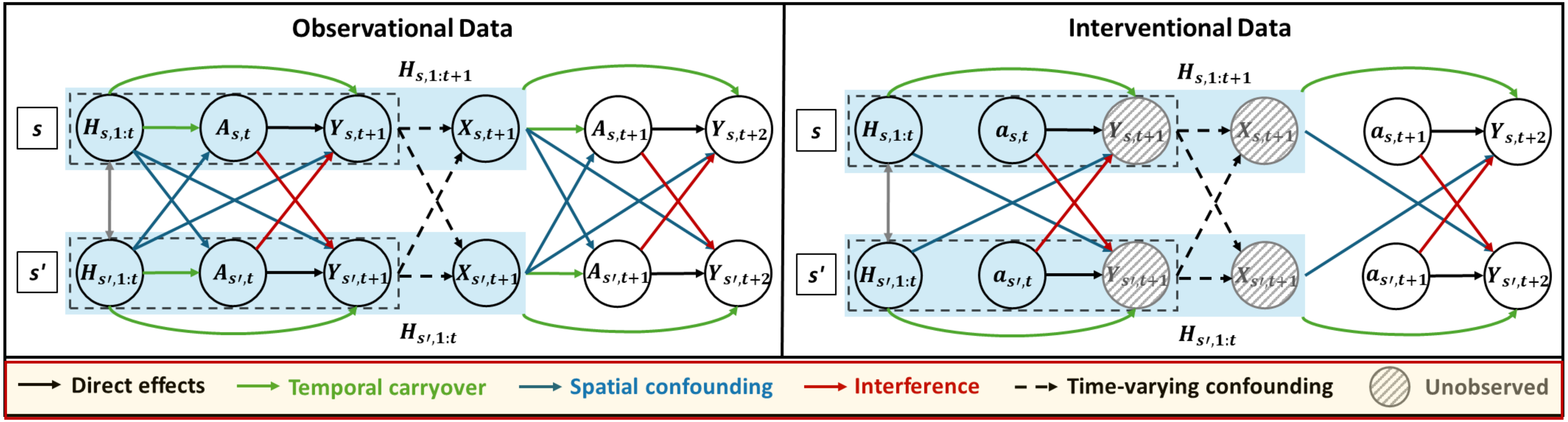
- We often have only one “realization” of space and time, rather than multiple parallel series from the same system.
- Challenging to isolate causal effects in this setting, since many methods rely on having multiple independent samples to tease out the impact of interventions.



Challenges in Spatiotemporal Causal Inference

2. Complex Space-Time Dependencies

- Observations at different locations and times can strongly influence one another, complicating standard causal analyses.
- Most ST causal inference works use strong modeling priors (e.g., linear models, Poisson processes, etc.).

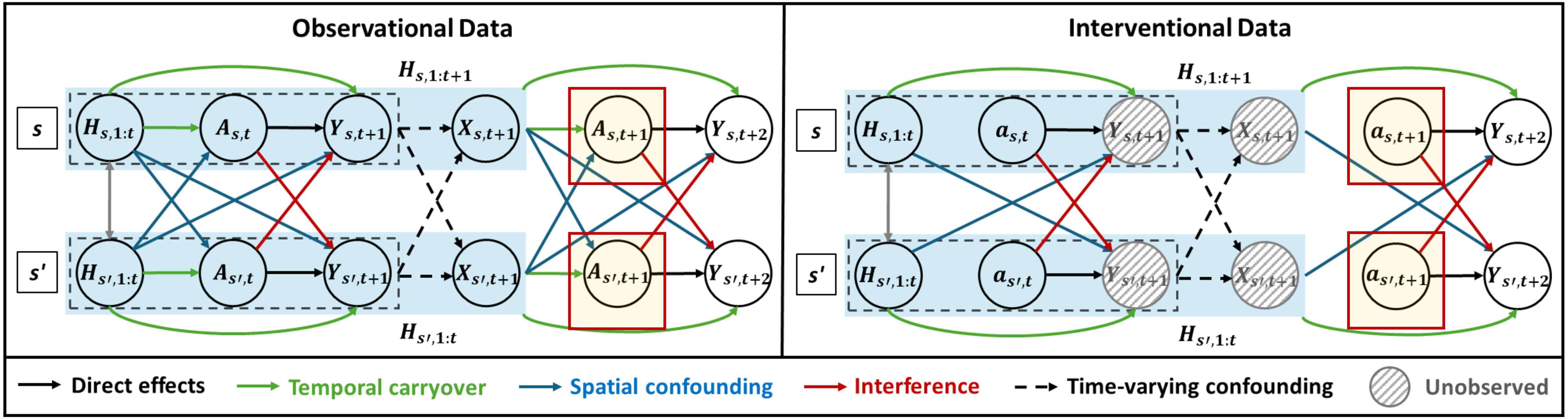


Challenges in Spatiotemporal Causal Inference

3(a). Observational vs. Interventional Data

- We need to learn features of an interventional distribution (new policy scenarios) from observational data, where interventions were applied differently (or non-randomly).
- In other words:

