Discovering Human Mobility Periodicity from Tensor Data with Sparse Autoregression

Xinyu Chen*
Department of Urban Studies and Planning, MIT
Qi Wang
Civil and Environmental Engineering, Northeastern University
Yunhan Zheng
Department of Civil and Environmental Engineering, MIT
Nina Cao
Department of Mechanical Engineering, MIT
HanQin Cai†
Department of Statistics and Data Science, University of Central Florida
and
Jinhua Zhao
Department of Urban Studies and Planning, MIT
Department of Civil and Environmental Engineering, MIT

October 30, 2025

Abstract

Human mobility regularity at daily and weekly cycles always demonstrates dynamic patterns and evolves with the external impact of pandemic disruption and policy intervention. This study discovers periodicity patterns from complex human mobility data across different spatial areas, years, travel modes, and variables in cities by using sparse autoregression. We introduce ℓ_0-norm induced sparsity and non-negativity constraints for identifying dominant auto-correlations from time series. Since human mobility data are in the form of tensors, we develop a multidimensional sparse autoregression method and apply it to large-scale metro passenger flow data in

*This research is based upon work supported by the U.S. Department of Energy's Office of Energy Efficiency and Renewable Energy (EERE) under the Vehicle Technology Program Award Number DE-EE0009211 and and DE-EE0011186. The Mens, Manus, and Machina (M3S) is an interdisciplinary research group (IRG) of the Singapore MIT Alliance for Research and Technology (SMART) center.

†The research is partially supported by NSF DMS 2304489.

Hangzhou, China and multimodal mobility data in New York City (NYC) and Chicago, USA. The analysis of NYC and Chicago ridesharing data from 2019 to 2024 demonstrates the spatiotemporal periodicity patterns and disruptive impact of the pandemic on weekly periodicity and the subsequent recovery trends in 2021 and 2022. Given the multimodal mobility data of Manhattan in 2024, we find that the periodicity patterns of ridesharing, taxi, subway, and bikesharing trips uncover temporal regularities and variability of these mobility data. This study highlights the interpretability of sparse autoregression for discovering spatiotemporal mobility patterns and offers a valuable tool for monitoring the shift of mobility regularity in urban systems.

Keywords: Tensor time series; Sparse autoregression; Periodicity quantification; Urban systems; Human mobility; Weekly periodicity

1 Introduction

Human mobility in cities often reveals recurring patterns, i.e., the predictable ebbs and flows of people moving through urban areas, that reflect the underlying structure of urban life (Gonzalez et al. 2008, Song et al. 2010, Simini et al. 2012). From daily commutes to weekend trips, these temporal regularities in human mobility underpin everything from urban transportation management (Çolak et al. 2016) and transit scheduling (Guihaire & Hao 2008) to predicting the spread of diseases (Balcan et al. 2009, Castillo-Chavez et al. 2016, Li et al. 2022, Du et al. 2018). In essence, human mobility regularity is often principled by daily and weekly cycles underlying periodic trips. An intuitive example of periodic trips in human mobility is daily commuting, where passenger flows typically exhibit morning and afternoon peaks, forming a roughly "M"-shaped pattern in trip time series data (Chen & Sun 2022), see e.g., the passenger flow time series in Figure 1B. As cities evolve in a long-term range, mobility patterns have become increasingly complex and less predictable, characterized by complicated regulatory policies, diverse transportation modes, and behavioral shifts driven by external factors (Tirachini 2020, Li et al. 2022). Another fact is that mobility flows can be severely disrupted by extreme events such as the COVID-19 pandemic, leading to a remarkable decline in mobility regularity (Li et al. 2022, Chen, Cai, Liu & Zhao 2025), and may recover gradually over time. Thus, understanding the rhythm of human mobility in urban areas with statistical tools is essential for grasping how our societies function and urban systems operate (Gonzalez et al. 2008, Song et al. 2010, Simini et al. 2012, Jiang et al. 2017, Alessandretti et al. 2020).

A paradigm for discovering long-term human mobility relies on collecting trip records and constructing tensor time series with a certain time resolution across different spatial locations and variables. In Figure 1A, each entry of human mobility tensor corresponds to the trip count value of a specific variable γ at a particular spatial location n and time step

