Reliable Causal Inference Under Unreliable Assumptions

Machine Learning Methods for Observational, Quasi-Experimental, and Structured Data

Miruna Oprescu
Cornell University • DOE Computational Science Graduate Fellow

B Exam • **Committee:** Nathan Kallus (Chair), Sarah Dean, Peter Frazier, Emma Pierson

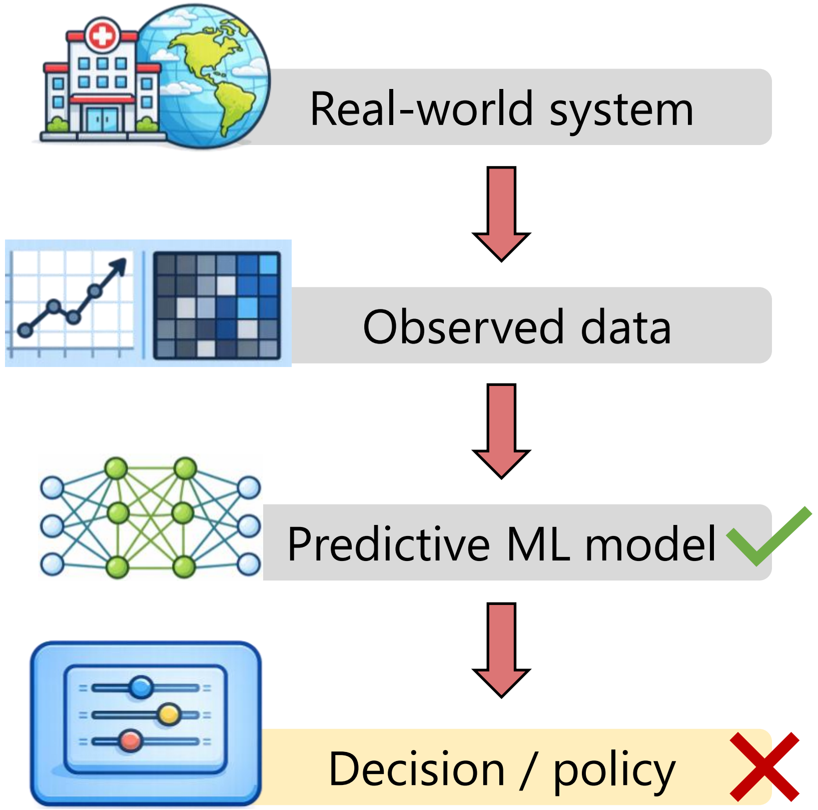
Machine Learning for High-Stakes Decision-Making

Healthcare
Should this ICU patient receive a more aggressive intervention?

Environmental policy
Would reducing smoke exposure improve health outcomes?

Scientific discovery
Which neural pathway should a treatment target?

AI systems
Should an AI assistant act, defer, or escalate to a human?



Predictive models learn correlations. Decisions require causal reasoning.

The Challenges of Causal Inference in Practice

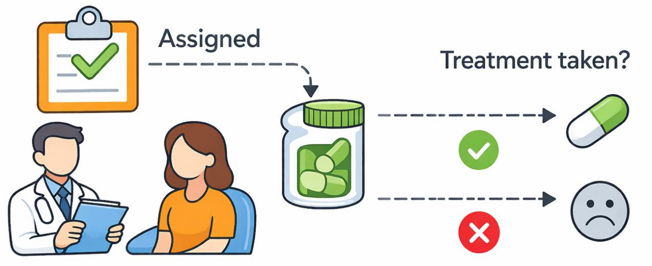
The gold standard is a randomized controlled trial, but that ideal is rarely realized in practice.

Observational Data

Intervention is not randomized.

- (Un)observed confounders create spurious correlations
- Point estimates can become misleading or uninformative.

Indirect Experiments



Assignment is randomized, but uptake is self-selected.

- Assignment does not guarantee treatment receipt
- Weak instruments and low compliance add noise

Structured Systems

Units interact across space, time, or networks.

- Outcomes depend on nearby treatments and past history
- Confounders evolve over time and affect future exposure

My work develops machine learning methods for causal inference in these imperfect settings.

My Research: ML Methods for Reliable Causal Inference

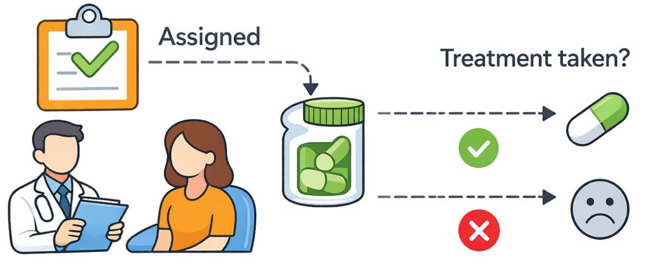
I develop ML methods that remain reliable when idealized assumptions do not fully hold.

Observational Data

- Debiased learning beyond average effects
- Partial ID and sharp bounds under hidden confounding

Oprescu *& Kallus (AISTATS '23)*
Oprescu *et al. (ICML '23; NeurIPS '24)*

Indirect Experiments



- Heterogeneous effects from weak experimental signals
- Adaptive experimentation with imperfect compliance

Oprescu *& Kallus (NeurIPS '24)*
Oprescu*, Kallus, & Cho (NeurIPS '25)*

Structured Systems

- Spatiotemporal effects under dynamic confounding
- Spatial and network effects under interference

Oprescu *et al. (NeurIPS '25)*
Khot, ***Oprescu*** *et al. (Preprint '25)*

Common thread: Flexible ML for reliable causal inference from imperfect data.

Today's Talk

1. Spatiotemporal Causal Inference
   - *Wildfire smoke and health effects*

2. Adaptive Experimentation Under Noncompliance
   - *Learning with imperfect control*

3. Future Directions
   - *Biomedical causal modeling, reliable evaluation, and decision-making under constraints*

Today's Talk

1. Spatiotemporal Causal Inference
    - *Wildfire smoke and health effects*
2. Adaptive Experimentation Under Noncompliance
    - *Learning with imperfect control*
3. Future Directions
    - *Biomedical causal modeling, reliable evaluation, and decision-making under constraints*

Example: Wildfire Smoke and Respiratory Health

We observe:

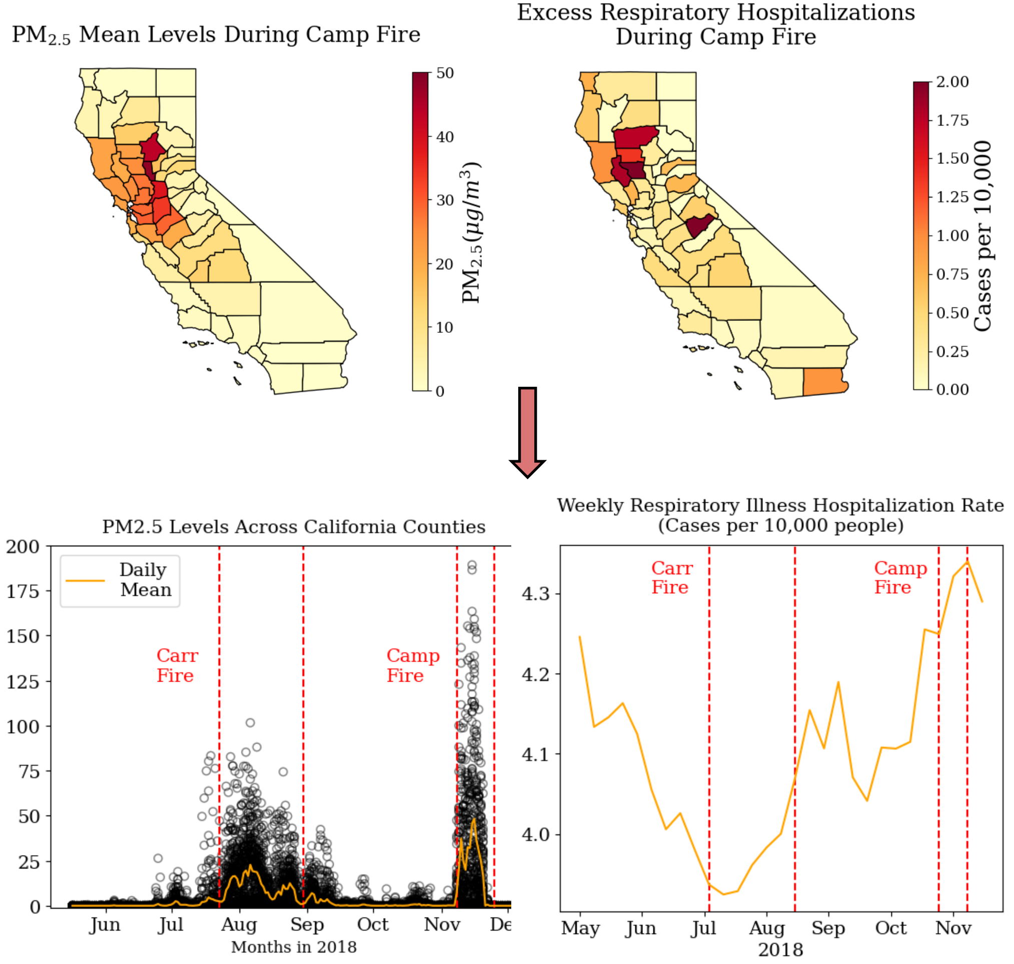
- Smoke ↑ and respiratory hospitalizations ↑

What would respiratory outcomes have looked like under a different smoke trajectory?

Why this is challenging:

- Smoke exposure varies across space
- Health effects unfold over time
- Nearby regions affect one another
- Exposure and health are shaped by evolving weather and prior conditions

California 2018 wildfire season



A Spatiotemporal Causal Model

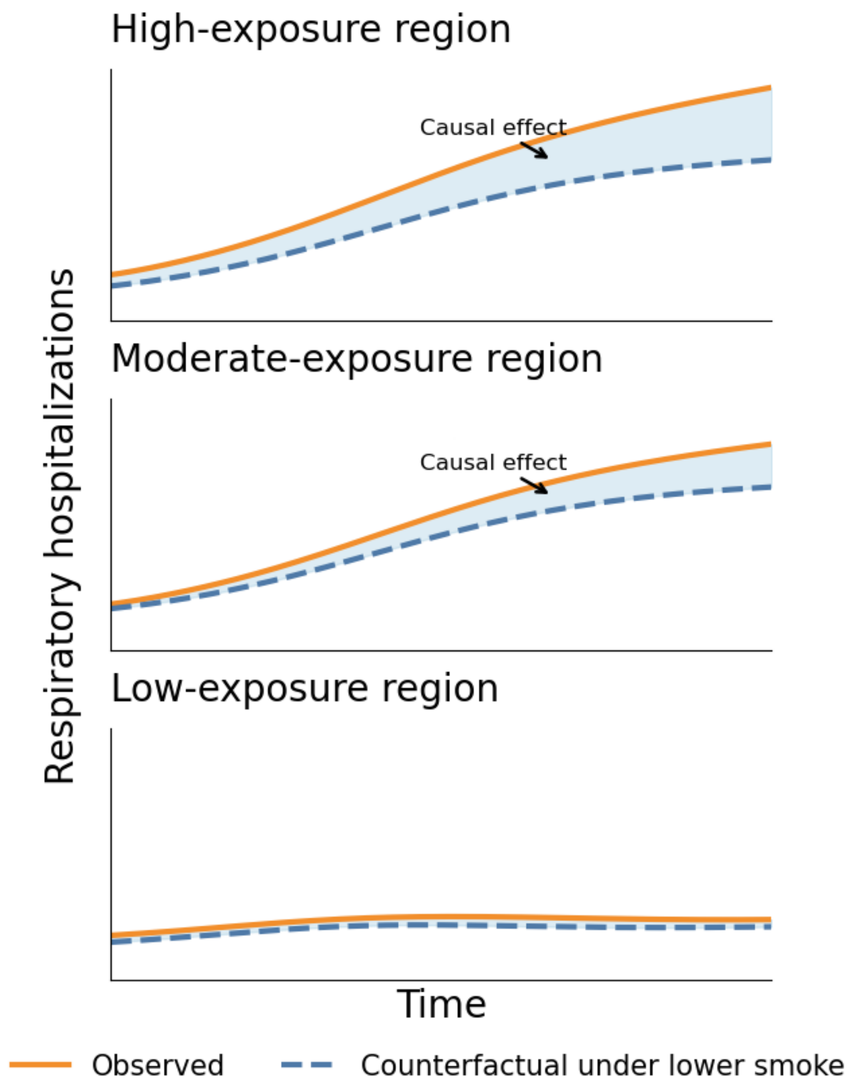
Notation

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

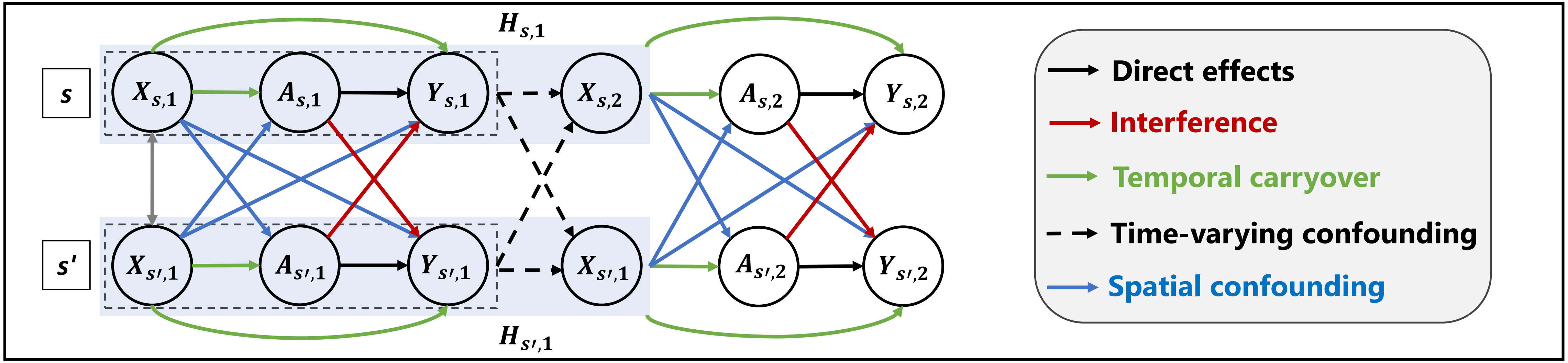
Counterfactual target

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

- Average potential outcome after τ time steps under a series of fixed interventions, $\boldsymbol{a}_{t:t+\tau-1}$, given history $\boldsymbol{h}_{1:t}$.
- "Starting from the observed history, what would have happened under a different smoke path?"