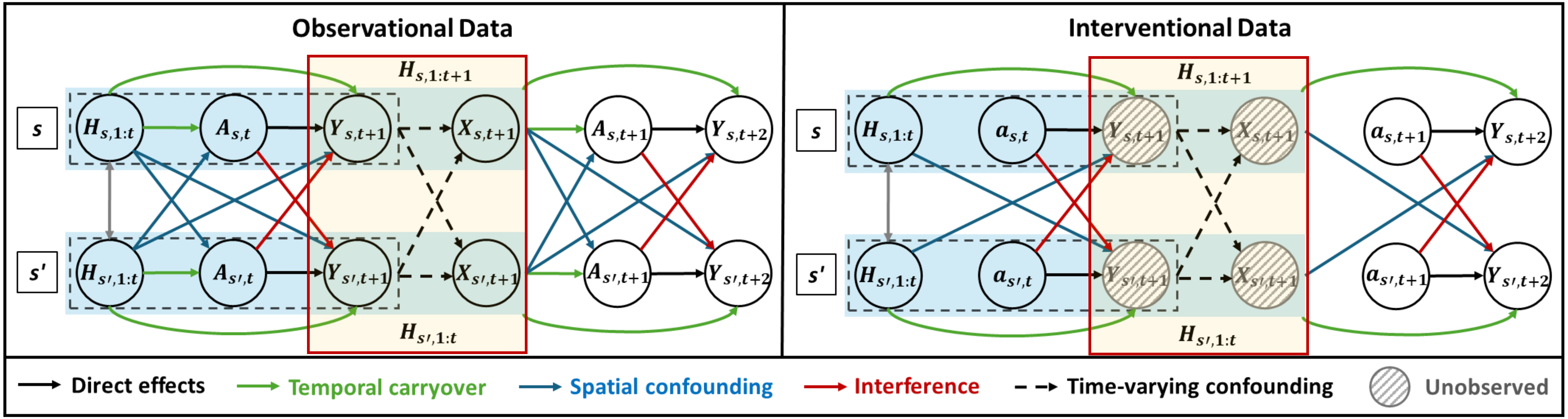
$$\mathbb{E}[\boldsymbol{Y}_{t+\tau}[\boldsymbol{a}_{t:t+\tau-1}] \mid \boldsymbol{H}_{1:t} = \boldsymbol{h}_{1:t}] \neq \mathbb{E}[\boldsymbol{Y}_{t+\tau} \mid \boldsymbol{H}_{1:t} = \boldsymbol{h}_{1:t}, \boldsymbol{A}_{t:t+\tau-1} = \boldsymbol{a}_{t:t+\tau-1}]$$



Challenges in Spatiotemporal Causal Inference

3(b). Time-Varying Confounders

- *A confounder* is any variable that affects both treatments and outcomes, and must be controlled to avoid biased causal estimates.
- A *time-varying confounder* is a variable that affects both future treatments and outcomes, creating feedback loops (e.g. past interventions shape future covariates, which in turn drive subsequent interventions and outcomes).



Talk Overview

1. **Identification** of Spatiotemporal Causal Effects
   - Representation-Based Time Invariance
   - Causal Inference with Time-Varying Confounders

2. **Estimation** of Spatiotemporal Causal Effects
   - GST-UNet Architecture
   - GST-UNet Training and Inference

3. **Empirical** Results
   - Synthetic Data
   - Effect of Wildfires on Respiratory Illness

Talk Overview

1. **Identification** of Spatiotemporal Causal Effects
    - Representation-Based Time Invariance
    - Causal Inference with Time-Varying Confounders

2. **Estimation** of Spatiotemporal Causal Effects
    - GST-UNet Architecture
    - GST-UNet Training and Inference

3. **Empirical** Results
    - Synthetic Data
    - Effect of Wildfires on Respiratory Illness

Using a Single Spatiotemporal Chain

Assumption 1: Representation-Based Time Invariance

- There exists an embedding $\phi: \mathcal{H} \times \mathcal{A} \to Z \subset \mathbb{R}^h$ such that, once we condition on $z = \phi(\boldsymbol{H}_{1:t}, \boldsymbol{A}_t)$ the distribution of $(\boldsymbol{X}_{t+1}, \boldsymbol{Y}_{t+1})$ does not explicitly depend on t. Formally:

$$p(\boldsymbol{X}_{t+1}, \boldsymbol{Y}_{t+1} \mid \phi(\boldsymbol{H}_{1:t}, \boldsymbol{A}_t) = z) = p(\boldsymbol{X}_{t'+1}, \boldsymbol{Y}_{t'+1} \mid \phi(\boldsymbol{H}_{1:t'}, \boldsymbol{A}_{t'}) = z)$$

Splicing the Single Time Series

- For each $t \in \{1, \ldots, T - \tau\}$, define a "prefix"

$$\boldsymbol{P}_t^\tau = (\boldsymbol{X}_{1:t+\tau}, \boldsymbol{A}_{1:t+\tau}, \boldsymbol{Y}_{1:t+\tau})$$

- Under representation-based time invariance, conditioning on $\phi(\boldsymbol{H}_{1:t}, \boldsymbol{A}_t)$ renders the distribution of $\boldsymbol{Y}_{t+\tau}$ independent of t.
- We can then write expectations over these prefixes as

$$\mathbb{E}_{\boldsymbol{P}}[\boldsymbol{Y}_{t+\tau} \mid \phi(\boldsymbol{H}_{1:t}, \boldsymbol{A}_t)]$$

Causal Inference with Time-Varying Confounders

Assumption 2 (Standard Causal Inference Assumptions)

- Consistency: $\boldsymbol{Y}_{t+\tau} = \boldsymbol{Y}_{t+\tau}[\boldsymbol{a}_{t:t+\tau-1}]$
- Positivity: $\boldsymbol{P}\big(A_{s,t} = a_{s,t} \mid \boldsymbol{H}_{1:t} = \boldsymbol{h}_{1:t} \big) > 0$ for any feasible $\boldsymbol{h}_{1:t}$.
- Sequential Unconfoundedness: $\boldsymbol{Y}_{t+1:T}[\boldsymbol{a}_{t+1:T}] \perp \boldsymbol{A}_t \mid \boldsymbol{H}_{1:t}$.

