



Quantifying Time Series Periodicity with Interpretable Machine Learning

Climate Systems & Urban Human Mobility

Xinyu Chen

Postdoctoral Associate, MIT

October 28, 2025

Cambridge, USA

Spatiotemporal Time Series Data

- Transport & mobility & climate application scenarios



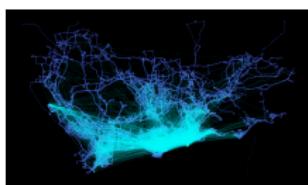
Highway (Portland)



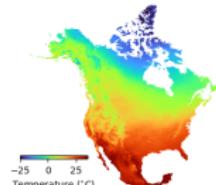
Uber movement (NYC)



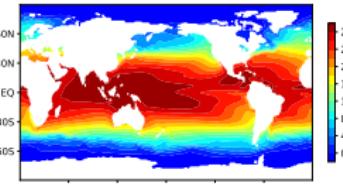
Uber movement (Seattle)



Taxi trajectory (Shenzhen)



Temperature (NA)

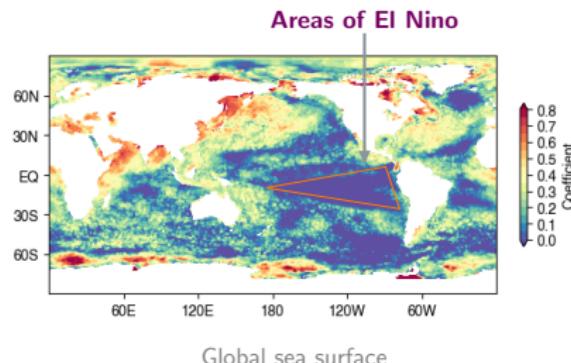
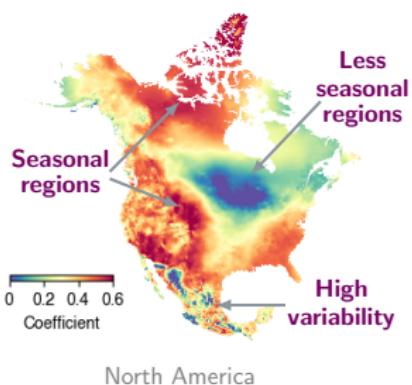


Temperature (sea surface)

- Challenges: Sparsity, high-dimensionality, multi-dimensionality, heavy tails, irregular sampling, and time-varying systems

Motivation

Yearly temperature seasonality patterns in 2010s

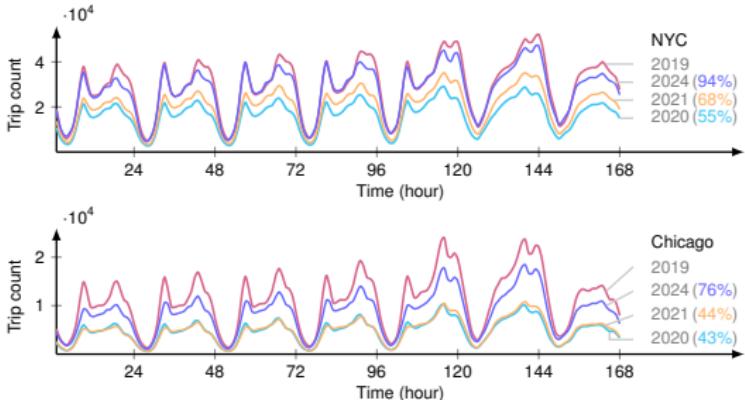
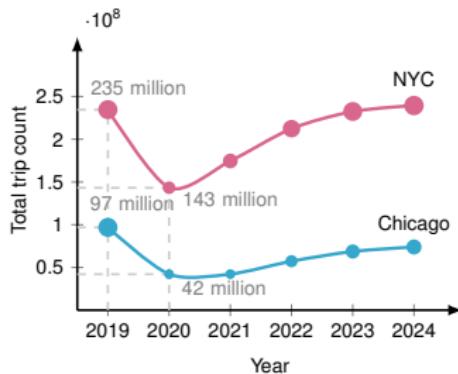


What motivates us most about periodicity?

- ① **Monitoring climate systems:** Empirically measure the periodicity of climate variables (e.g., temperature, precipitation).
- ② **Discovering spatiotemporal patterns:** Identify periodicity pattern shift and special climate events.

Motivation

Ridesharing trip data



What motivate us most about periodicity?

- ① Resilience and stability of systems:** Empirically measure the periodicity and predictability of urban systems.
- ② Optimization of transport systems:** Optimize resources (e.g., public transit, taxi, ridesharing, micromobility) to meet transport demand efficiently.
- ③ Design of sustainable transport & infrastructure:** Implement energy-efficient solutions (e.g., congestion pricing) tailored to peak hours.



Interpretable Time Series Autoregression



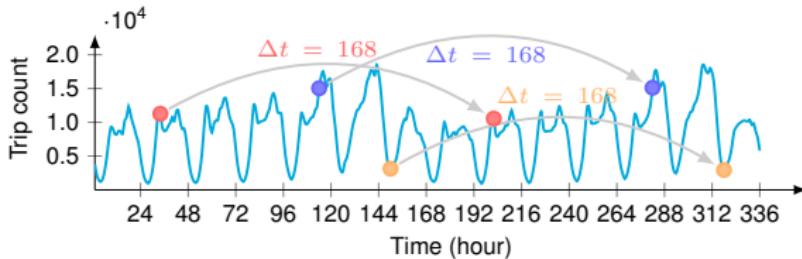
<https://github.com/xinychen/integers>

- Interpretable ML
- ℓ_0 -norm optimization
- Climate system seasonality
- Sparse autoregression
- Mixed-integer programming
- Human mobility regularity