Challenges in Spatiotemporal Causal Inference

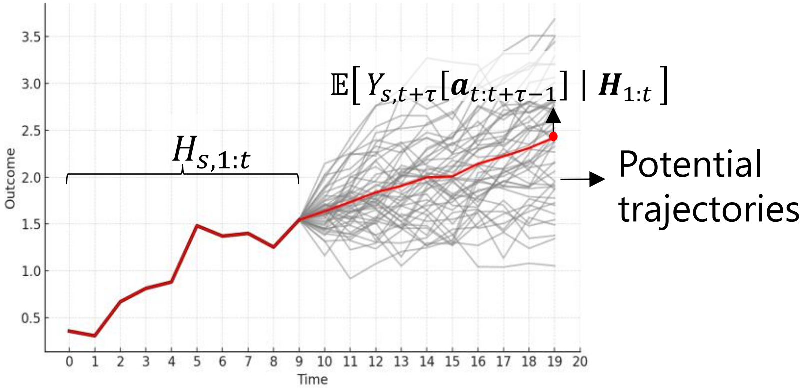


1. **Interference**
2. **Temporal carryover**
3. **Time-varying confounders**
   - Past conditions affect both future exposure and future outcomes, creating feedback loops (e.g., $\text{weather}_t \rightarrow \text{PM}_{2.5,\ t} \rightarrow \text{health}_{t+1}$, $\text{weather}_{t+1} \rightarrow \text{PM}_{2.5,\ t+1}$ etc.)
4. **Single observed trajectory (not many i.i.d. trajectories)**

Estimating Counterfactual Trajectories

Time-varying confounders

- Iterative G-Computation handles this by **recursively integrating** over future counterfactual histories



Single spatiotemporal trajectory

- We learn a shared representation across time and use it at each recursive step, such that:

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

- 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.

Last step: predict terminal outcome

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

Recursive step *k*: replace future outcomes with their counterfactuals

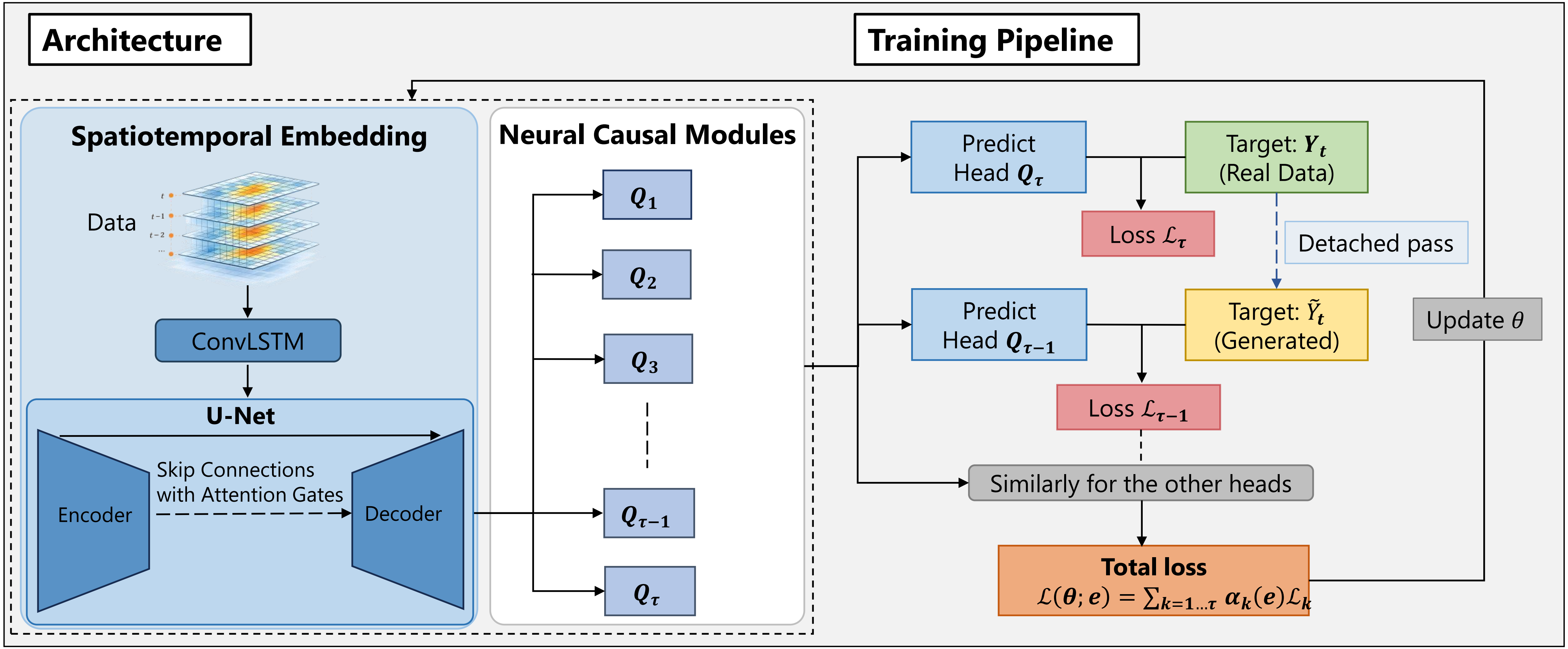
$$Q_k(\boldsymbol{H}_{1:t+k-1}, \boldsymbol{A}_{t+k-1}) = \mathbb{E}[Q_{k+1}(\boldsymbol{H}^{\boldsymbol{a}}_{1:t+k}, \boldsymbol{A}_{t+k}) \mid \phi(\boldsymbol{H}_{1:t+k-1}, \boldsymbol{A}_{t+k-1})]$$

Result: obtain the counterfactual outcome at the current history

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

GST-UNet: Neural G-Computation for Spatiotemporal Data

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



Simulation Results on Synthetic Data

Wildfire smoke is simulated using an advection–diffusion process

Table 1: RMSE ± SD across test trajectories. Bold indicates lowest error per column; color shows improvement (RMSE **decrease** or **increase**) over best baseline (excluding ablations).