Theorem 1 (Identification under Assumptions 1&2 – Part 1)

Let $\boldsymbol{H}^{\boldsymbol{a}}_{1:t+k} = (\boldsymbol{H}_{1:t+k}, [\boldsymbol{A}_{1:t-1}, \boldsymbol{a}_{t:t+k-1}], \boldsymbol{Y}_{1:t+k})$. Then:

$$\mathbb{E}[\boldsymbol{Y}_{t+\tau}[\boldsymbol{a}_{t:t+\tau-1}] \mid \boldsymbol{H}_{1:t} = \boldsymbol{h}_{1:t}]$$

$$= \int \mathbb{E}_P[Y_{t+\tau} \mid \phi(\boldsymbol{h}^{\boldsymbol{a}}_{1:t+\tau-1}, \boldsymbol{a}_{t+\tau-1})] \prod_{k=1}^{\tau} p(x_{t+k}, y_{t+k} \mid \phi(\boldsymbol{h}^{\boldsymbol{a}}_{1:t+k-1}, \boldsymbol{a}_{t+k-1}))d(x_{t+k}, y_{t+k})$$

$$= \underbrace{\mathbb{E}_P[\ldots \mathbb{E}_P[Y_{t+\tau} \mid \phi(\boldsymbol{H}^{\boldsymbol{a}}_{1:t+\tau-1}, \boldsymbol{a}_{t+\tau-1})] \mid \phi(\boldsymbol{H}^{\boldsymbol{a}}_{1:t+\tau-2}, \boldsymbol{a}_{t+\tau-2})] \ldots \mid \phi(\boldsymbol{H}_{1:t}, \boldsymbol{a}_t) = \phi(\boldsymbol{h}_{1:t}, \boldsymbol{a}_t)]}_{\text{Iterative G-Computation}}$$

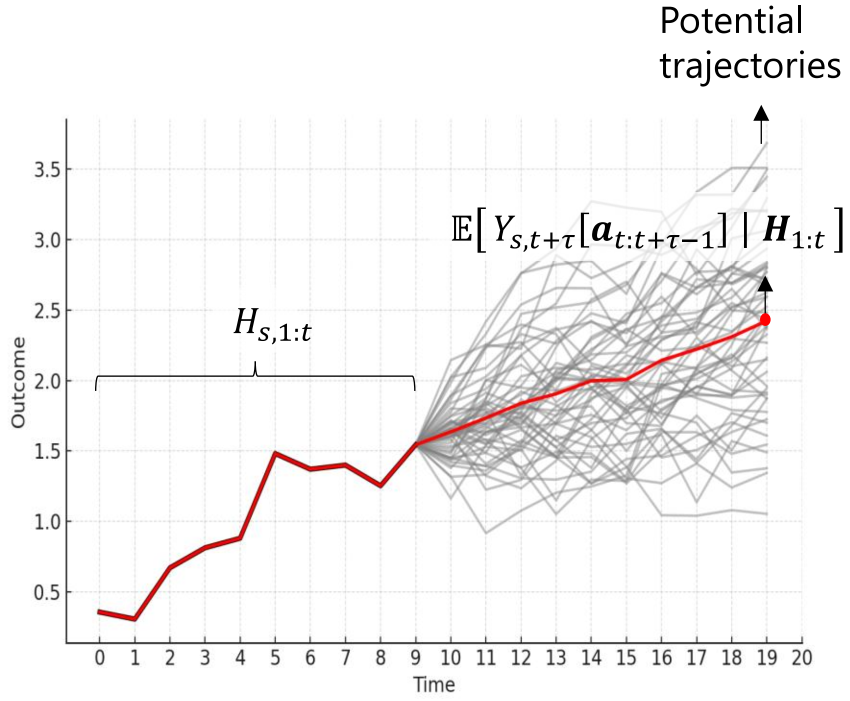
Bonus: Time-Varying Confounders without Interference

Estimation Strategy via IPW Estimator (No Interference)

$$\hat{Y}_{t+\tau} = \left(\prod_{l=t}^{t+\tau} \frac{\mathbb{I}[A_l = a_l]}{\hat{\pi}(a_l \mid H_{1:l})} \right) Y_{t+\tau}$$

$$\hat{\mathbb{E}}[\, Y_{t+\tau}[a_{t:t+\tau-1}] \mid H_{1:t} = h_{1:t}\,] = \mathbb{E}[\hat{Y}_{t+\tau} \mid H_{1:t} = h_{1:t}]$$

- There is also a doubly robust alternative (see [6]).
- Can work with *unstructured interference* under additional assumptions, such as known exposure function and exposure ignorability.



Talk Overview

1. **Identification** of Spatiotemporal Causal Effects
    - Representation-Based Time Invariance
    - Causal Inference with Time-Varying Confounders

2. **Estimation** of Spatiotemporal Causal Effects
    - GST-UNet Architecture
    - GST-UNet Training and Inference

3. **Empirical** Results
    - Synthetic Data
    - Effect of Wildfires on Respiratory Illness

Estimation of Spatiotemporal Causal Effects via G-Computation

Task: Estimate

$$\mathbb{E}_P[\ldots\mathbb{E}_P[Y_{t+\tau} \mid \phi(\boldsymbol{H}^{\boldsymbol{a}}_{1:t+\tau-1}, \boldsymbol{a}_{t+\tau-1})] \mid \phi(\boldsymbol{H}^{\boldsymbol{a}}_{1:t+\tau-2}, \boldsymbol{a}_{t+\tau-2})] \ldots \mid \phi(\boldsymbol{H}_{1:t}, \boldsymbol{a}_t) = \phi(\boldsymbol{h}_{1:t}, \boldsymbol{a}_t)]$$

Iterative G-Computation via Recursive Regression [3]

