ICML
International Conference
On Machine Learning

spatiotemporal
Machine Learning

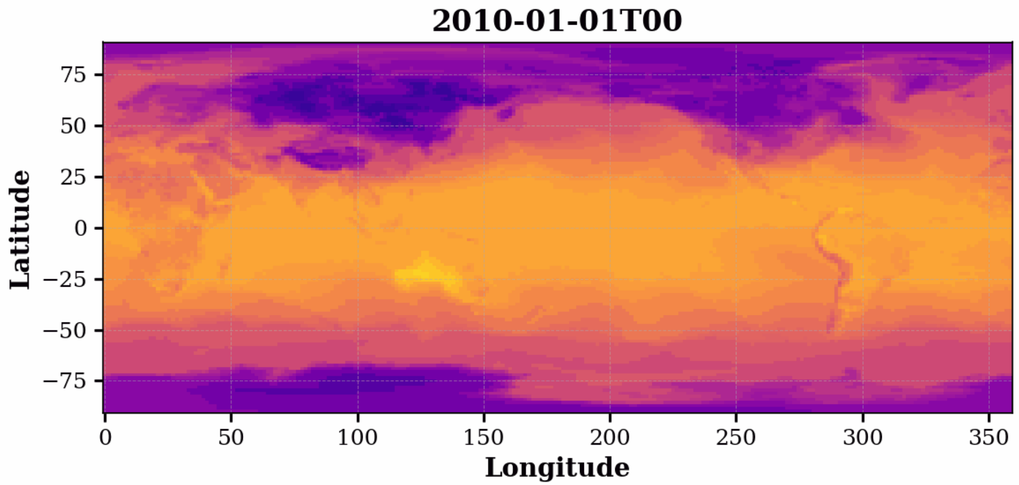
Discovering Latent Causal Graphs from Spatiotemporal Data

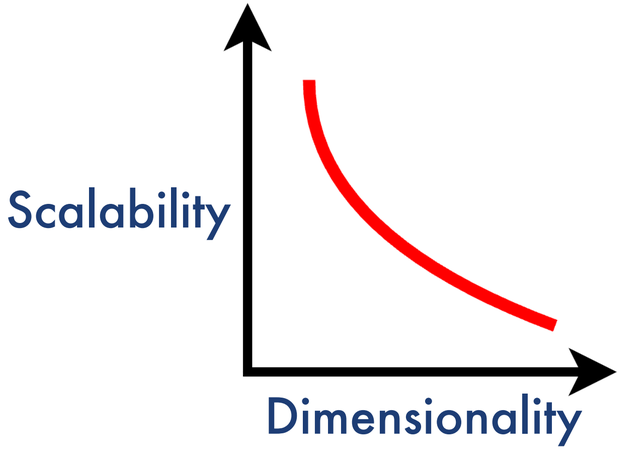
Kun Wang*, Sumanth Varambally*, Duncan Watson-Parris, Yian Ma, Rose Yu