Valorizing Autoregression

- Time series autoregression on $\mathbf{x} \in \mathbb{R}^T$ with order $d \in \mathbb{Z}^+$

$$\mathbf{w} := \arg \min_{\mathbf{w}} \sum_{t=d+1}^T \left(\mathbf{x}_t - \sum_{k=1}^d w_k \mathbf{x}_{t-k} \right)^2$$



Periodicity of hourly rideshare trip time series

- Sparse** coefficient vector \mapsto **Interpretability?**

$$\underbrace{\mathbf{w}}_{\text{sparsity } \|\mathbf{w}\|_0 \triangleq 3} = (\underbrace{0.33}_{k=1}, 0, \dots, 0, \underbrace{0.20}_{k=167}, \underbrace{0.46}_{k=168})^\top \in \mathbb{R}^{168}$$

Sparse Autoregression

- Identify the dominant auto-correlations
 - $\tau \in \mathbb{Z}^+$: Upper bound of the number of nonzero entries in $w \in \mathbb{R}^d$

$$\tilde{x} \approx A \begin{pmatrix} w_1 \\ w_2 \\ w_3 \\ w_4 \end{pmatrix} \quad \left\| w \right\|_0 \leq \tau$$

$$w := \arg \min_{\|w\|_0 \leq \tau} \sum_{t=d+1}^T \left(x_t - \sum_{k=1}^d w_k x_{t-k} \right)^2$$

$$= \arg \min_{\|w\|_0 \leq \tau} \|\tilde{x} - Aw\|_2^2$$

- ℓ_0 -norm optimization is NP-hard
- Formulate it as a mixed-integer programming
 - Introduce binary decision variables $\beta \in \{0, 1\}^d$

$$\min_w \|\tilde{x} - Aw\|_2^2 \quad \iff \quad \min_{w, \beta} \|\tilde{x} - Aw\|_2^2$$

$$\text{s.t. } \underbrace{\|w\|_0 \leq \tau}_{\clubsuit \text{ sparsity of } w} \quad \iff \quad \text{s.t. } \underbrace{-\beta \leq w \leq \beta}_{\text{bounds being either 0 or } \pm 1}, \quad \underbrace{\|\beta\|_1 \leq \tau}_{\clubsuit \text{ sparsity of } \beta}$$

ℓ_0 -Norm Optimization

Sparse Autoregression Done Right

$$\min_{\mathbf{w}, \boldsymbol{\beta}} \underbrace{\sum_{t=d+1}^T \left(x_t - \sum_{k=1}^d w_k x_{t-k} \right)^2}_{\text{Time series autoregression}} \quad \text{s.t. } \underbrace{-\beta_k \leq w_k \leq \beta_k}_{\text{Lower and upper bounds}}, \quad \underbrace{\sum_{k=1}^d \beta_k \leq \tau}_{\clubsuit \text{ Sparsity}} , \quad \underbrace{\beta_k \in \{0, 1\}}_{\text{Binary variable}}$$

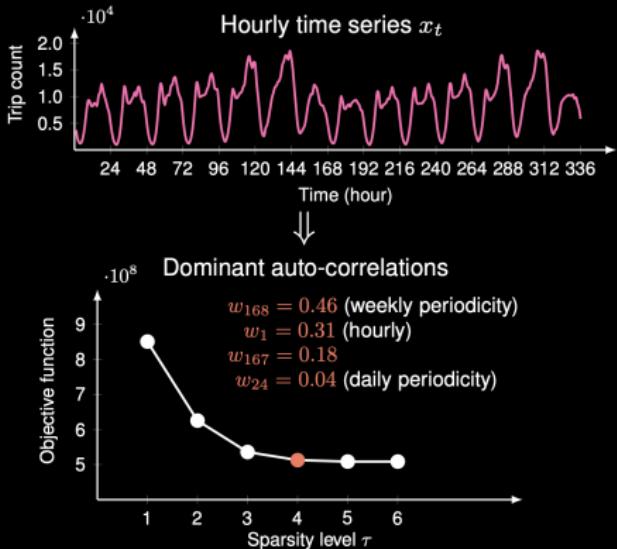
- $\mathbf{w} \in \mathbb{R}^d$: Auto-correlations
- $\boldsymbol{\beta} \in \{0, 1\}^d$: Sparsity pattern
- $d = 168$: Autoregression order

```

1 import numpy as np
2 from docplex.mp.model import Model
3
4 def sparse_ar(x, d, tau):
5     model = Model('Sparse Autoregression')
6     T = x.shape[0]
7     w = [model.continuous_var(name = f'w_{k}') for k in range(d)]
8     beta = [model.binary_var(name = f'beta_{k}') for k in range(d)]
9     model.minimize(model.sum((x[t] - model.sum(w[k] * x[t - k - 1]
10                                for k in range(d))) ** 2
11                                for t in range(d, T)))
12     model.add_constraint(model.sum(beta[k] for k in range(d)) <= tau)
13     for k in range(d):
14         model.add_constraint(w[k] <= beta[k])
15         model.add_constraint(w[k] >= -beta[k])
16     solution = model.solve()
17     return np.array(solution.get_values(w))

```

<https://github.com/xinychen/integers>

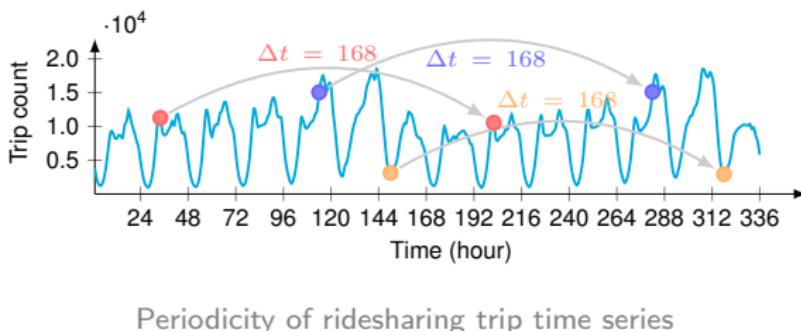


Solution Quality → Better Interpretability?