1. Last Step:

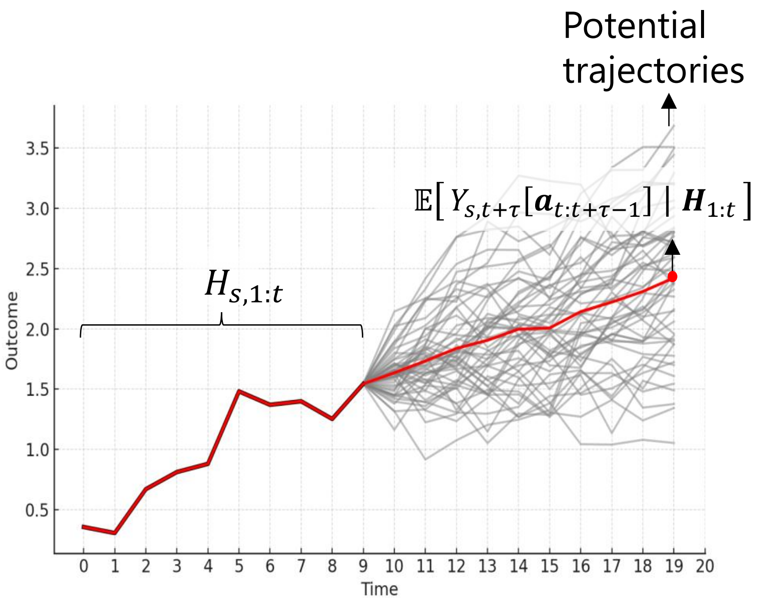
$$Q_\tau(\boldsymbol{H}_{1:t+\tau-1}, \boldsymbol{A}_{t+\tau-1}) = \mathbb{E}_P[\boldsymbol{Y}_{t+\tau} \mid \phi(\boldsymbol{H}_{1:t+\tau-1}, \boldsymbol{A}_{t+\tau-1})]$$

2. Recursive Steps (for $k = \tau - 1, \ldots, 1$):

$$Q_k(\boldsymbol{H}_{1:t+k-1}, \boldsymbol{A}_{t+k-1}) = \mathbb{E}_P[Q_{k+1}(\boldsymbol{H}^{\boldsymbol{a}}_{1:t+k}, \boldsymbol{A}_{t+k}) \mid \phi(\boldsymbol{H}_{1:t+k-1}, \boldsymbol{A}_{t+k-1})]$$

3. Result:

$$\mathbb{E}_P[\boldsymbol{Y}_{t+\tau}[\boldsymbol{a}_{t:t+\tau-1}] \mid \phi(\boldsymbol{H}_{1:t}, \boldsymbol{a}_t) = \phi(\boldsymbol{h}_{1:t}, \boldsymbol{a}_t)] = Q_1(\boldsymbol{h}_{1:t}, \boldsymbol{a}_t)$$



Estimation of Spatiotemporal Causal Effects via G-Computation

Task: Estimate

$$\mathbb{E}_P[\ldots \mathbb{E}_P[Y_{t+\tau} \mid \phi(\boldsymbol{H}^{\boldsymbol{a}}_{1:t+\tau-1}, \boldsymbol{a}_{t+\tau-1})] \mid \phi(\boldsymbol{H}^{\boldsymbol{a}}_{1:t+\tau-2}, \boldsymbol{a}_{t+\tau-2})] \ldots \mid \phi(\boldsymbol{H}_{1:t}, \boldsymbol{a}_t) = \phi(\boldsymbol{h}_{1:t}, \boldsymbol{a}_t)]$$

Iterative G-Computation via Recursive Regression [3]

1. Last Step:

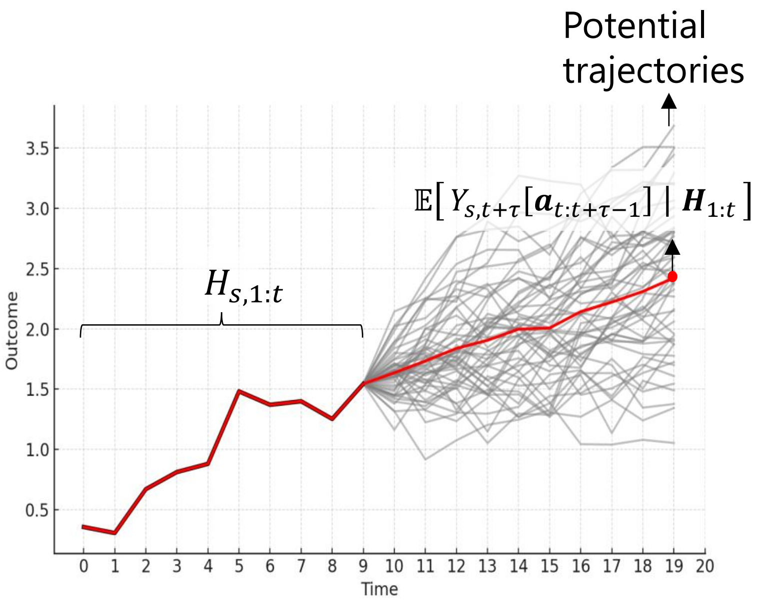
$$\hat{Q}_\tau(\boldsymbol{H}_{1:t+\tau-1}, \boldsymbol{A}_{t+\tau-1}) = \widehat{\mathbb{E}}_P[\boldsymbol{Y}_{t+\tau} \mid \hat{\phi}(\boldsymbol{H}_{1:t+\tau-1}, \boldsymbol{A}_{t+\tau-1})]$$

2. Recursive Steps (for $k = \tau - 1, \ldots, 1$):

$$\hat{Q}_k(\boldsymbol{H}_{1:t+k-1}, \boldsymbol{A}_{t+k-1}) = \widehat{\mathbb{E}}_P[\hat{Q}_{k+1}(\boldsymbol{H}^{\boldsymbol{a}}_{1:t+k}, \boldsymbol{A}_{t+k}) \mid \hat{\phi}(\boldsymbol{H}_{1:t+k-1}, \boldsymbol{A}_{t+k-1})]$$