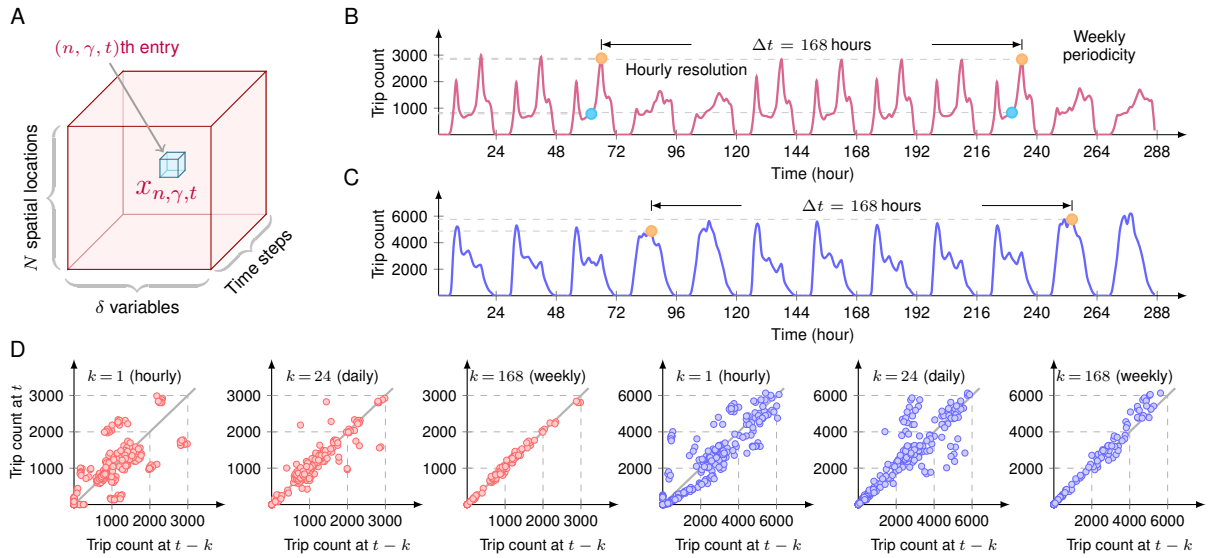

Figure 1: Human mobility data in the form of tensor time series in panel A. Two example passenger flow time series in panels B-C demonstrate clear weekly periodicity by observing auto-correlation plots in panel D with specific time lags $k = 1, 24, 168$, corresponding to hourly, daily, and weekly cycles, respectively.

t. The variable dimension γ is flexible and can encode different aspects of mobility, such as inflow and outflow directions at a particular spatial location. Moreover, the temporal lags such as data collected from different years can also be incorporated as distinct variables within this tensor structure. To explore human mobility regularity underlying these trip time series, a few critical but unsolved questions arise as follows. First, there is no standard statistical method to quantify and compare the strength of mobility periodicity with different noise levels and temporal structures. For example, the two time series of passenger flow in Figures 1B–C show different periodic trends, the question that remains is how to compute the comparable metric for time series periodicity. Second, there is no established approach to interpret periodicity patterns across both space and time. Previous research has adopted methods such as Fourier transform (Li et al. 2010, Prabhala et al. 2014, Piccardi et al. 2024), Bayesian structural time series methods (Hyndman & Athanasopoulos 2018), entropy-based

measures (Goulet-Langlois et al. 2017, Teixeira et al. 2021), dynamic mode decomposition (Brunton & Kutz 2022), principal component analysis (Piccardi et al. 2024), and clustering of similar patterns (Manley et al. 2018, Zhang & Song 2022) to identify temporal regularities and cyclical patterns from time series. However, these approaches are not well-suited to quantifying periodicity across spatiotemporal domains and discovering dynamic patterns that related to human mobility regularity.

Recently, interpretable machine learning methods (Murdoch et al. 2019, Rudin et al. 2022, Brunton & Kutz 2022) offer new opportunities to improve robustness and interpretability of data-driven exploration in real-world systems. In terms of model development, sparse linear regression (Tibshirani 1996, Jenatton et al. 2011) can be used to reformulate the optimization problem of sparse identification of nonlinear dynamical systems (Brunton et al. 2016), which laid the foundation of several machine learning frameworks that designed for variable and feature selection in scientific discovery. In that venue, sparsity is one of the most important aspects to reinforce interpretability of machine learning methods (Murdoch et al. 2019, Rudin et al. 2022). However, existing sparse methods face critical limitations when applied to large-scale human mobility data. First, classical sparse regression methods are primarily designed for variable selection from univariate data and struggle to scale to multidimensional data with spatiotemporal dependencies. Second, mobility data are inherently noisy, irregular, and heterogeneous across spatiotemporal domains, which challenges standard sparse regression methods that lack built-in mechanisms for structured robustness. Third, even when sparsity is enforced, the resulting models do not yield interpretable or comparable metrics for time series periodicity across different spatial locations, variables, and time phases.

To bridge these gaps, we develop a multidimensional sparse autoregression method by reformulating the optimization problem of the interpretable time series autoregression

proposed by Chen, Digalakis Jr, Ding, Zhuang & Zhao (2025). The method characterizes time series of mobility flows across space and time, captures dominant auto-correlations explicitly, and provides an interpretable measure of time series periodicity at daily and weekly cycles. We carry out extensive data-driven exploration of periodicity patterns on large-scale human mobility data collected from different cities over several years. First, we demonstrate that our method effectively discovers the weekly periodicity from metro passenger flow time series with complicated dimensions, including station, inflow/outflow direction, and time step. The results clearly identify the periodicity pattern difference between inflow and outflow directions of passenger flows. Although time resolution is an influential factor for quantifying time series periodicity, the periodicity patterns are comparable across different stations and flow directions with the same time resolution. Second, we utilize millions of ridesharing trip records in NYC and Chicago from 2019 to 2024 and aggregate these data as mobility tensors. The method clearly discovers spatial patterns of weekly periodicity, demonstrating a remarkable decrease of human mobility regularity in 2020 and the recovery in the post-pandemic period. Third, we compare the periodicity patterns learned from massive trip records of ridesharing, yellow taxi, subway, and bikesharing in 2024 collected from Manhattan, NYC. The spatial patterns of daily and weekly periodicity across four different travel modes bring insightful evidence for understanding human mobility regularity.