- Sparse autoregression

$$\min_{\mathbf{w} \geq 0} \|\tilde{\mathbf{x}} - \mathbf{A}\mathbf{w}\|_2^2 \quad \text{s.t. } \|\mathbf{w}\|_0 \leq \tau$$

- Subspace pursuit (SP) sometimes fails

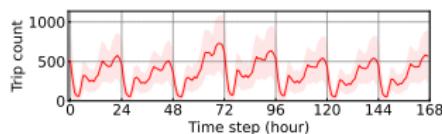
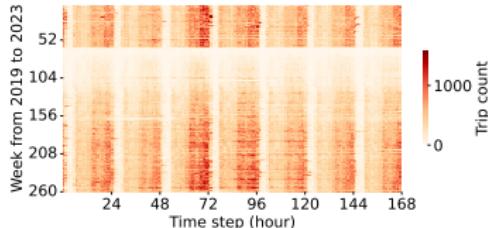


- Exact solution w/ mixed-integer programming (MIP)
- An intuitive example (sparsity $\tau = 2$):

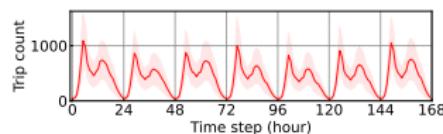
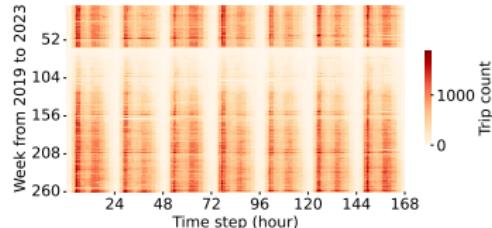
$$\underbrace{\mathbf{w} = (\dots, \underbrace{0.02}_{k=53}, \dots, \underbrace{0.96}_{k=168})^\top}_{\text{obj. } = 8.32 \times 10^7 \text{ (SP)}} \quad \text{vs.} \quad \underbrace{\mathbf{w} = (\underbrace{0.22}_{k=1}, \dots, \underbrace{0.77}_{k=168})^\top}_{\clubsuit \text{ obj. } = 6.25 \times 10^7 \text{ (MIP)}}$$

John F. Kennedy International Airport

- Daily & weekly periodicity: **dropoff > pickup** trips at JFK airport
 - Pickup trips are relevant to flight delay, baggage claim, and other factors.
 - Dropoff trips to airport are highly related to flight schedules.



Pickup trips from airport



Dropoff trips to airport

- Sparse coefficient vectors (**sparsity $\tau = 3$**):

$$\mathbf{w} = (\underbrace{0.31}_{k=1}, \dots, \underbrace{0.28}_{k=24}, \dots, \underbrace{0.41}_{k=168})^\top \quad \text{vs.} \quad \mathbf{w} = (\underbrace{0.18}_{k=1}, \dots, \underbrace{0.35}_{k=24}, \dots, \underbrace{0.47}_{k=168})^\top$$

High-Dimensional Sparse Autoregression

- On high-dimensional time series with a large N :

$$\begin{aligned} & \underbrace{\min_{\{\boldsymbol{w}_n\}_{n=1}^N, \boldsymbol{\beta}}}_{(N+1)d \text{ decision var.}} \quad \underbrace{\sum_{n=1}^N \|\tilde{\boldsymbol{x}}_n - \boldsymbol{A}_n \boldsymbol{w}_n\|_2^2}_{\text{multivariate time series}} \\ \text{s.t.} \quad & \underbrace{0 \leq \boldsymbol{w}_n \leq \boldsymbol{\beta},}_{\text{bounds being either 0 or 1}} \quad \underbrace{\|\boldsymbol{\beta}\|_1 \leq \tau,}_{\text{sparsity of } \boldsymbol{\beta}} \quad \boldsymbol{\beta} \in \{0, 1\}^d \end{aligned}$$

- How to handle millions of time series (e.g., $N \geq 10^6$)?

High-Dimensional Sparse Autoregression

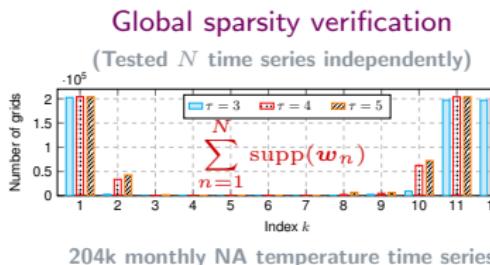
- On high-dimensional time series with a large N :

$$\begin{aligned}
 & \underbrace{\min_{\{\mathbf{w}_n\}_{n=1}^N, \boldsymbol{\beta}}}_{(N+1)d \text{ decision var.}} \quad \underbrace{\sum_{n=1}^N \|\tilde{\mathbf{x}}_n - \mathbf{A}_n \mathbf{w}_n\|_2^2}_{\text{multivariate time series}} \\
 \text{s.t.} \quad & \underbrace{0 \leq \mathbf{w}_n \leq \boldsymbol{\beta},}_{\text{bounds being either 0 or 1}} \quad \underbrace{\|\boldsymbol{\beta}\|_1 \leq \tau,}_{\text{sparsity of } \boldsymbol{\beta}} \quad \boldsymbol{\beta} \in \{0, 1\}^d
 \end{aligned}$$