3. Result:

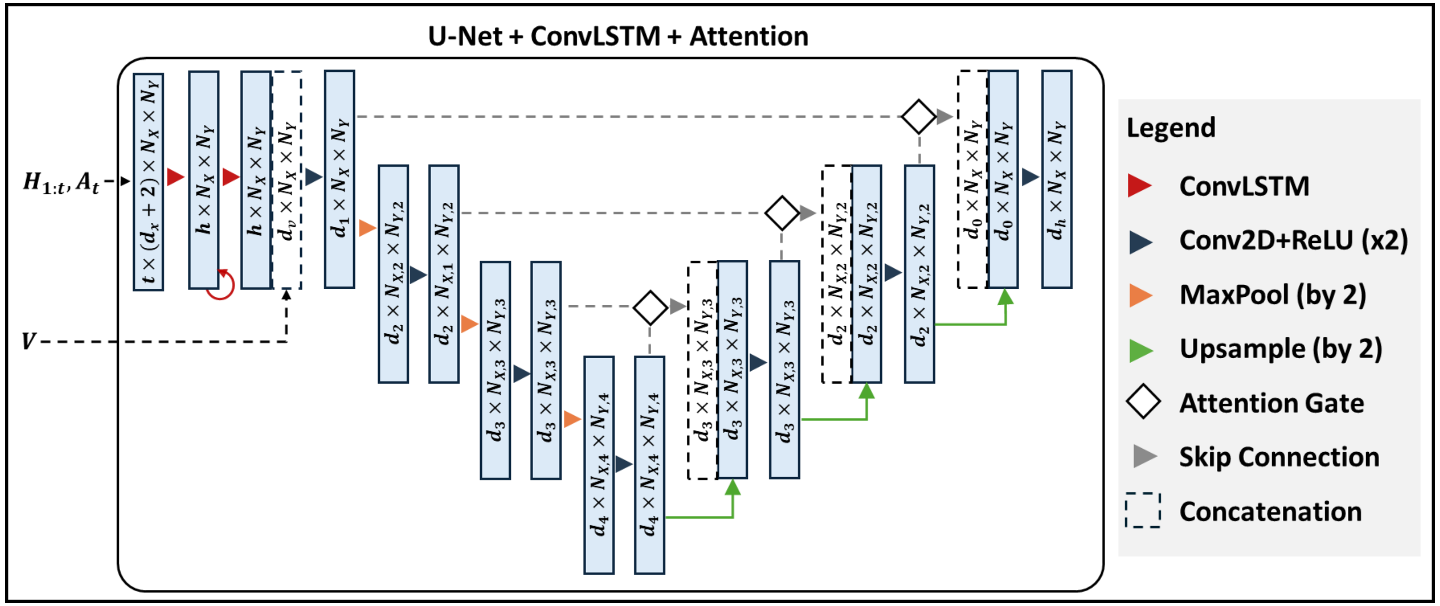
$$\widehat{\mathbb{E}}_P[\boldsymbol{Y}_{t+\tau}[\boldsymbol{a}_{t:t+\tau-1}] \mid \phi(\boldsymbol{H}_{1:t}, \boldsymbol{a}_t) = \phi(\boldsymbol{h}_{1:t}, \boldsymbol{a}_t)] = \hat{Q}_1(\boldsymbol{h}_{1:t}, \boldsymbol{a}_t)$$



Learning the Spatiotemporal Embedding ϕ

Approach: Use neural networks to capture spatiotemporal patterns

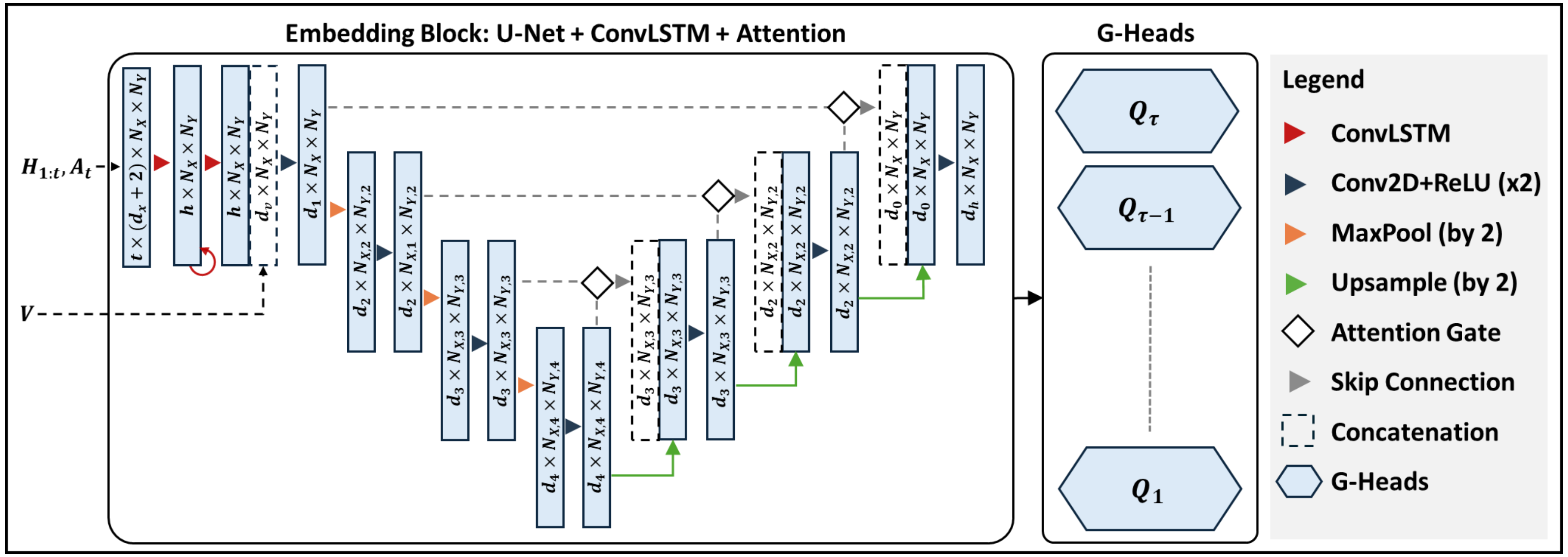
- **U-Net** for spatial dependencies [1]
  - Encoder-decoder architecture that captures multi-scale spatial features.
- **ConvLSTM** for temporal dynamics
  - Merges convolution and LSTM to model temporal dynamics within a single series.
- **Attention** to highlight key spatial regions and time steps [2].