The rest of this study is organized as follows. Section 2 introduces the preliminaries of time series auto-correlations and sparse autoregression. Section 3 describes the multidimensional sparse autoregression method for quantifying human mobility periodicity. To carry out a thorough empirical study, we perform extensive experiments on metro passenger flow data (see Section 4), NYC and Chicago ridesharing data (see Section 5), and multimodal Manhattan mobility data (see Section 6). Finally, Section 7 draws the conclusion.

2 Preliminaries

2.1 Time Series & Auto-Correlation Analysis

In urban systems, human mobility data often involve both spatial and temporal dimensions. As shown in Figure 1A, trip time series can be represented as a third-order tensor of non-negative integers $x_{n,\gamma,t} \in \mathbb{Z}^+$ where $n \in [N]$ indexes spatial locations, $\gamma \in [\delta]$ indexes different variables or time phases, and $t \in [T_\gamma]$ indexes time steps with a certain time resolution. For notational convenience, we use $[h] = \{1, 2, \ldots, h\}$ for any integer $h > 0$ and $\mathbb{Z}^+$ as the set of nonzero integers throughout this work. Each univariate time series $\{x_{n,\gamma,t}\}_{t\in[T_\gamma]}$ for any spatial location $n \in [N]$ and variable $\gamma \in [\delta]$ can be seamlessly modeled by autoregressive process (Hamilton 2020). Consequently, autoregression methods allow a preliminary analysis of inherent periodicity patterns in trip time series of human mobility (Chen & Sun 2022). However, it is unclear about how to identify periodicity patterns of multidimensional tensor time series and make these patterns interpretable and comparable in the spatiotemporal context.

As exemplified by Figure 1B, real-world passenger flow time series demonstrate a clear weekly periodicity by manually selecting some data points with a weekly cycle. To intuitively analyze the dominant auto-correlations, Figure 1D presents auto-correlation plots that visualize the relationship between trip counts at a given time step t and the previous time step $t - k$, in which $k \in \mathbb{Z}^+$ represents a certain time lag. In the case of hourly time series, $k = 1, 24, 168$ correspond to hourly, daily, and weekly cycles, respectively. When these time series data points are denoted by $x_t, t \in [T]$, the auto-correlation plots in Figure 1D can be formulated as $x_t \approx w_k x_{t-k}$, measuring the self-similarity of time series trends with a certain time lag $k \in \mathbb{Z}^+$. Scatters that align closely along the antidiagonal curve (i.e., equality $x_t - x_{t-k} = 0$) imply positive auto-correlation and temporal regularities. Moreover, the

"closeness" of data points to the antidiagonal curve at time lag $k = 168$ justifies a clear weekly periodicity of time series. According to the rule of antidiagonal "closeness", the time series in Figure 1B is more periodic than Figure 1C at a weekly cycle. Furthermore, the auto-correlation plots of both time series with time lag $k = 24$ are less close to the antidiagonal curve compared to the time lag $k = 168$. Thus, the weekly periodicity of both time series is stronger than the daily periodicity.

2.2 Sparse Autoregression

Time series autoregression is a classical statistical method that formulates each time series data point as a linear combination of its past data points, plus noise (Hamilton 2020). For any time series $\{x_t\}_{t\in[T]}$ of length T, the dth-order autoregression is $x_t \approx \sum_{k\in[d]} w_k x_{t-k}$ whose optimization problem is minimizing the sum of squared autoregressive errors as shown in Figure 2A. Herein, $\boldsymbol{w} = (w_1, w_2, \cdots, w_d)^\top \in \mathbb{R}^d$ is the coefficient vector of length $d \in \mathbb{Z}^+$, denoting a sequence of coefficients or auto-correlations with order $d < T$. To address the issue of high dimensionality and difficult-to-interpret temporal dependencies, sparse autoregression such as sparse vector autoregression (Valdés-Sosa et al. 2005, Basu & Michailidis 2015, Davis et al. 2016) has been verified as an effective method that assumes many of autoregressive coefficients being zero. Furthermore, these sparse autoregression methods can be used to reformulate the estimation problem of network Granger causality (Basu et al. 2015).