τ	Model	$\beta_1 = 0.0$	$\beta_1 = 0.5$	$\beta_1 = 1.0$	$\beta_1 = 1.5$	$\beta_1 = 2.0$
5	UNet+	**0.28 ± 0.00**	0.36 ± 0.00	0.54 ± 0.01	0.71 ± 0.01	0.81 ± 0.01
	STCINet	0.29 ± 0.00	0.38 ± 0.01	0.62 ± 0.01	0.80 ± 0.01	0.90 ± 0.01
	IPWUNet	0.60 ± 0.01	0.58 ± 0.01	0.58 ± 0.01	0.59 ± 0.01	0.59 ± 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
	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
	GST-UNet	0.33 ± 0.00 (**+17.9%**)	**0.35 ± 0.00** (**-2.7%**)	**0.40 ± 0.00** (**-21.6%**)	**0.44 ± 0.00** (**-25.4%**)	**0.40 ± 0.01** (**-32.2%**)
10	UNet+	**0.28 ± 0.00**	0.61 ± 0.00	1.18 ± 0.00	1.45 ± 0.00	1.71 ± 0.01
	STCINet	0.31 ± 0.00	0.68 ± 0.00	1.25 ± 0.00	1.47 ± 0.01	1.60 ± 0.01
	IPWUNet	0.78 ± 0.01	0.80 ± 0.01	0.96 ± 0.01	1.19 ± 0.02	1.08 ± 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
	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
	GST-UNet	0.38 ± 0.00 (**+35.7%**)	**0.55 ± 0.00** (**-9.8%**)	**0.68 ± 0.00** (**-29.2%**)	**0.73 ± 0.01** (**-38.7%**)	**0.85 ± 0.01** (**-21.3%**)

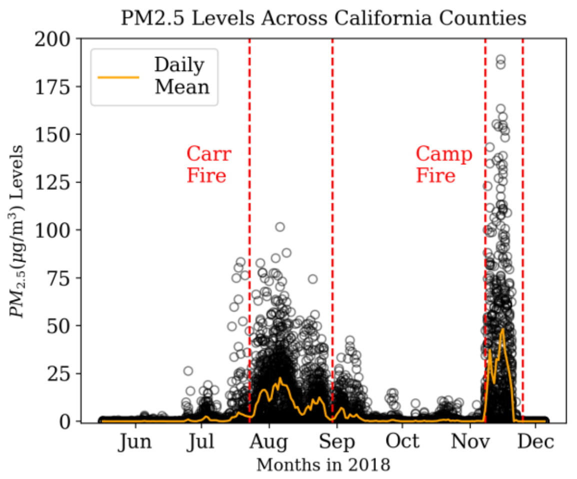
Case Study: Effect of Wildfire Smoke on Respiratory Illness during the 2018 California Camp Fire

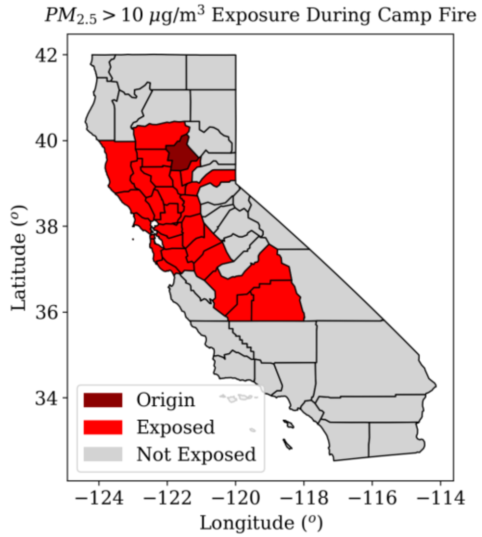
Data (2018 California, county-level data):

- **Covariates:** demographics, wind, temperature, precipitation, humidity, shortwave radiation
- **Intervention:** $PM_{2.5} > 10\ \mu g/m^3$ (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

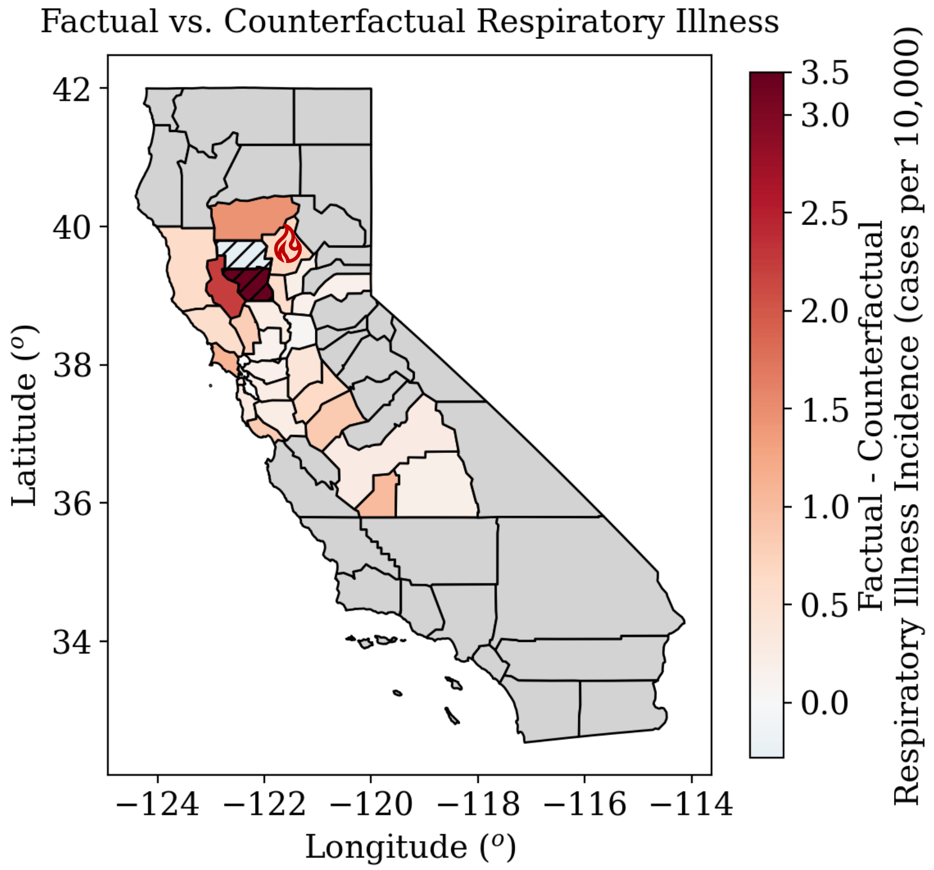
Results

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).