Introducing the GST-UNet (Our Work)

G-computation **S**patio-**T**emporal **UNet** (**GST-UNet**):

- **Spatiotemporal Embedding:** U-Net + ConvLSTM + attention gates.
- **Neural Causal Modules:** G-computation heads (e.g. shallow feed-forward networks or convolutional layers) for iterative adjustment.
- **Key Innovation:** Flexible, end-to-end approach that avoids strong modeling assumptions and properly accounts for time-varying confounders.



GST-UNet End-to-End Architecture

GST-UNet Training and Inference

Algorithm 1 GST-UNet Training and Inference

Input: Horizon τ, prefix data $\{\mathbf{P}_t^{\tau}\}_{t=1}^{T-\tau}$, interventions $\mathbf{a}_{t:t+\tau-1}$, curriculum schedule $\alpha_k^{(e)}$, total epochs E.

Initialize: parameters θ (U-Net embedding + G-heads).

for $e = 1 \dots E$ **do**

  for $k = \tau \dots 1$ **do**

    (Supervision) For each prefix i, predict outcomes:

$$\widehat{Y}_{t+k}^{(i)} = Q_k(\phi(\mathbf{H}_{1:t+k-1}^{(i)}, \mathbf{A}_{t+k-1}^{(i)}); \theta).$$

    (Generation (detached)) For each prefix i, generate pseudo-outcomes:

$$\widetilde{Y}_{t+k}^{(i)} = \begin{cases} Q_k\Big(\phi((\mathbf{H}_{1:t+k-1}^{\mathbf{a}})^{(i)}, \mathbf{a}_{t+k-1}^{(i)}); \theta\Big), & k < \tau, \\ Y_{t+\tau}^{(i)}, & k = \tau. \end{cases}$$

where the observed $\mathbf{A}_{t:t+k-2}$'s were replaced with $\mathbf{a}_{t:t+k-2}$ in the history.

  end for

  (Loss aggregation) Compute the overall MSE loss

$$\mathcal{L}(\theta; e) = \frac{1}{\tau} \sum_{k=1}^{\tau} \alpha_k^{(e)} \sum_i \left(\widehat{Y}_{t+k}^{(i)} - \widetilde{Y}_{t+k+1}^{(i)}\right)^2.$$

  (Backward pass) Update θ by backpropagation.

end for

(Inference) Given a $\mathbf{h}_{1:t}$, return $Q_1(\phi(\mathbf{h}_{1:t}, \mathbf{a}_t); \widehat{\theta})$.

GST-UNet Training and Inference

Curriculum Training: $\mathcal{L}(\theta; e) = \frac{1}{\tau}\sum_{k=1}^{\tau} \boldsymbol{\alpha_k^{(e)}} \sum_i \left(\widehat{\boldsymbol{Y}}_{t+k}^{(i)} - \widetilde{\boldsymbol{Y}}_{t+k+1}^{(i)}\right)^2$

Curriculum Options:

- No curriculum: $\alpha_k^{(e)} = 1.$

  Issue: the later heads $(1, 2, \dots)$ train on noise while the earlier heads $(\tau, \tau - 1, \dots)$ learn. Can (and will) converge to suboptimal solution.

- Sequential head training: $\alpha_k^{(e)} = \mathbb{I}[e_k \leq e < e_{k+1}]$ for some increasing e_k.

  Issue: each Q_k head might attempt to tailor ϕ to its own objective (ϕ is much more expressive than Q_k), leading to misaligned training signals.

- **Hybrid curriculum:** let $p(e) = \min\{\tau, \lceil\frac{e}{e_c}\rceil\}$, where e_c is the curriculum period.

$$\alpha_k^{(e)} = \begin{cases} \dfrac{1}{p(e)}, & if\ k \in \{\tau, \tau - 1, \dots, \tau - p(e) + 1\} \\ 0, & otherwise \end{cases}$$

Talk Overview

1. **Identification** of Spatiotemporal Causal Effects
    - Representation-Based Time Invariance
    - Causal Inference with Time-Varying Confounders

2. **Estimation** of Spatiotemporal Causal Effects
    - GST-UNet Architecture
    - GST-UNet Training and Inference

3. **Empirical** Results
    - Synthetic Data
    - Effect of Wildfires on Respiratory Illness

Simulation Results on Synthetic Data

- **Data:** We generate $T = 200$ steps of a 64×64 grid of observational data from:

$$\boldsymbol{X}_t = \alpha_0 + \alpha_1 \boldsymbol{X}_{t-1} + \alpha_2 \boldsymbol{A}_{t-1} + \alpha_3 K_X * \boldsymbol{X}_{t-1} + \epsilon_X$$

$$\boldsymbol{A}_t \sim Bern\left(\sigma\left(\beta_1\left(\beta_0 + \frac{1}{L}\sum_{l=0}^{L-1} K_A * \boldsymbol{X}_{t-l}\right)\right)\right)$$

$$\boldsymbol{Y}_t = \gamma_0 + \gamma_1 (K_{YA} * \boldsymbol{A}_{t-1}) + \gamma_2 \frac{1}{L}\sum_{l=1}^{L} K_{YX} * \boldsymbol{X}_{t-l} + \gamma_3 \boldsymbol{Y}_{t-1} + \epsilon_Y$$