In this work, the sparse autoregression starts from the core idea of interpretable time series autoregression. The objective of sparse autoregression is to find a set of coefficients that minimize the sum of squared autoregressive errors in the presence of sparsity and non-negativity constraints (Chen, Digalakis Jr, Ding, Zhuang & Zhao 2025). In particular, sparsity, modeled by the ℓ_0-norm of coefficient vector $\boldsymbol{w}$ and denoted by $\|\boldsymbol{w}\|_0$, limits the

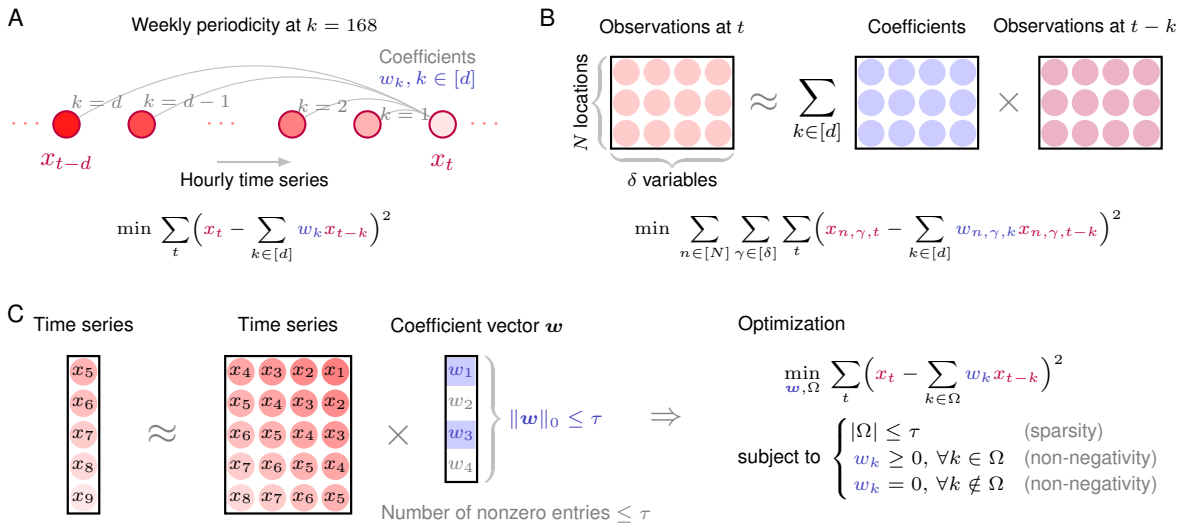

Figure 2: Illustration of time series autoregression. (A-B) Autoregression on the univariate and multidimensional time series. (C) Sparse autoregression on the univariate time series.

number of nonzero coefficients with an upper bound $\tau \in \mathbb{Z}^+$, satisfying $\tau < d$. This implies that only a small number of time lags significantly contribute to predicting the current time series data point. The sparse autoregression assumes the non-negative coefficient vector $\boldsymbol{w} \in \mathbb{R}^d$ because the periodic cycles are often present in the positive auto-correlations. As a result, the optimization problem of sparse autoregression is constructed as follows,

$$\begin{aligned} \min_{\boldsymbol{w}\ge 0} \quad & \sum_t \left(x_t - \sum_{k\in[d]} w_k x_{t-k}\right)^2 \\ \text{s.t.} \quad & \|\boldsymbol{w}\|_0 \le \tau, \end{aligned} \tag{1}$$

where the upper bound τ in the constraint is defined as sparsity level. If the time series such as Figures 1B–C demonstrates remarkable periodicity, then the non-negativity constraint allows the automatic identification of positive auto-correlations. Let Ω be the index set of nonzero coefficients in $\boldsymbol{w} \in \mathbb{R}^d$, the ℓ_0-norm can be converted into cardinality of index set Ω, formally denoted by $|\Omega|$. Thus, the optimization problem in Eq. (1) becomes

$$\begin{aligned} \min_{\boldsymbol{w}\ge 0,\, \Omega} \quad & \sum_t \left(x_t - \sum_{k\in\Omega} w_k x_{t-k}\right)^2 \\ \text{s.t.} \quad & |\Omega| \le \tau, \quad w_k = 0, \forall k \notin \Omega, \end{aligned} \tag{2}$$

including both coefficient vector $\boldsymbol{w} \in \mathbb{R}^d$ and index set Ω.

3 Methodology

In this section, we elaborate on ℓ_0-norm induced sparse autoregression for discovering temporal regularities from human mobility data. As demonstrated by Figure 2B, multidimensional time series autoregression specifies that the coefficient vectors $\{\boldsymbol{w}_{n,\gamma}\}_{n\in[N],\gamma\in[\delta]}$ over N spatial locations and δ variables can be optimized simultaneously. The index set Ω is assumed to be globally sparse across $N\delta$ coefficient vectors and constrained by the upper bound $\tau \in \mathbb{Z}^+$ on the cardinality. On the mobility tensor data $\{x_{n,\gamma,t}\}_{n\in[N],\gamma\in[\delta],t\in[T_\gamma]}$ as shown in Figure 1A, we formulate the optimization problem of multidimensional sparse autoregression with order $d \in \mathbb{Z}^+$ as follows,