* equal contribution

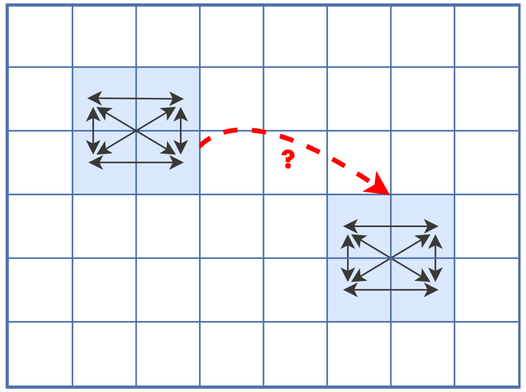
UC San Diego

Motivation





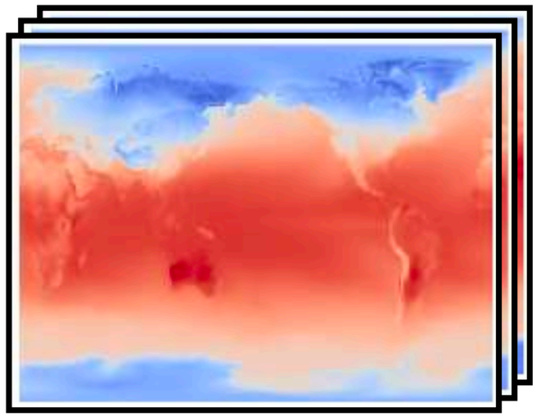
High dimensionality limits scalability



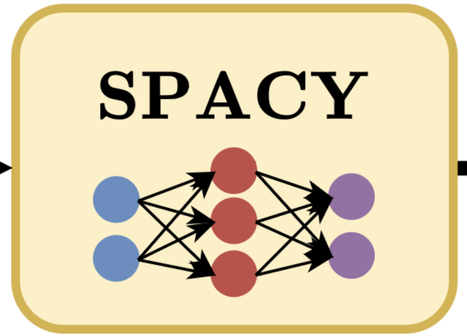
Strong correlations between nearby locations can obscure real causal links

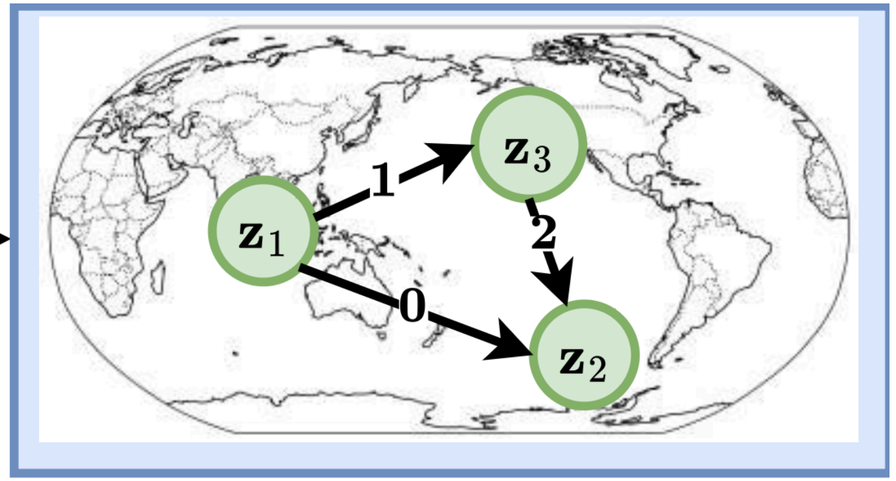
ICML
International Conference
On Machine Learning