- How to handle millions of time series (e.g., $N \geq 10^6$)?
- Two-stage optimization (♣):

① Learn sparsity patterns in $\boldsymbol{\beta} \in \{0, 1\}^d$

$$\begin{aligned}
 & \min_{\mathbf{w}, \boldsymbol{\beta}} \underbrace{\text{tr}(\mathbf{w} \mathbf{w}^\top \mathbf{P})}_{\text{quadratic}} - \underbrace{2 \mathbf{w}^\top \mathbf{q}}_{\text{linear}} \\
 \text{s.t.} \quad & 0 \leq \mathbf{w} \leq \boldsymbol{\beta}, \quad \|\boldsymbol{\beta}\|_1 \leq \tau
 \end{aligned}$$



② Quadratic programming with index set $\Omega = \text{supp}(\boldsymbol{\beta})$

$$\mathbf{w}_n := \arg \min_{\mathbf{w} \geq 0} \|\tilde{\mathbf{x}}_n - \mathbf{A}_n \mathbf{w}\|_2^2 \quad \text{s.t. } w_k = 0, \forall k \notin \Omega$$

Climate Seasonality Patterns

(arXiv:2506.22895)

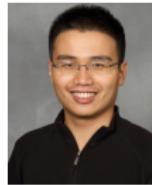
- North America temperature
- Climate variable seasonality
- Sea surface temperature
- Spatiotemporal patterns



Xinyu Chen
MIT



Vassilis Digalakis Jr.
BU



Lijun Ding
UCSD



Dingyi Zhuang
MIT



Jinhua Zhao
MIT

Understanding Climate Systems

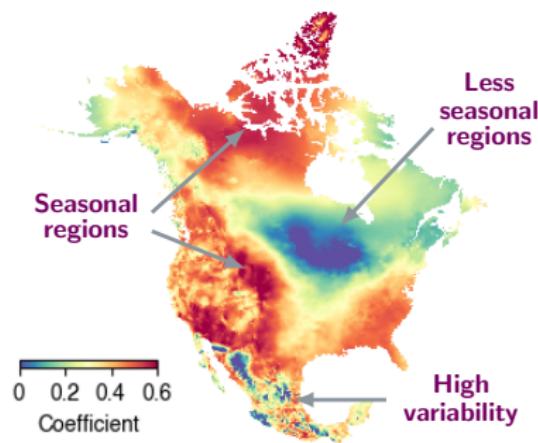
Quantify yearly seasonality by $\{w_{m,n,\gamma,k}\}$ at index $k = 12$

$$\min_{\{\mathbf{w}_{m,n,\gamma}\}, \boldsymbol{\beta}} \sum_{m=1}^M \sum_{n=1}^N \sum_{\gamma=1}^{\delta} \|\tilde{\mathbf{x}}_{m,n,\gamma} - \mathbf{A}_{m,n,\gamma} \mathbf{w}_{m,n,\gamma}\|_2^2$$

longitude
latitude | decade monthly
 ↓ ↓ ↓ temperature

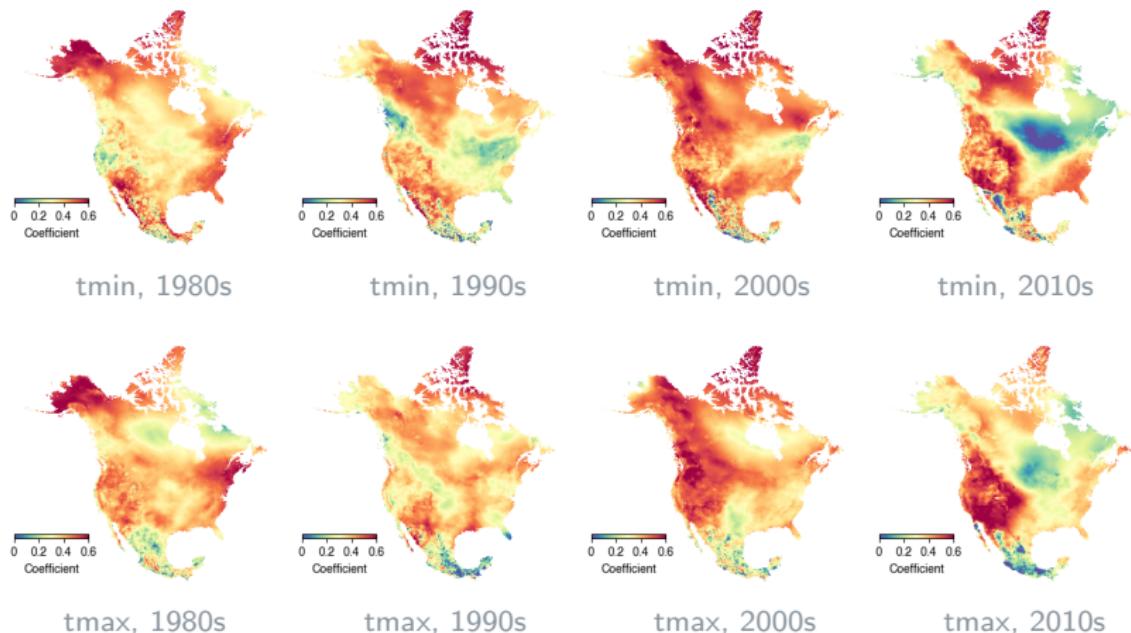
s.t. $0 \leq \mathbf{w}_{m,n,\gamma} \leq \boldsymbol{\beta}$ sparsity constraint
 $\|\boldsymbol{\beta}\|_1 \leq \tau$
 $\boldsymbol{\beta} \in \{0, 1\}^d$ binary decision var.