Today's Talk

1. Spatiotemporal Causal Inference
   - *Wildfire smoke and health effects*

2. Adaptive Experimentation Under Noncompliance
   - *Learning with imperfect control*

3. Future Directions
   - *Biomedical causal modeling, reliable evaluation, and decision-making under constraints*

Example: Adaptive Designs for Clinical Trials

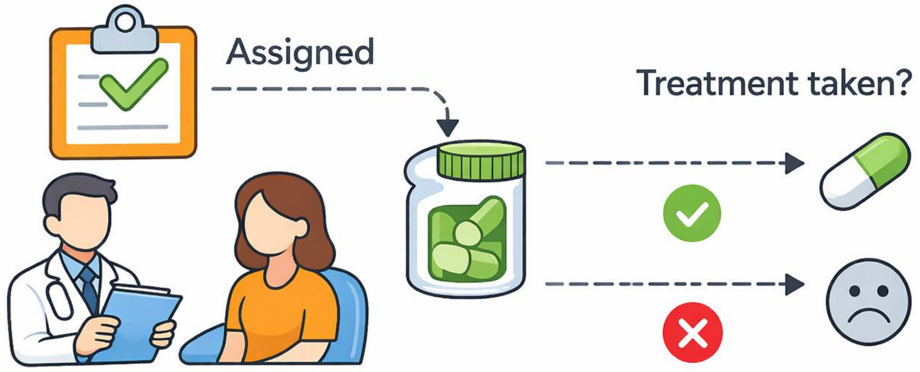
Clinical Trial

- Randomized trials are the gold standard
  for estimating treatment effects.
- Adaptive designs can improve efficiency
  Updating the trial using interim data might reduce expected sample size or trial duration.
- But assignment does not guarantee uptake
  Patients might not actually take the assigned treatment.

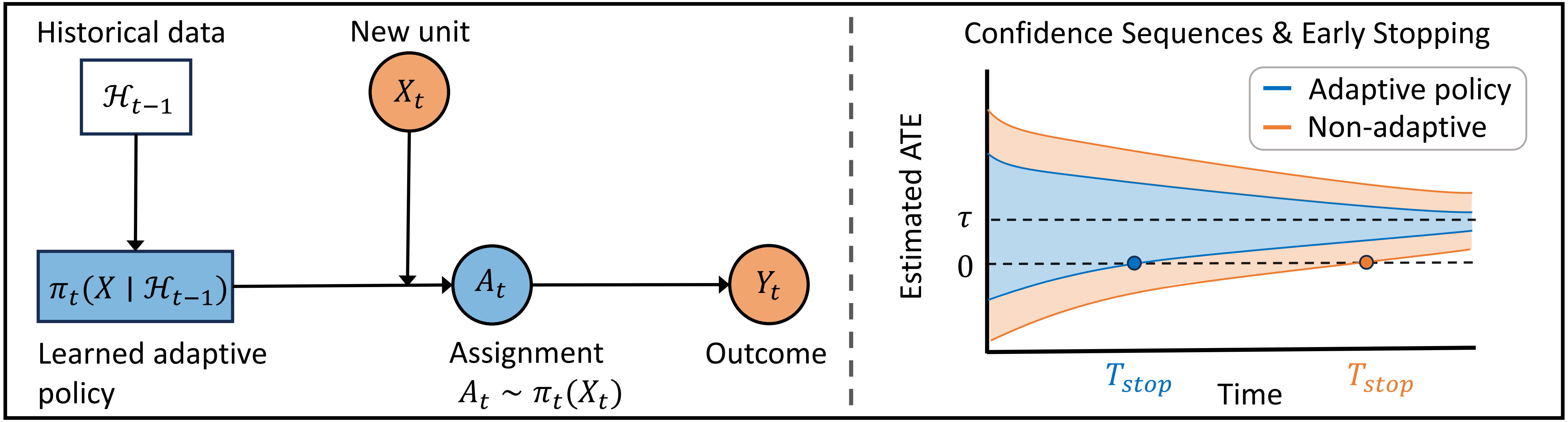
How should we design experiments when treatment can be randomized, but not fully enforced?

"An **adaptive design** may provide the same statistical power with a smaller expected sample size or shorter expected duration than a comparable non-adaptive design."

FDA, Adaptive Designs for Clinical Trials of Drugs and Biologics (2019)



Adaptive Design with Perfect Compliance

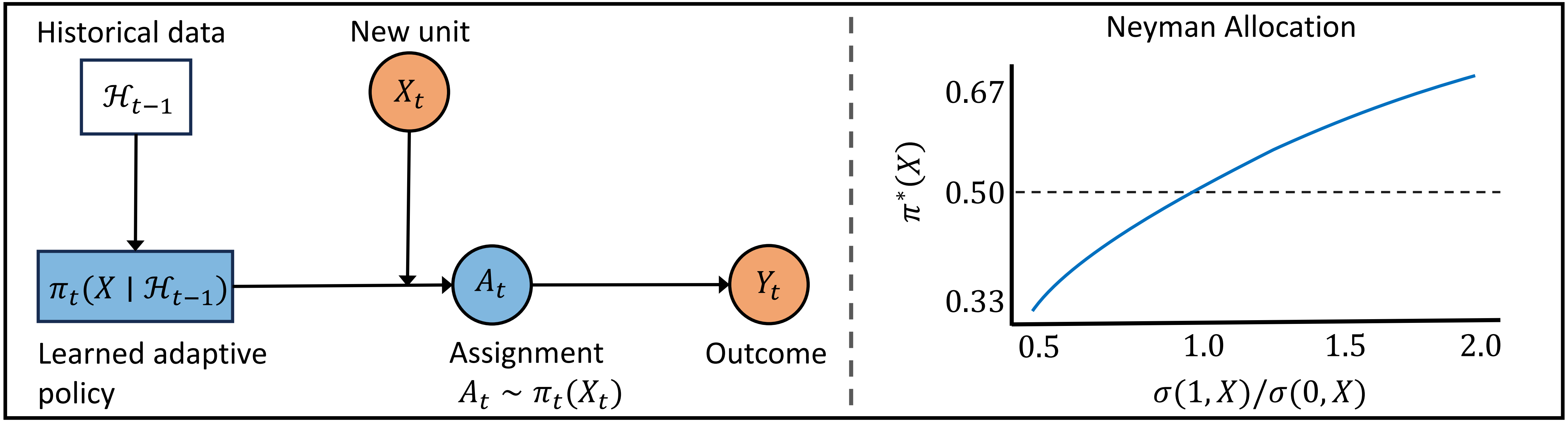


- **Setting:** Observe X_t , assign binary $A_t \sim \pi_t(X \mid \mathcal{H}_{t-1})$, get Y_t.
- **Goal:** Learn an adaptive policy $\pi_t(X \mid \mathcal{H}_{t-1})$ at time t that minimize the variance of the average treatment effect (ATE):

$$ATE = \mathbb{E}[Y \mid A = 1] - \mathbb{E}[Y \mid A = 0]$$

- **Motivation:** *Early stopping* by driving faster uncertainty reduction.