Overview



Spatiotemporal Data





Inferred Spatial Modes
and Latent Causal Graph

spatiotemporal
Machine Learning

ICML
International Conference
On Machine Learning

SPACY

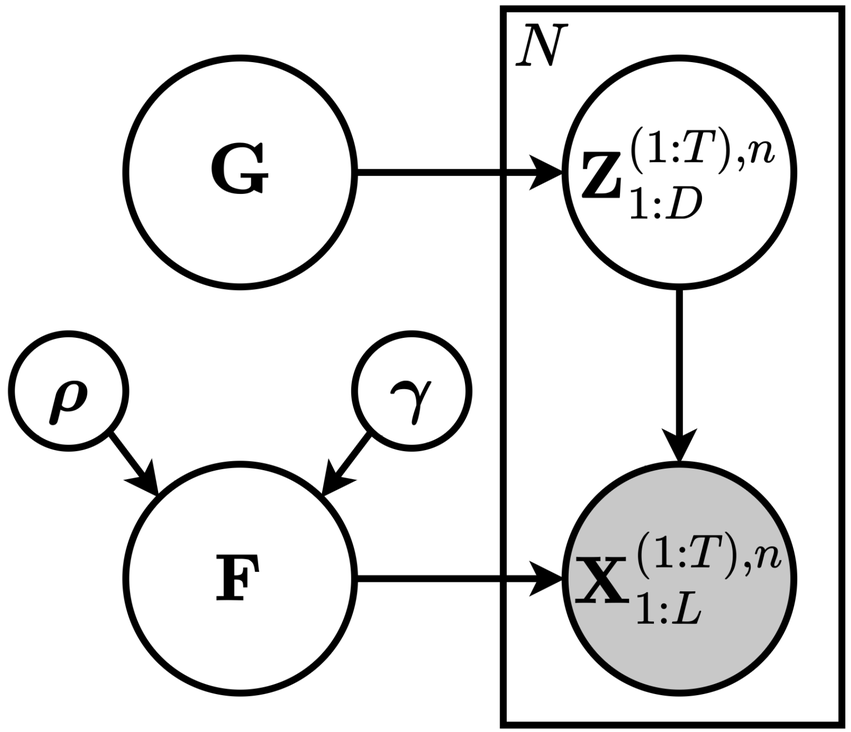
spatiotemporal
Machine Learning

Observed spatiotemporal data $\mathbf{X}_{1:L}^{(1:T),n}$ with L grid-points

Latent time series $\mathbf{Z}_{1:D}^{(1:T),n}$

The causal graph $\mathbf{G}$

Spatial Factors $\mathbf{F}$



$$\mathbf{X}_{\ell}^{(t)} = g_{\ell}\left([\mathbf{FZ}]_{\ell}\right) + \varepsilon_{\ell}^{(t)}$$

ICML
International Conference
On Machine Learning

SPACY

spatiotemporal Machine Learning

Spatial Factors $\mathbf{F}$ implemented as RBF kernels:

$$\mathbf{F}_{\ell d} = \mathrm{RBF}_d\left(x_\ell; \rho_d, \gamma_d\right) = \exp\left(-\frac{\|x_\ell - \rho_d\|^2}{\exp\left(\gamma_d\right)}\right)$$

Evidence Lower Bound (ELBO):

Conditional Likelihood Latent-causal conformity

$$\log p_\theta\left(\mathbf{X}^{(1:T),1:N}\right) \geq \sum_{n=1}^{N}\left\{\mathbb{E}_{q_\phi(\mathbf{Z}^{(1:T),n}|\mathbf{X}^{(1:T),n})q_\phi(\mathbf{G})q_\phi(\mathbf{F})}\left[\log p_\theta\left(\mathbf{X}^{(1:T),n}\,|\,\mathbf{Z}^{(1:T),n},\mathbf{F}\right) + \left[\log p_\theta\left(\mathbf{Z}^{(1:T),n}\,|\,\mathbf{G}\right) - \log q_\phi\left(\mathbf{Z}^{(1:T),n}\,|\,\mathbf{X}^{(1:T),n}\right)\right]\right]\right\}$$

$$-\mathrm{KL}\left(q_\phi(\mathbf{G})\,||\,p(\mathbf{G})\right) - \mathrm{KL}\left(q_\phi(\mathbf{F})\,||\,p(\mathbf{F})\right) = \mathrm{ELBO}(\theta,\phi)$$