North America Temperature



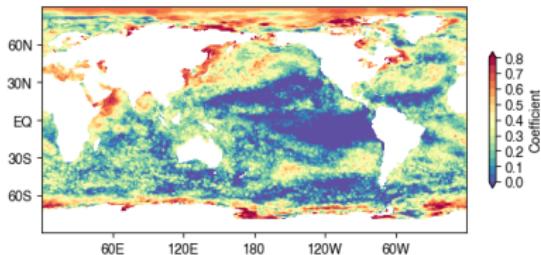
Yearly temperature **seasonality** pattern in 2010s

North America Temperature

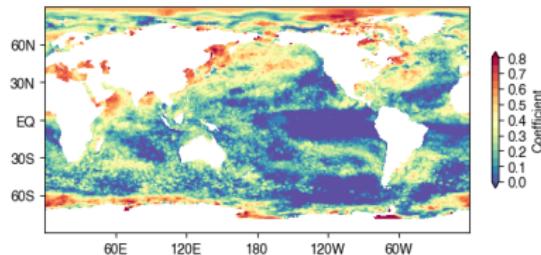


- Identify yearly periodicity at $k = 12$ from temperature data ($\tau = 3$)
 - ❶ Stronger yearly seasonality in high-latitude areas
 - ❷ Less seasonal temperature in south areas (e.g., Mexico)
 - ❸ Seasonality patterns in 2000s & 2010s are different from 1980s & 1990s

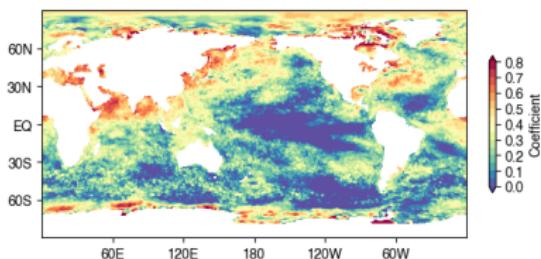
Sea Surface Temperature



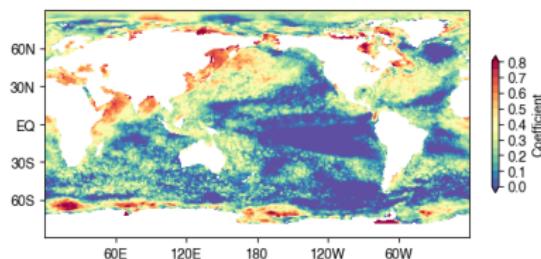
1980s



1990s



2000s



2010s

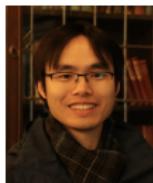
- Identify yearly periodicity at $k = 12$ from SST data ($\tau = 3$)
 - ➊ The areas of El Niño events are less seasonal/predictable
 - ➋ Arctic becomes less seasonal/predictable in the past 20 years

Human Mobility Regularity

Applications and Case Studies

(arXiv:2508.03747)

- Daily & weekly periodicity
- NYC & Chicago ridesharing
- Multi-modal mobility
- Network resilience
- Post-pandemic recovery
- Metro passenger flow