$$\begin{aligned} \min_{\{\boldsymbol{w}_{n,\gamma}\geq 0\}_{n\in[N],\gamma\in[\delta]},\Omega} \quad & \sum_{n\in[N]}\sum_{\gamma\in[\delta]}\sum_{t\in[T_\gamma]}\left(x_{n,\gamma,t}-\sum_{k\in\Omega} w_{n,\gamma,k}x_{n,\gamma,t-k}\right)^2 \\ \text{s.t. } & |\Omega|\leq\tau, \quad w_{n,\gamma,k}=0,\ \forall n\in[N],\gamma\in[\delta],k\notin\Omega, \end{aligned} \tag{3}$$

where $w_{n,\gamma,k}$, $n \in [N], \gamma \in [\delta], k \in [d]$ are the non-negative coefficients. In literature, such an optimization problem with ℓ_0-norm induced sparsity constraint can be addressed by mixed-integer optimization techniques (Bertsimas et al. 2016, 2020, 2024, Tillmann et al. 2024). Recently, estimating the coefficient vectors in Eq. (3) has an efficient two-stage algorithm developed by Chen, Digalakis Jr, Ding, Zhuang & Zhao (2025), including 1) the estimation of global sparsity patterns, and 2) the learning process of individual coefficient vectors. In this case, the optimization problem for estimating a global coefficient vector $\boldsymbol{w} \in \mathbb{R}^d$ from tensor data $\{x_{n,\gamma,t}\}_{n\in[N],\gamma\in[\delta],t\in[T_\gamma]}$ is given by

$$\begin{aligned} \min_{\boldsymbol{w}\geq 0,\Omega} \quad & \sum_{n\in[N]}\sum_{\gamma\in[\delta]}\sum_{t\in[T_\gamma]}\left(x_{n,\gamma,t}-\sum_{k\in\Omega} w_k x_{n,\gamma,t-k}\right)^2 \\ \text{s.t. } & |\Omega|\leq\tau, \quad w_k\geq 0,\ \forall k\notin\Omega, \end{aligned} \tag{4}$$

which can be converted into a mixed-integer optimization by replacing the index set Ω with binary decision variables. Given the index set Ω that optimized from Eq. (4), then learning the individual coefficient vector $\boldsymbol{w}_{n,\gamma} \in \mathbb{R}^d$, $\forall n \in [N], \gamma \in [\delta]$ can be simplified as the following quadratic optimization:

$$\begin{aligned} \min_{\boldsymbol{w}_{n,\gamma} \geq 0} \quad & \sum_{t \in [T_\gamma]} \left(x_{n,\gamma,t} - \sum_{k \in \Omega} w_{n,\gamma,k} x_{n,\gamma,t-k} \right)^2 \\ \text{s.t.} \quad & w_{n,\gamma,k} = 0, \, \forall k \notin \Omega, \end{aligned} \tag{5}$$

for any spatial location $n \in [N]$ and variable $\gamma \in [\delta]$.

In the model setting, the order d should be set as inherent periodic cycles or greater, e.g., $d \geq 168$ for hourly time series which might demonstrate weekly periodicity as illustrated in Figure 2A. In what follows, we consider the order $d = 168$ for the hourly human mobility datasets. While the constraint $|\Omega| \leq \tau$ requires selecting a subset of τ coefficients out of d possible ones, we extract the coefficients at time lag $k = 168$ and interpret these coefficients as the weekly periodicity across N spatial locations and δ variables. In addition, if τ is very small, then the time lag $k = 168$ might not be in the index set Ω, making it difficult to quantify weekly periodicity. In contrast, given a large τ, the quantity of auto-correlations within the index set Ω would have marginal differences.

4 Metro Passenger Flow Periodicity

To demonstrate the effectiveness of the multidimensional sparse autoregression method, we begin with a case study of metro passenger flow dataset, spanning from January 2 to 25, 2019 in Hangzhou, China.[1] This dataset demonstrates clear periodicity patterns related to commuting, making it an ideal testbed for discovering meaningful weekly periodicity across different subway stations and flow directions. The dataset offers millions of individual trip

[1]The anonymized trip record dataset is available at https://doi.org/10.5281/zenodo.3145404.

records in inflow and outflow directions across 81 subway stations. We aggregate these trip records as multidimensional time series that represent the hourly trip counts of 81 stations, 2 flow directions, and 576 time steps across 24 days. As a result, the mobility tensor is of size $81 \times 2 \times 576$ in which inflow and outflow directions are two variables, denoted by $\gamma = 1$ and 2, respectively.

Figure 3A visualizes the daily average trip counts of 81 subway stations as a heatmap where stations such as #15 and #9 in downtown areas severed a large amount of passengers. Although the hourly time series in this metro passenger flow dataset might have different levels of trip counts, the sparse autoregression method only takes into account the "closeness" between data points $x_{n,\gamma,t}$ and $x_{n,\gamma,t-k}$ with time lag $k \in [d]$. In the experiment, we set the sparsity level as $\tau = 4$, depending on how many dominant auto-correlations we expect to identify. The index set is optimized as $\Omega = \{1, 24, 144, 168\}$, corresponding to hourly, daily, 6-day, and weekly cycles, respectively. The coefficients at time lag $k = 168$ refer to the weekly periodicity, and the periodicity patterns are both interpretable and comparable across different subway stations and flow directions.