ICML
International Conference
On Machine Learning

Identifiability

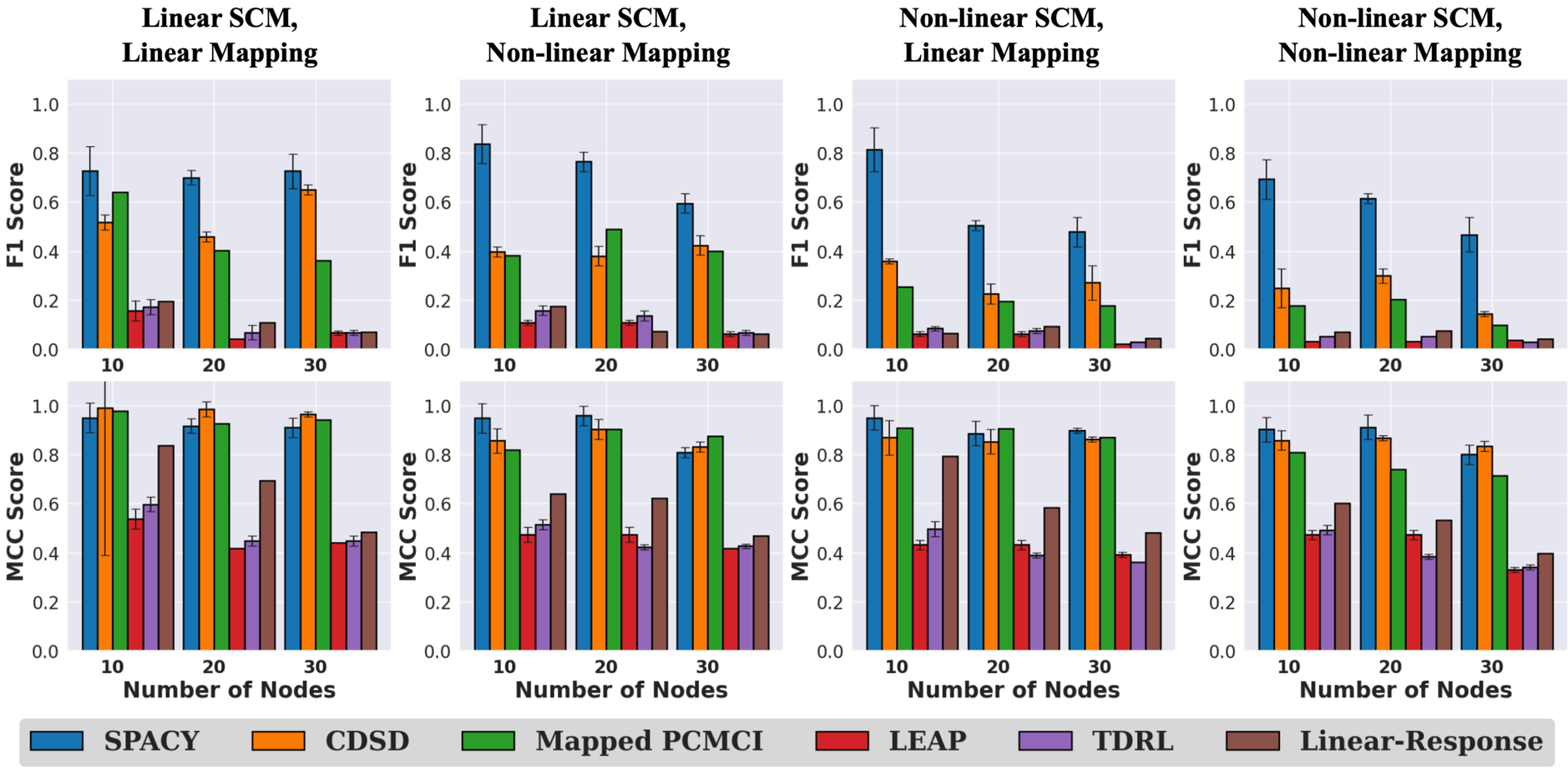
spatiotemporal
Machine Learning

Theorem (informal version). *Suppose we have two spatiotemporal processes* $\mathbf{X}^{(t)}(\ell), \widehat{\mathbf{X}^{(t)}}(\ell)$ *in the infinite resolution grid* $\mathcal{G} = [0,1]^K$ *described by the equation* $\mathbf{X}^{(t)}(\ell) = g_\ell\left(\mathbf{F}_\ell^\top \mathbf{Z}^{(t)}\right) + \varepsilon_\ell^{(t)}$ *for* $\ell \in \mathcal{G}$. *For certain smoothness conditions on* g_ℓ, *if* $p\left(\mathbf{X}^{(t)}(\ell) \mid \mathbf{Z}^{(t)}, \mathbf{F}_\ell\right) = p\left(\widehat{\mathbf{X}}^{(t)}(\ell) \mid \widehat{\mathbf{Z}}^{(t)}, \widehat{\mathbf{F}}_\ell\right), \forall \ell \in \mathcal{G}, t \in [T]$, *then the latent process* $\mathbf{Z}$ *is identifiable up to permutation and scaling.*