Xinyu Chen
MIT



Qi Wang
NEU



Yunhan Zheng
MIT → PKU



Nina Cao
MIT



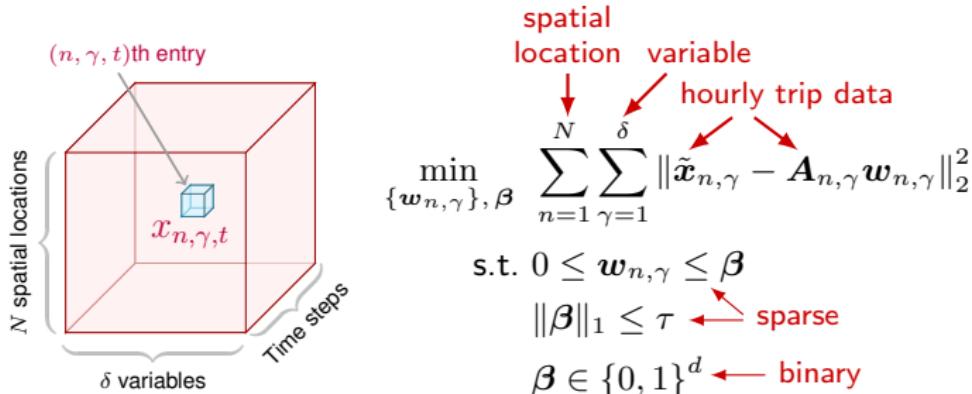
HanQin Cai
UCF



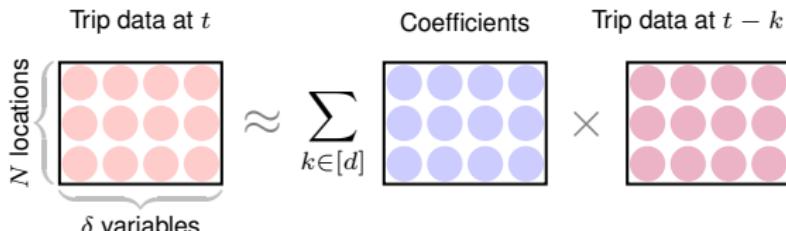
Jinhua Zhao
MIT

Envisioning Human Mobility

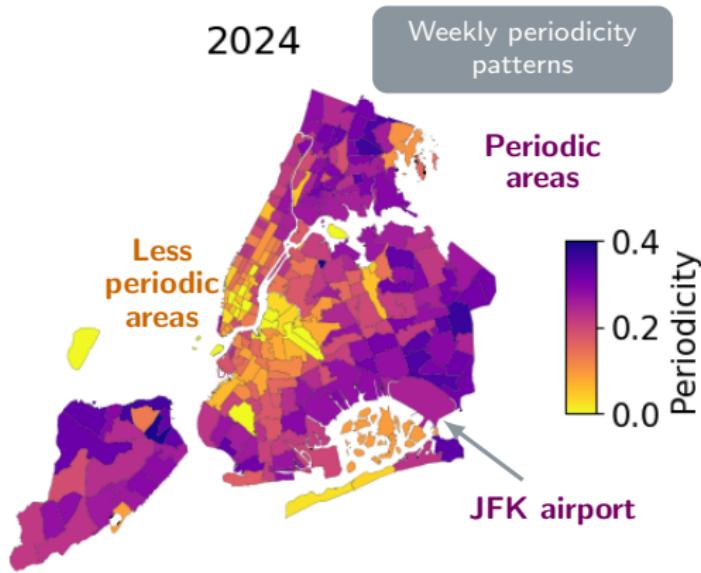
- Human mobility data



- Quantify **weekly periodicity** by $\{\mathbf{w}_{n,\gamma,k}\}$ at index $k = 168$