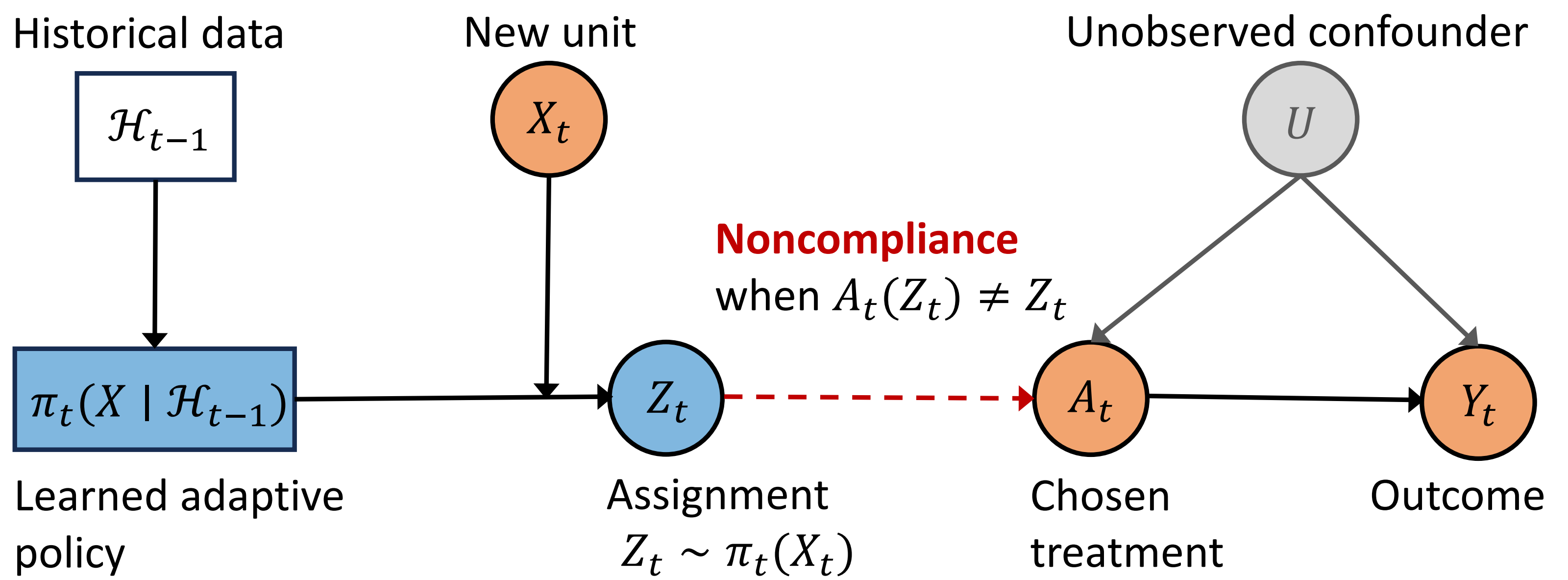
Adaptive Design with Perfect Compliance



- **Classical Result:** *Neyman allocation* — assign more where outcomes are noisier.

$$\pi^*(X) = \frac{\sqrt{\mathrm{Var}(Y \mid A = 1, X)}}{\sqrt{\mathrm{Var}(Y \mid A = 0, X)} + \sqrt{\mathrm{Var}(Y \mid A = 1, X)}} := \frac{\sigma(1, X)}{\sigma(0, X) + \sigma(1, X)}$$

Adaptive Design with Noncompliance



Motivating applications

- **Noncompliance:** Recommendation Z_t is randomized, treatment A_t is not.
- **Issue:** Naive A/B on A_t is invalid.
- **Fix:** Use Z_t as an *instrument* to identify the ATE and adapt its policy instead.

Optimal Adaptive Policy with Noncompliance

Optimal adaptive policy must balance:

- Outcome uncertainty (like Neyman)
- Compliance, i.e. how strongly the instrument shifts treatment:

$$\delta^A(x) = \mathbb{E}[\,A \mid X = x, Z = 1\,] - \mathbb{E}[\,A \mid X = x, Z = 0\,]$$

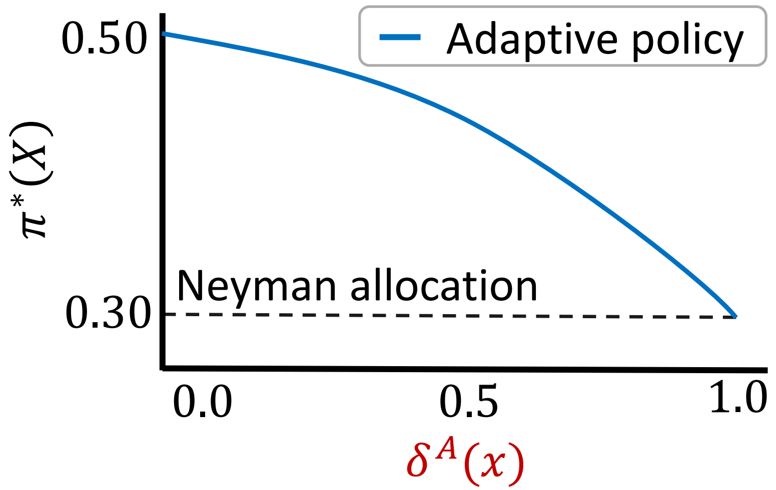
Fixed policy that minimizes asymptotic variance*:

$$\boldsymbol{\pi}^*(\boldsymbol{X}) = \frac{\sqrt{\mathrm{Var}(Y - A\delta(X) \mid Z = 1, X)}}{\sqrt{\mathrm{Var}(Y - A\delta(X) \mid Z = 0, X)} + \sqrt{\mathrm{Var}(Y - A\delta(X) \mid Z = 1, X)}}$$

where:

$$\delta(X) = \frac{\delta^Y(X)}{\delta^A(X)} = \frac{\mathbb{E}[\,Y \mid X = x, Z = 1\,] - \mathbb{E}[\,Y \mid X = x, Z = 0\,]}{\underbrace{\mathbb{E}[\,A \mid X = x, Z = 1\,] - \mathbb{E}[\,A \mid X = x, Z = 0\,]}_{\delta^A(x)\ \text{(compliance factor)}}}$$

Adaptive policy vs compliance



*ATE Identification from Wang & Tchetgen Tchetgen (2018): $\mathrm{ATE} = \mathbb{E}[\delta(x)]$.

- Under IV relevance, exclusion, randomization given X and unconfounded compliance

AMRIV: Adaptive + Robust Estimation

AMRIV = **A**daptive **M**ultiply-**R**obust estimator for **IV** settings

Burn-in period

For $\boldsymbol{t = 1 \dots T_0}$:

- Collect initial randomized data with a fixed policy

E.g.: $\pi_0(X) = 1/2$

- Build initial history

$$H_{T_0} = \{(X_i, Z_i, A_i, Y_i)\}_{i=1}^{T_0}$$

Adaptive collection

For $\boldsymbol{t = T_0 \dots T}$:

Estimate the optimal policy $\boldsymbol{\pi^*}$

- Estimate $\mathrm{Var}(Y - A\delta(X) \mid Z = z, X)$ as $\hat{\sigma}(z, X)$ from H_{t-1}.

$$\tilde{\pi}_t(X \mid H_{t-1}) = \frac{\hat{\sigma}(1, X)}{\hat{\sigma}(0, X) + \hat{\sigma}(1, X)}$$

- Truncate: $\boldsymbol{\pi_t}(\boldsymbol{X} \mid \boldsymbol{H_{t-1}}) := \min(1 - \epsilon_t, \max(\epsilon_t, \tilde{\pi}_t))$

Collect data and impute treatment effect

- See X_t, set $Z_t \sim \pi_t(X_t)$, observe A_t, Y_t.
- Estimate nuisances $\hat{\eta}_t$ from H_{t-1} via cross-fitting.
- Estimate per-round effect as

$$\boldsymbol{\phi_t} = \phi(X_t, Z_t, A_t, Y_t, ; \pi_t, \hat{\eta}_t)$$

where ϕ is the **efficient influence function (EIF)** of τ.