Figure 3B shows the weekly periodicity of inflow time series across $N = 81$ stations. As can be seen, passenger flow time series of most stations demonstrate remarkable weekly periodicity greater than 0.8. In contrast, Figure 3C visualizes the weekly periodicity of outflow time series, in which time series of the end stations in subway lines present less remarkable weekly periodicity than the stations in downtown areas. Analyzing the weekly periodicity of metro passenger flow also reveals the mobility patterns of both inflow and outflow directions. Table 1 summarizes a comparison of weekly periodicity between inflow and outflow directions. Observing Figures 3B–C and Table 1, inflow time series at the end stations of subway lines (e.g., $\{27, 28\}$, $\{30, 31, 32, 33, 34\}$, $\{35, 36, 37\}$, and $\{63, 64, 65, 66, 67\}$) relevant to commuting trips show stronger weekly periodicity than

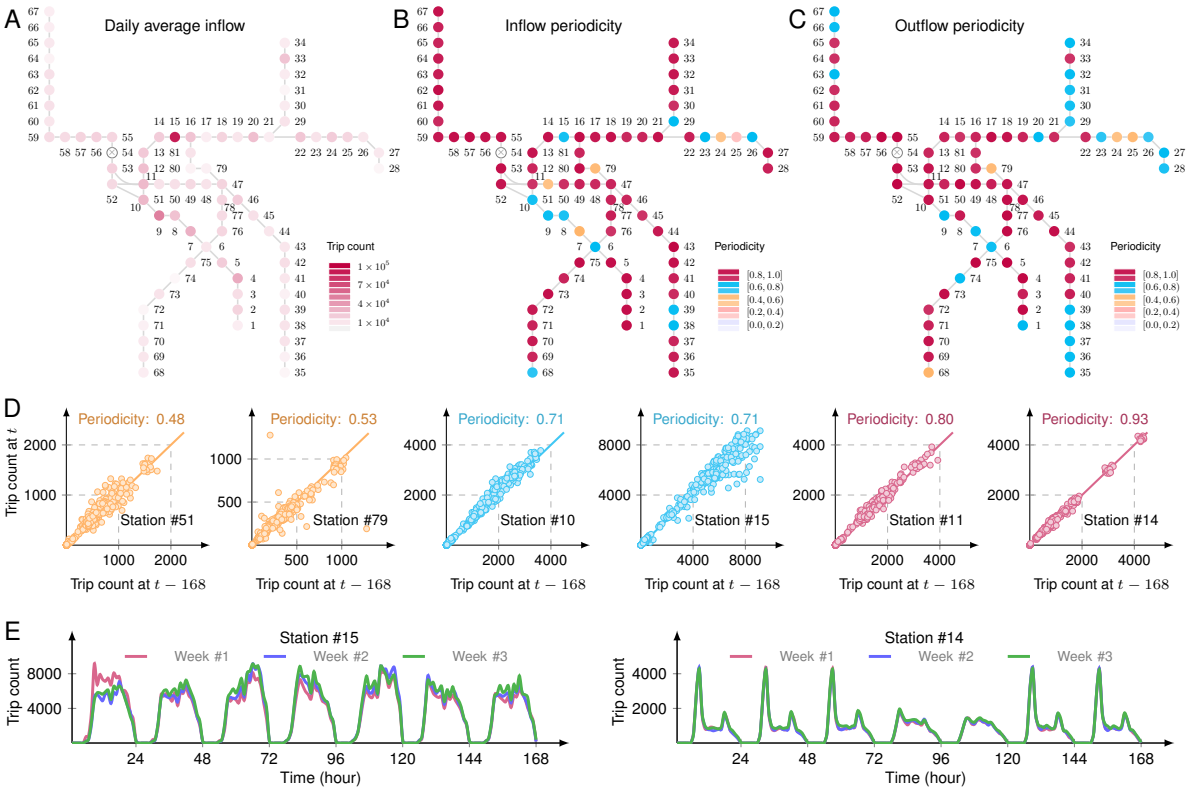

Figure 3: Weekly periodicity of hourly metro passenger flow time series with the multidimensional sparse autoregression method. (A) Daily average inflow trips of 81 metro stations. Note that the trip records of the station #54 are missing. (B) Weekly periodicity of inflow time series. 64 stations show high periodicity (≥ 0.8). (C) Weekly periodicity of outflow time series. 53 stations show high periodicity (≥ 0.8). (D) Auto-correlation plots with a weekly cycle of 6 selected inflow stations. (E) Inflow time series of stations #14 and #15, shown across three consecutive weeks from January 2 to 22, 2019.

outflow time series. Conversely, outflow time series of stations in downtown areas (e.g., $\{7, 8, 10, 15, 51\}$) seem to be more periodic than inflow time series.

Figure 3D shows the auto-correlation plots for inflow time series of 6 selected stations with time lag $k = 168$ (i.e., weekly cycle). The source emphasizes that time series with high weekly periodicity also show a strong self-similarity between data points at times t and

Table 1: Weekly periodicity of the Hangzhou metro passenger flow data at the selected end stations of subway lines.