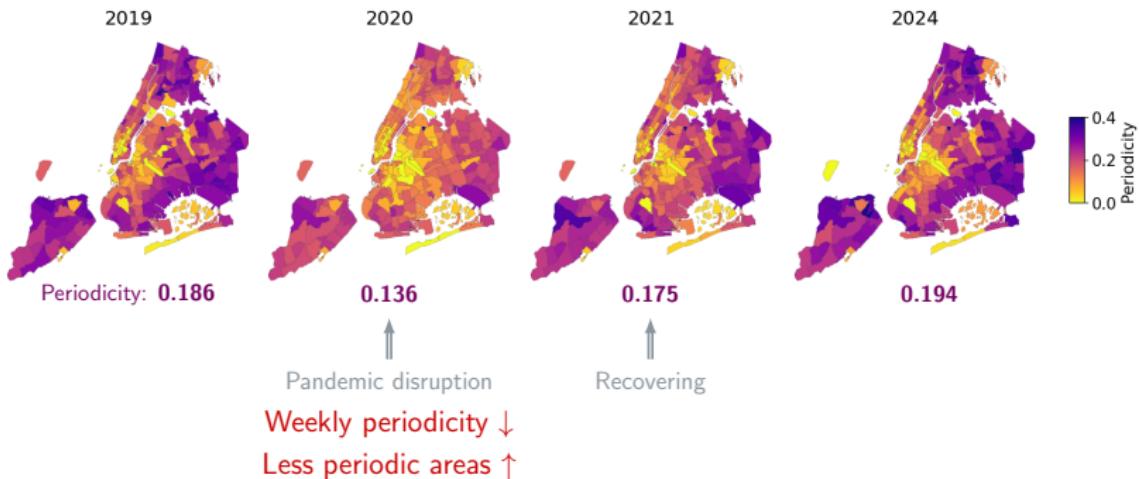


NYC Ridesharing

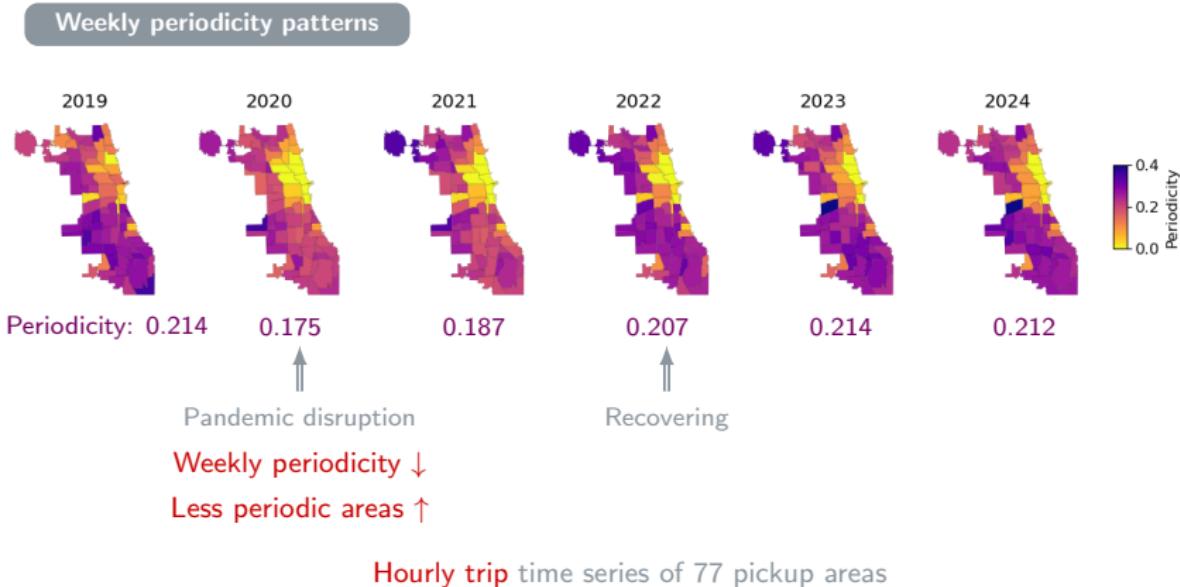


NYC Ridesharing

Weekly periodicity patterns

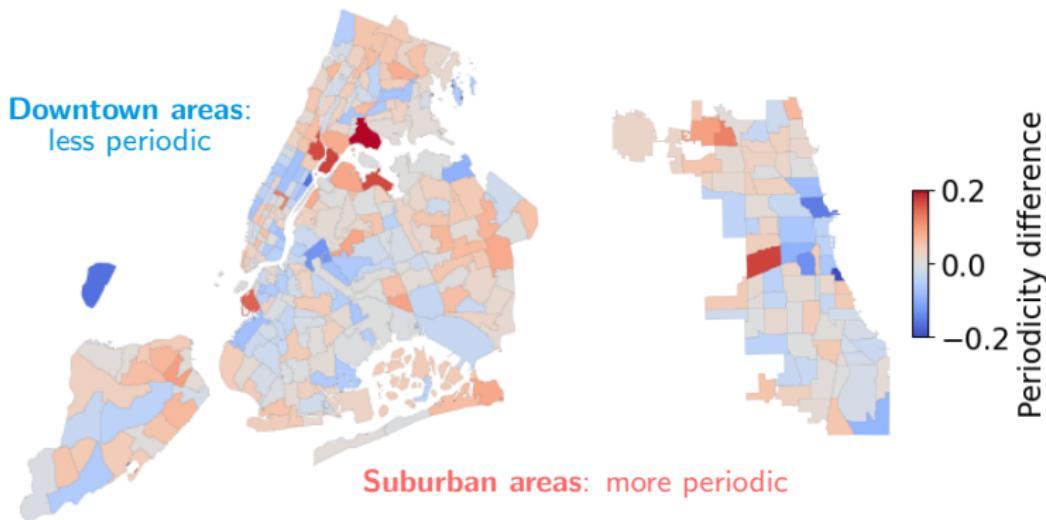


Chicago Ridesharing



Post-Pandemic Recovery

2024's periodicity minus 2019's periodicity



Weekly Periodicity of Manhattan Mobility

Mobility data of 2024

North areas:

Ridesharing > yellow taxi

Overall:

Subway > others

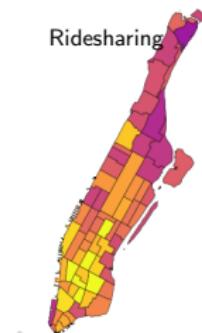
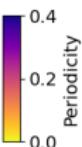
Ridesharing

Yellow taxi

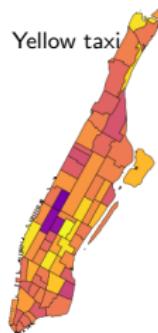
Subway

Bikesharing
(member)

Bikesharing
(all)



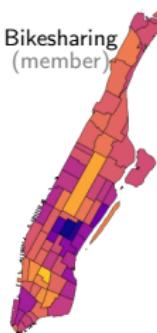
Periodicity: 0.122



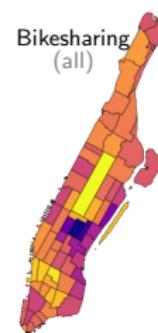
0.120



0.363



0.191



0.151

South areas:
Yellow taxi > Ridesharing

Remarkable weekly periodicity difference
between membership and all travels

Daily Periodicity of Manhattan Mobility (Weekdays)

Mobility data of 2024

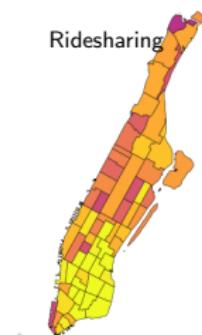
Overall:

Yellow taxi > ridesharing

Subway:

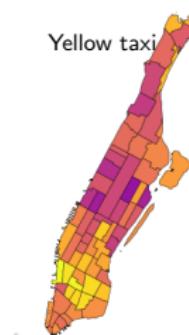
North areas > south areas

Ridesharing



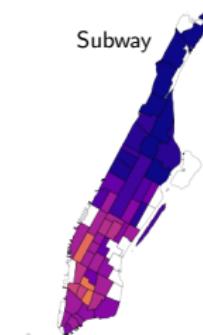
Periodicity: 0.072

Yellow taxi



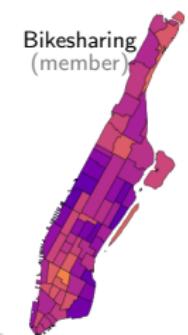
0.127

Subway



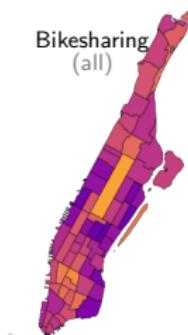
0.305

Bikesharing
(member)



0.232

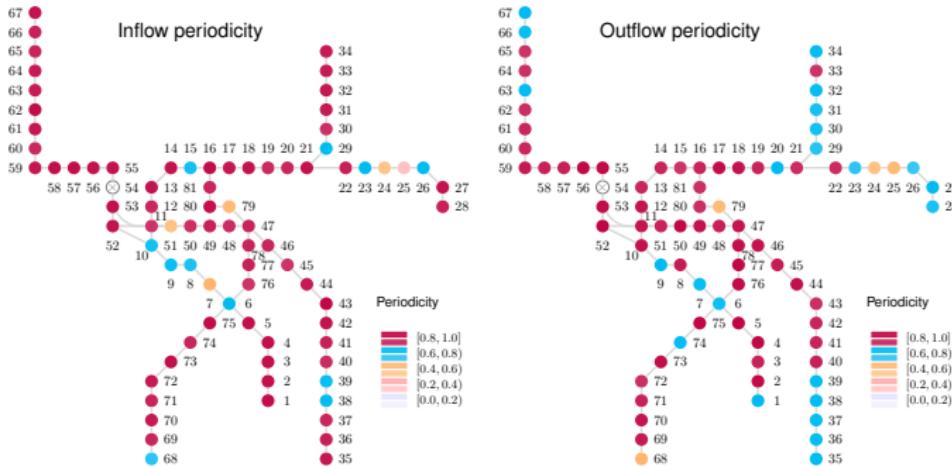
Bikesharing
(all)