Estimation

At $\boldsymbol{t = T}$:

- Estimate effect as:

$$\hat{\tau}_T^{AMRIV} = \frac{1}{T} \sum_{t=1}^{T} \phi_t$$

- Build confidence sequences:

$$[L_t, U_t]$$

- Decide whether to continue or stop data collection

Theoretical Guarantees

- Efficient: smallest possible variance.

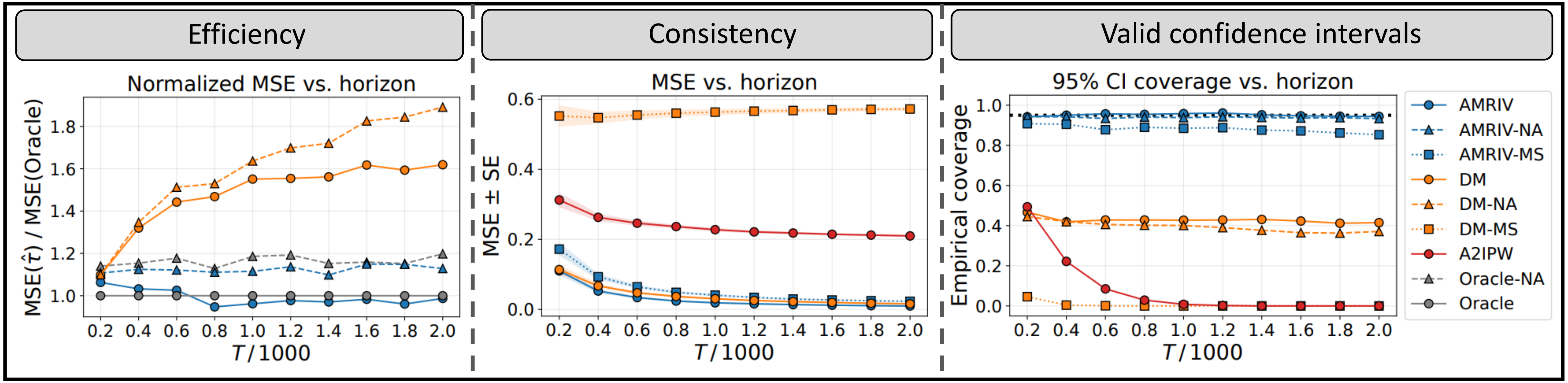
$$\sqrt{T}\left(\hat{\tau}_T^{AMRIV} - \tau\right) \to \mathcal{N}\left(0, V_{eff}(\pi)\right)$$

  with $\pi = \pi^*$ achieving the minimum bound.

- Robust: converges if some models are wrong.
  - Consistent if either $\delta(X)$ or $\delta_A(X)$ is learned consistently; AMRIV is $O_p(T^{-1/2})$ if both $\delta(X)$ and $\delta_A(X)$ are $o_p(T^{-1/4})$.
- Anytime-valid: safe early stopping.
  - Can build anytime valid asymptotic confidence sequences which enables peek-safe early stopping.

TL;DR: We get the tightest peek-safe valid confidence intervals under adaptivity and noncompliance.

Experimental Results



- **Efficiency:** Adaptivity improves efficiency of all estimators.
- **Robustness:** AMRIV-MS is consistent even when one of the nuisances is misspecified, whereas the direct method DM-MS is not.
- **Validity:** AMRIV achieves nominal (95%) coverage unlike non-robust methods.

Today's Talk

1. Spatiotemporal Causal Inference
   - *Wildfire smoke and health effects*

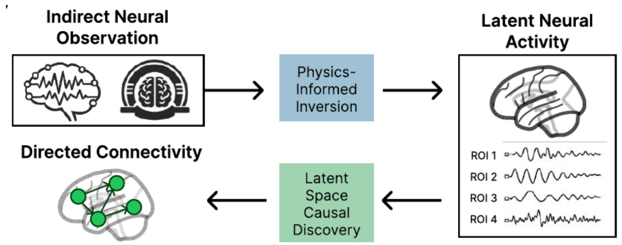
2. Adaptive Experimentation Under Noncompliance
   - *Learning with imperfect control*

3. Future Directions
   - *Biomedical causal modeling, reliable evaluation, and decision-making under constraints*

Future Directions

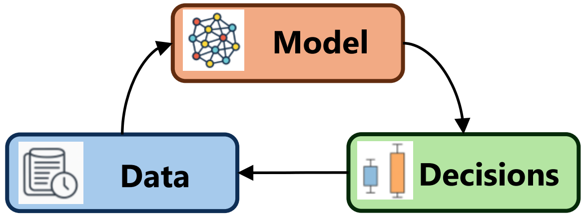
Looking ahead, I am interested in the next generation of reliable causal learning problems.

Scientific / Biomedical Causal Modeling



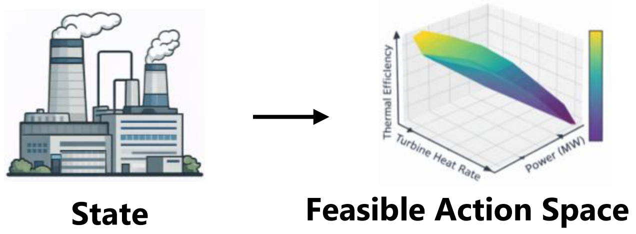
- Causal discovery from indirect observations
- Structured, high-dimensional scientific systems
- Physics-informed + causal reasoning

Reliable Evaluation for Adaptive / AI Systems



- Feedback loops: model decisions change data
- What is identifiable under adaptive data collection?
- Valid uncertainty under evolving policies

Causal Inference with Physical and Operational Constraints



- Continuous actions under physical constraints
- Feasible set defined by system dynamics
- Causal policy learning with constraint structure

Unifying theme: Reliable causal inference in complex, real-world systems.

Q&A

- **Special thanks to my committee:**
  Nathan Kallus (Chair), Sarah Dean, Peter Frazier, Emma Pierson.
- **Collaborators on this work (in alphabetical order):**
  Brian M Cho, Nathan Kallus, Xihaier Luo, David K Park, Shinjae Yoo.

- **Funding support:**
  The Department of Energy Computational Science Graduate Fellowship.