Station #	Inflow	Outflow	Station #	Inflow	Outflow	Station #	Inflow	Outflow
23	**0.77**	0.65	31	**0.90**	0.70	39	0.73	**0.74**
24	0.46	**0.48**	32	**0.91**	0.74	40	0.83	**0.88**
25	0.24	**0.54**	33	**0.90**	0.80	62	**0.95**	0.82
26	**0.78**	0.62	34	**0.90**	0.76	63	**0.89**	0.77
27	**0.93**	0.78	35	**0.91**	0.76	64	**0.89**	0.81
28	**0.87**	0.69	36	**0.88**	0.74	65	**0.86**	0.80
29	**0.77**	0.66	37	**0.85**	0.74	66	**0.87**	0.73
30	**0.81**	0.71	38	**0.76**	0.75	67	**0.93**	0.79

$t-168$, indicating positive auto-correlations. The weekly periodicity is visually confirmed by auto-correlation plots and time series overlaps. For instance, the inflow time series of station #14 shows a high weekly periodicity of 0.93, and its auto-correlation plot demonstrates a strong alignment along the antidiagonal curve, see Figure 3D. In particular, the comparison between stations #14 and #15 in Figure 3E answers why the time series periodicity is a fair metric to reveal repeating patterns of human mobility. The weekly periodicity of inflow time series of station #15 is 0.71, which is smaller than station #14. As shown in Figure 3E, the inflow time series of station #15 does not align perfectly across three weeks, indicating lower weekly periodicity compared to the more consistent time series trends of station #14.

To take into account the influential factor of time resolution when constructing time series, Figure 4 shows the weekly periodicity of 30-minute time series of the Hangzhou metro passenger flow dataset. The order and sparsity level of the multidimensional sparse autoregression method are set as $d = 336$ and $\tau = 4$, respectively. The index set is optimized

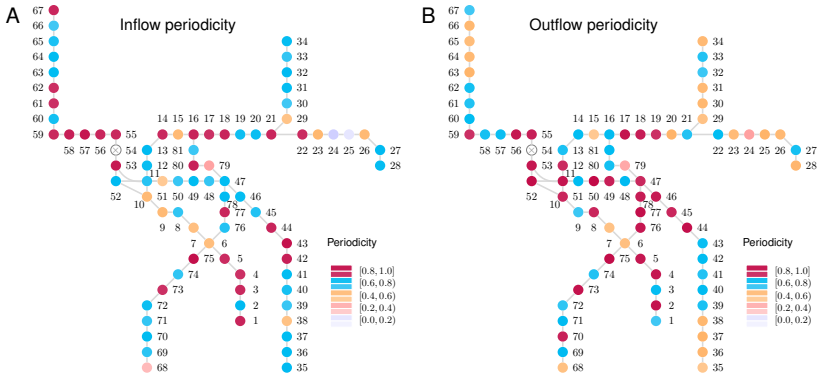

Figure 4: Weekly periodicity of 30-minute metro passenger flow time series with the multidimensional sparse autoregression method.

as $\Omega = \{1, 48, 335, 336\}$, and consequently, the weekly periodicity in Figure 4 is represented by the auto-correlations at time lag $k = 336$. Compared to the hourly time series, the weekly periodicity of 30-minute time series is less remarkable because 30-minute time series demonstrate stronger local auto-correlations than hourly time series. As shown in Figure 4, inflow time series are more periodic than outflow time series in the end stations of subway lines, while the inflow is less periodic than outflow in downtown areas.

5 Ridesharing Trip Periodicity

In this section, we analyze millions of ridesharing trip records of NYC[2] and Chicago[3] that are publicly available on the open data portals. In NYC, the original ridesharing trip records since February 2019 are collected with pickup/dropoff areas (i.e., 265 spatial areas in total) and time information. By aggregating these trip records according to the labels including pickup/dropoff area, year, and hourly time step, one can construct multidimensional tensor time series as shown in Figure 1A, representing hourly trip counts. The ridesharing trip

[2]The NYC ridesharing trip record dataset is available at https://www.nyc.gov/site/tlc/about/tlc-trip-record-data.page.

[3]The Chicago ridesharing trip record dataset is available at https://data.cityofchicago.org/.

records of NYC from 2019 to 2024 are processed as a mobility tensor. Here, $\gamma = 1, 2, \ldots, 6$ refer to the years from 2019 to 2024, respectively, and $T_\gamma \in \mathbb{Z}^+$ is the length of hourly time series in a certain year γ. Analogously, we process the ridesharing trip records of Chicago as a mobility tensor whose entries are the hourly trip counts, and there are 77 spatial areas for anonymizing pickup/dropoff locations.