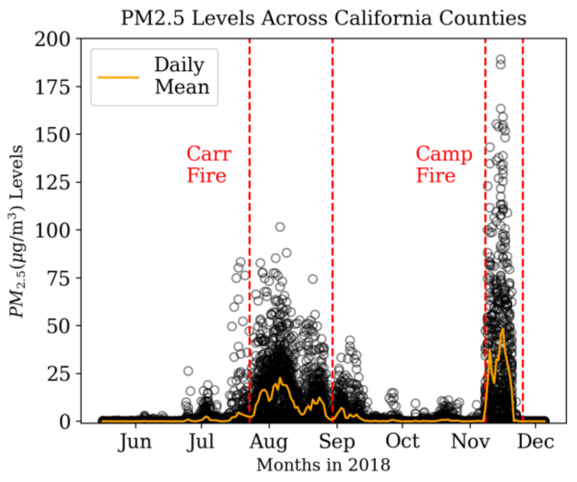
Note: "*" is the convolution operation and β_1 controls the time-varying confounding.

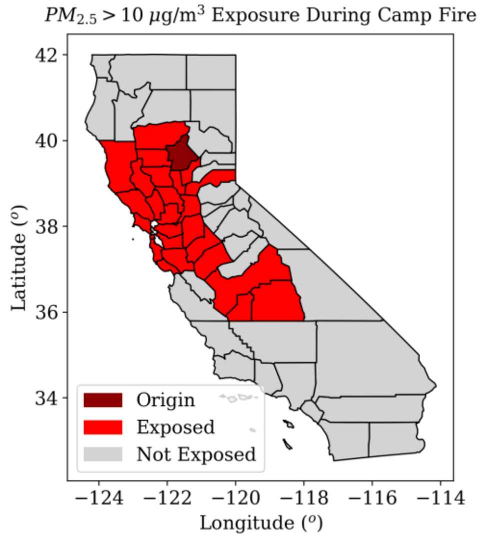
- **Results (RMSE):**

τ	Model	$\beta_1 = 0.0$	$\beta_1 = 0.5$	$\beta_1 = 1.0$	$\beta_1 = 1.5$	$\beta_1 = 2.0$	$\beta_1 = 2.5$	$\beta_1 = 3.0$
5	UNet+	**0.28 ± 0.00**	0.36 ± 0.00	0.54 ± 0.01	0.71 ± 0.01	0.81 ± 0.01	0.87 ± 0.01	0.97 ± 0.01
	STCINet	0.29 ± 0.00	0.38 ± 0.01	0.62 ± 0.01	0.80 ± 0.01	0.90 ± 0.01	1.03 ± 0.01	1.07 ± 0.01
	GST-UNet w/o Attention	0.50 ± 0.00	0.46 ± 0.00	0.51 ± 0.00	0.45 ± 0.01	0.47 ± 0.01	0.45 ± 0.01	0.52 ± 0.01
	GST-UNet w/o Curriculum	0.69 ± 0.00	0.64 ± 0.00	0.63 ± 0.00	0.61 ± 0.01	0.61 ± 0.01	0.61 ± 0.01	0.61 ± 0.01
	GST-UNet	0.33 ± 0.00 (+17.9%)	**0.35 ± 0.00** (**-2.8%**)	**0.40 ± 0.00** (**-25.9%**)	**0.44 ± 0.00** (**-38.0%**)	**0.40 ± 0.01** (**-50.6%**)	**0.42 ± 0.01** (**-51.7%**)	**0.50 ± 0.01** (**-48.5%**)
10	UNet+	**0.28 ± 0.00**	0.61 ± 0.00	1.18 ± 0.00	1.45 ± 0.00	1.71 ± 0.01	1.73 ± 0.01	1.75 ± 0.01
	STCINet	0.31 ± 0.00	0.68 ± 0.00	1.25 ± 0.00	1.47 ± 0.01	1.60 ± 0.01	1.66 ± 0.01	1.94 ± 0.01
	GST-UNet w/o Attention	0.42 ± 0.00	0.60 ± 0.00	0.61 ± 0.00	0.79 ± 0.01	1.07 ± 0.01	0.91 ± 0.01	1.02 ± 0.01
	GST-UNet w/o Curriculum	0.62 ± 0.00	0.88 ± 0.00	1.02 ± 0.00	1.08 ± 0.01	1.12 ± 0.01	1.15 ± 0.01	1.17 ± 0.01
	GST-UNet	0.38 ± 0.00 (+35.7%)	**0.55 ± 0.00** (**-9.8%**)	**0.68 ± 0.00** (**-42.4%**)	**0.73 ± 0.01** (**-49.7%**)	**0.85 ± 0.01** (**-50.3%**)	**0.85 ± 0.01** (**-50.9%**)	**0.85 ± 0.01** (**-51.4%**)

Case Study: Effect of Wildfire Smoke on Respiratory Illness during the 2018 California Camp Fire

- **Data (2018 California, county-level data [4]):**
  - **Covariates:** wind, temperature, precipitation, humidity, shortwave radiation
  - **"Treatment":** $PM_{2.5} > 10\ \mu g/m^3$ (unhealthy)
  - **Outcome:** Respiratory hospitalizations.
- **Counterfactual/ Policy-Relevant Question:**
  - How did unhealthy $PM_{2.5}$ (Camp Fire smoke) affect respiratory hospitalization?
  - If Camp Fire never occurred (i.e. $PM_{2.5}$ never exceeded 10 $\mu g/m^3$), how would the daily respiratory hospitalizations differ during the same time period?