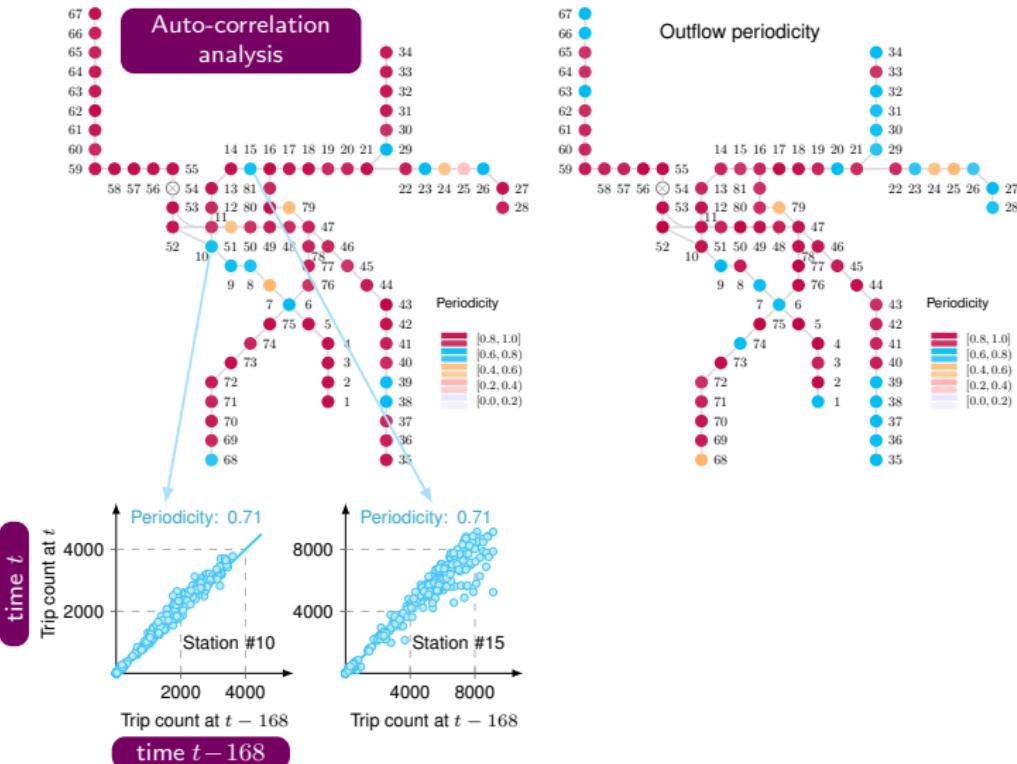
0.214

Marginal daily periodicity difference
between membership and all travels

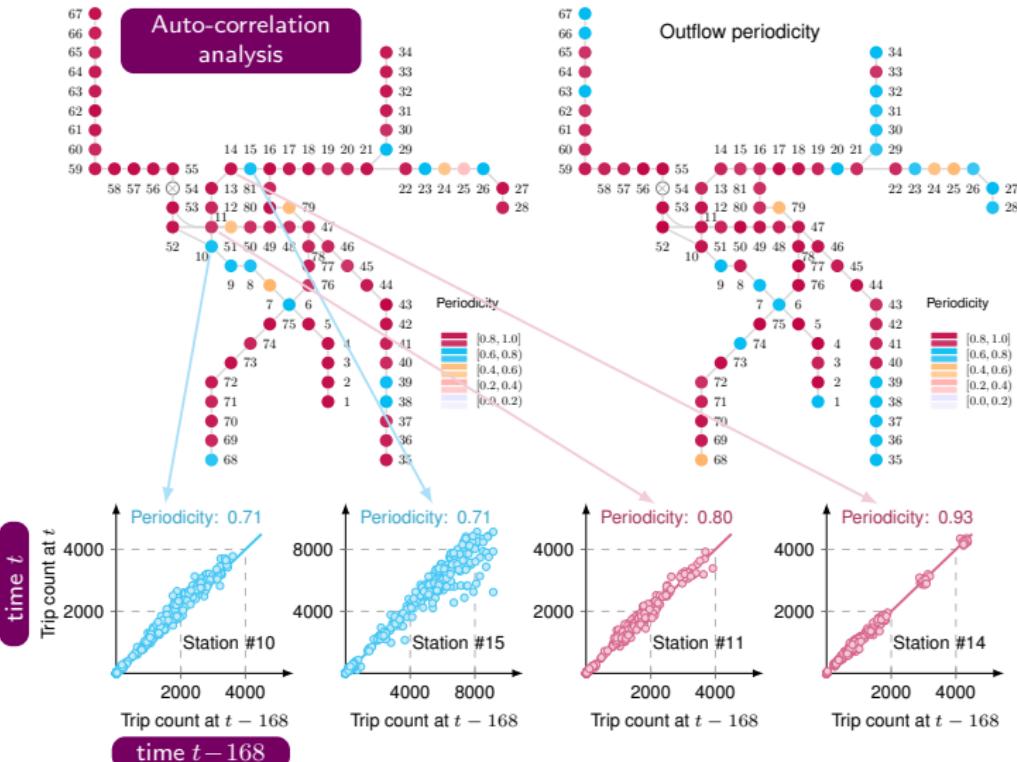
Hangzhou Metro Passenger Flow



Hangzhou Metro Passenger Flow



Hangzhou Metro Passenger Flow



"Closeness" to the anti-diagonal curve $y = x$



Thanks for your attention!

Any Questions?

Slides: <https://xinychen.github.io/slides/intro.pdf>

About me:

- 🏠 Homepage: <https://xinychen.github.io>
- /github/ GitHub: <https://github.com/xinychen>
- ✉️ How to reach me: xinychen@mit.edu