ICML
International Conference
On Machine Learning

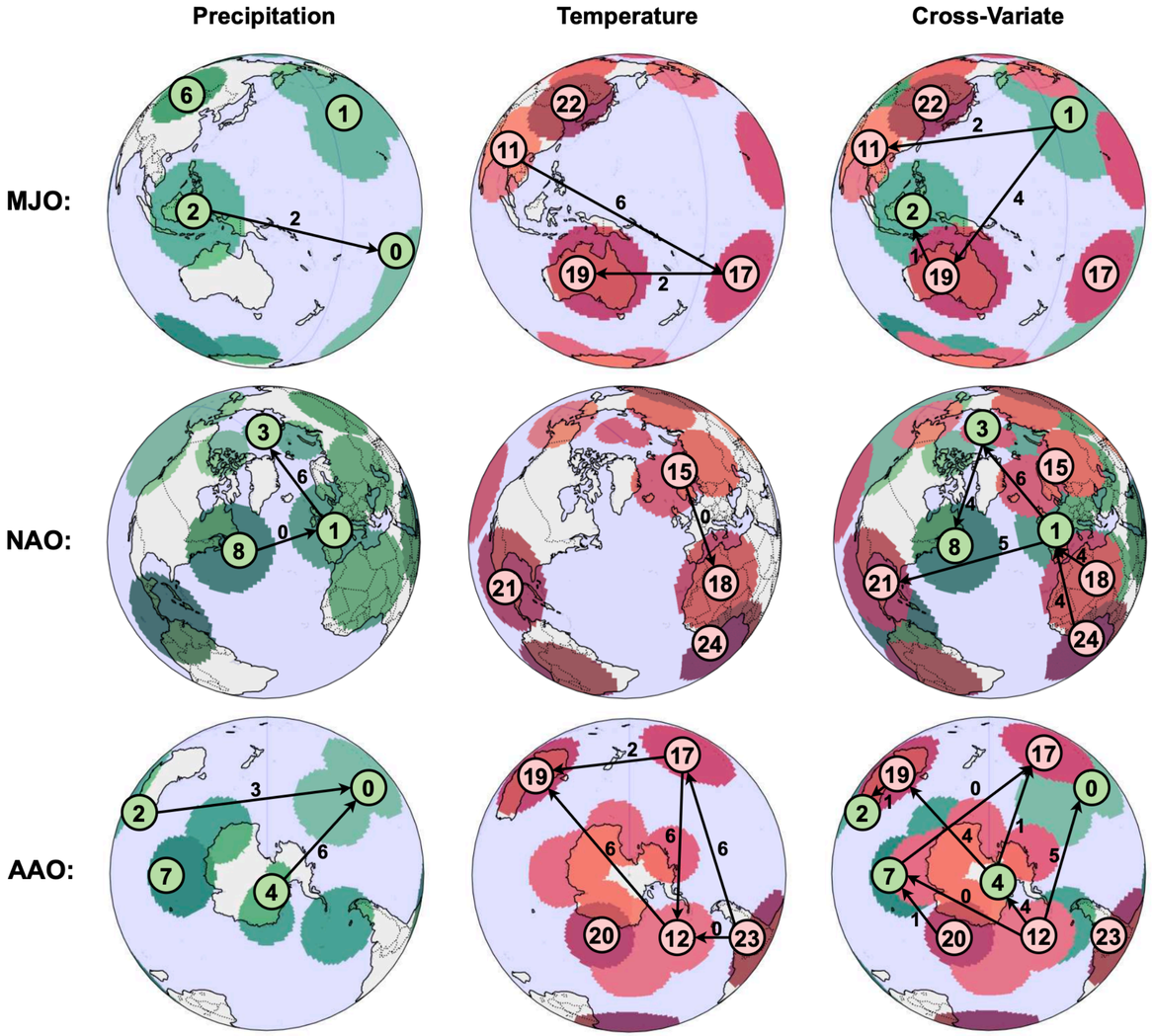
Results - Synthetic Data

spatiotemporal
Machine Learning



ICML
International Conference
On Machine Learning

Results - Climate Data



spatiotemporal
Machine Learning

ICML
International Conference
On Machine Learning

Thank You!

E-mail: svarambally@ucsd.edu

spatiotemporal
Machine Learning