Case Study: Effect of Wildfire Smoke on Respiratory Illness during the 2018 California Camp Fire

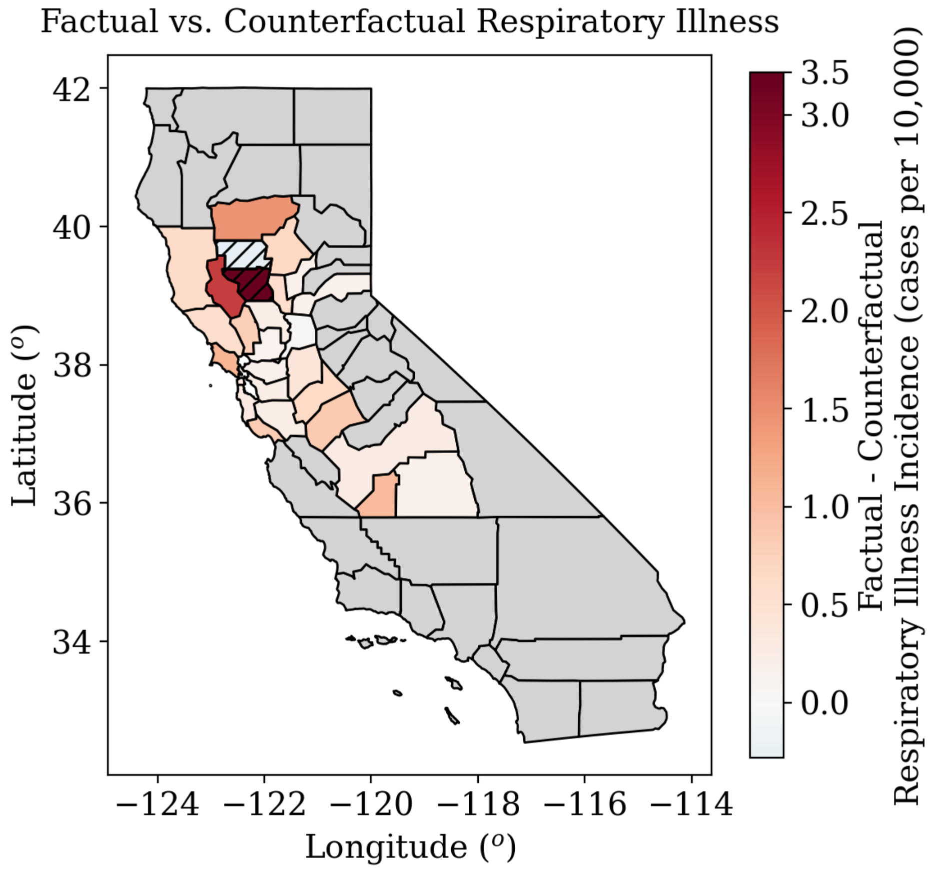
Results

The GST-UNet estimates that the peak period of the Camp Fire (November 8–17, 2018) contributed to an excess 4650 ([1890, 6535] 95% CI) (465 per day)[1] respiratory-related hospitalizations in the affected counties.

Baseline Predictions

- **UNet+:** 3911 ([–899, 5202] 95% CI)
- **STCINet:** 343 ([–3077, 3281] 95% CI)

[1] **Note:** This result aligns qualitatively with [4], who used a synthetic controls method and found about 259 excess daily cases from November 8–December 5 (including lower-intensity days, hence a smaller daily estimate).



Observed minus predicted daily respiratory admissions at Camp Fire peak. Hashed areas mark small-population counties (<30,000).

References

[1] O. Ronneberger, P. Fischer, T. Brox. *U-Net: Convolutional Networks for Biomedical Image Segmentation*. MICCAI 2015.

[2] O. Oktay, J. Schlemper, L. Le Folgoc, et al.
Attention U-Net: Learning Where to Look for the Pancreas. MIDL 2018.

[3] J. Robins and M. Hernán.
Estimation of the causal effects of time-varying exposures.
In Chapman & Hall/CRC Handbooks of Modern Statistical Methods, 2008.

[4] N. Letellier, M. Hale, K. U. Salim, et al.
Applying a two-stage generalized synthetic control approach to quantify the heterogeneous health effects of extreme weather events: A 2018 large wildfire in California event as a case study.
Environmental Epidemiology, 2025.

[5] D. Frauen, K. Hess, and S. Feuerriegel. *Model-agnostic meta-learners for estimating heterogeneous treatment effects over time*. ICLR 2025.

Thank You!

Paper: *GST-UNet: Spatiotemporal Causal Inference with Time-Varying Confounders.*

Miruna Oprescu, David K. Park, Xihaier Luo, Shinjae Yoo, Nathan Kallus (Under Review, 2025).

Contact: miruna@cs.cornell.edu.

Case Study: Effect of Wildfire Smoke on Respiratory Illness during the 2018 California Camp Fire

- Estimated county-level increases in respiratory ED visits attributable to the wildfire event, with 95% bootstrap confidence intervals.
- Population is reported in units of 10,000. Counties marked with * (hashed on the map) have smaller populations, which leads to greater uncertainty.

County	Mean	2.5%	97.5%	Population ($\times 10^4$)	Interval Width / Population
Tehama	37	-126	158	6.4	44.4
Butte	168	30	325	23.0	12.8
Glenn*	-52	-262	39	2.8	107.6
Colusa*	13	-158	107	2.1	124.0
Sutter	-18	-170	70	9.6	24.9
Napa	81	-41	192	13.9	16.8
Lake	103	-66	203	6.4	41.8
Solano	38	-79	173	44.6	5.6
Sacramento	202	-107	484	153.9	3.8