Appendix

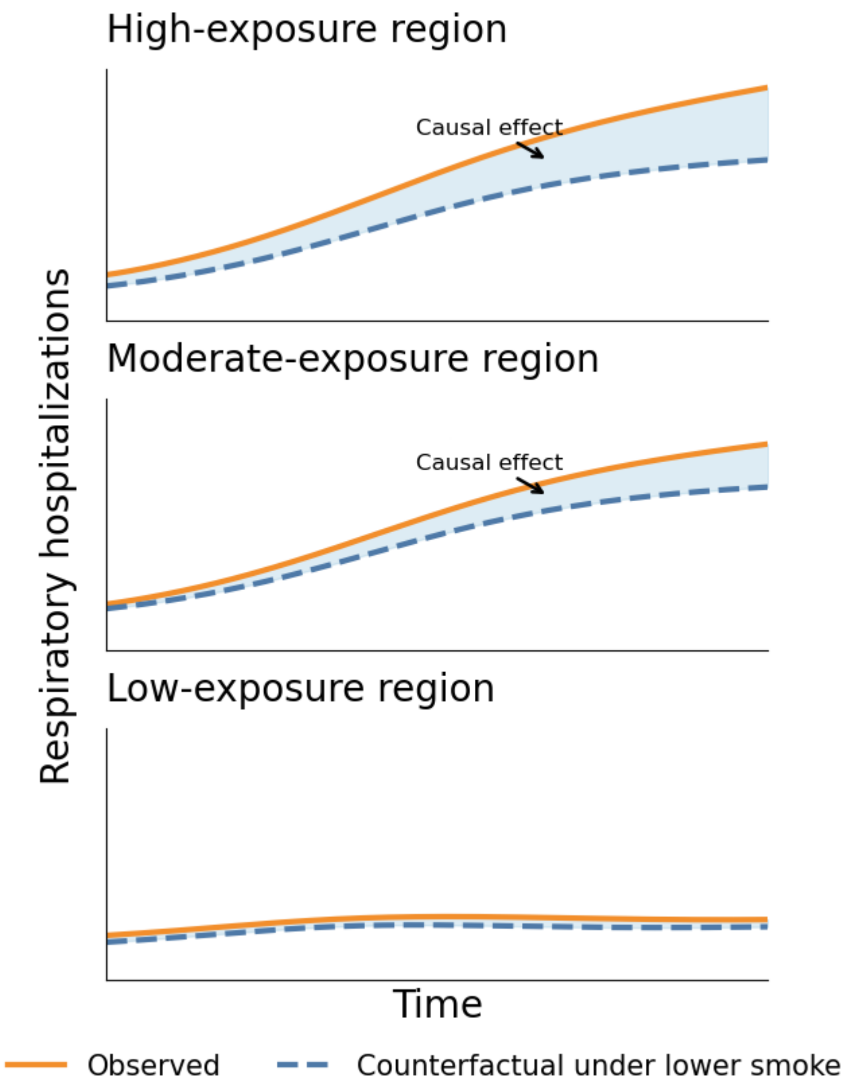
What Do We Want to Estimate?

How would health outcomes change under a different smoke exposure history?

- Not just forecasting future hospitalizations
- Estimate the effect of a potential intervention path across space and time
- Compare observed outcomes to a counterfactual trajectory

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

  - Average potential outcome after τ time steps under a series of fixed interventions, $\boldsymbol{a}_{t:t+\tau-1}$, given history $\boldsymbol{h}_{1:t}$.



Using a Single Spatiotemporal Chain

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 $t \in \{1, \dots, T - \tau\}$, define a "prefix"

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

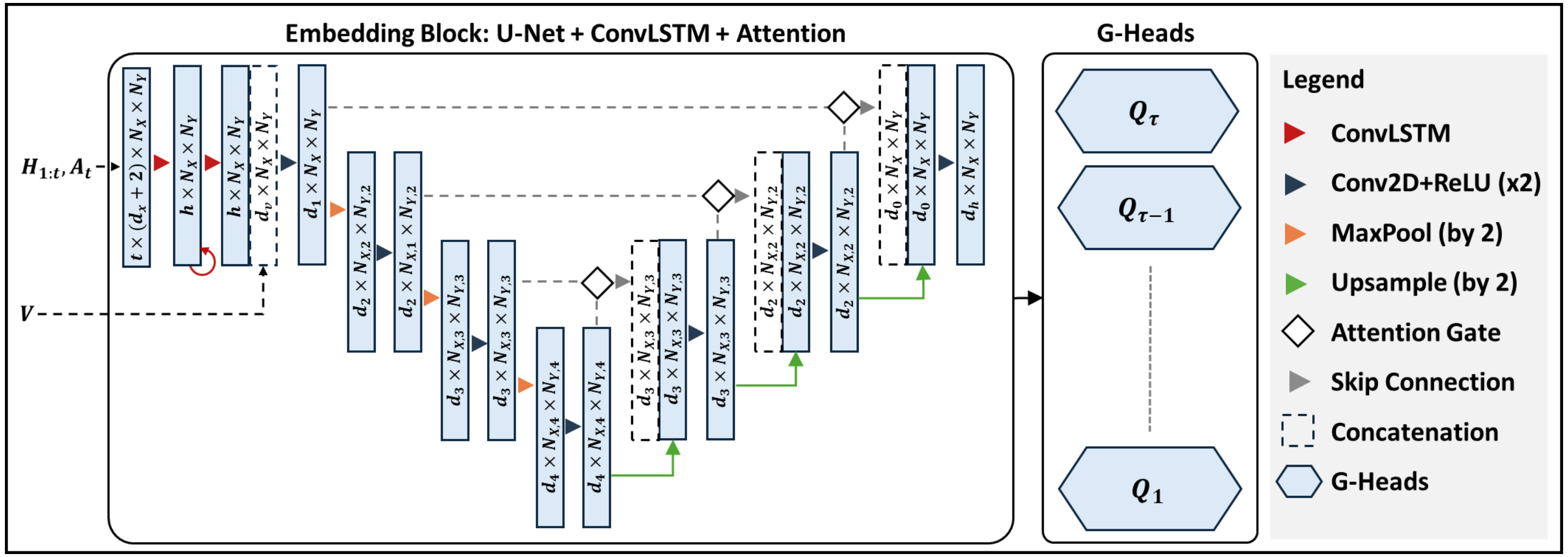
- Under representation-based time invariance, conditioning on $\phi(\boldsymbol{H}_{1:t}, \boldsymbol{A}_t)$ renders the distribution of $\boldsymbol{Y}_{t+\tau}$ independent of t.
- We can then write expectations over these prefixes as

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

GST-UNet: Neural G-Computation for Spatiotemporal Data

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

- **Spatiotemporal Embedding ϕ:** U-Net + ConvLSTM + attention gates.
- **Neural Causal Modules Q_k:** 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

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

Curriculum Options:

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

  Issue: the later heads $(1, 2, \ldots)$ train on noise while the earlier heads $(\tau, \tau - 1, \ldots)$ 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.

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

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 $e = 1 \ldots E$ **do**
  for $k = \tau \ldots 1$ **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

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

  (**Backward pass**) Update θ by backpropagation.
end for
(**Inference**) Given a $\mathbf{h}_{1:t}$, return $Q_1(\phi(\mathbf{h}_{1:t}, \mathbf{a}_t); \widehat{\theta})$.