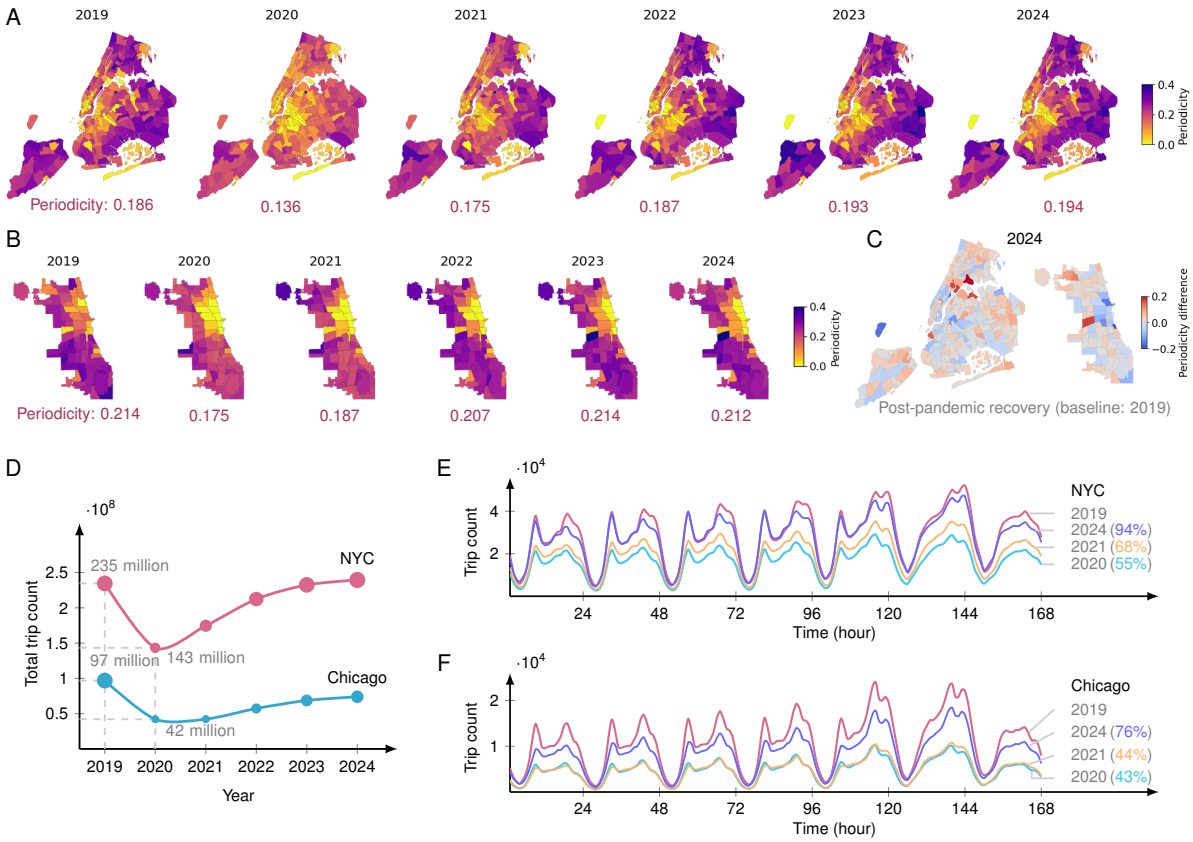

Figure 5: Weekly periodicity of ridesharing trip data with the multidimensional sparse autoregression method. (A) Weekly periodicity of NYC ridesharing pickup trip data across 265 areas. (B) Weekly periodicity of Chicago ridesharing pickup trip data across 77 areas. (C) Weekly periodicity of 2024 minus 2019. (D) Annual ridesharing trip counts of NYC and Chicago from 2019 to 2024. (E-F) Hourly trip count time series averaged across weeks in NYC and Chicago, starting from Monday to Sunday.

In 2020 and 2021, the COVID-19 pandemic had a significant disruptive impact on hu-

man mobility regularity. Both NYC and Chicago experienced a remarkable reduction of ridesharing trips in 2020 and a substantial recovery in the following years, see Figure 5D. To examine how such disruptions affected urban human mobility, we quantify periodicity of ridesharing trips from 2019 to 2024 across different spatial areas and years through the multidimensional sparse autoregression method. In the experiments, we set the sparsity level as $\tau = 6$ on both datasets. The index set is optimized as $\Omega = \{1, 23, 24, 143, 167, 168\}$ whose cardinality is $|\Omega| = 6$, covering the time lags of hourly, daily, and weekly cycles. The resulting weekly periodicity across urban areas extracts spatial patterns and the long-term evolution of human mobility regularity.

In terms of the periodicity results of NYC ridesharing trips in Figure 5A, weekly periodicity measured by the coefficients $w_{n,\gamma,168}, n \in [265]$ of ridesharing pickup trips in 2019 (i.e., $\gamma = 1$) depicts that trip time series of downtown areas such as the south areas of Manhattan are less periodic than suburban areas. Ridesharing trips in 2020 marked a significant shift from 2019, with the pandemic severely disrupting periodicity patterns and reducing the average of weekly periodicity across 265 spatial areas from 0.186 to 0.136. By 2021, the ridesharing trips in NYC began recovering, transitioning from pandemeic-induced irregularity towards the post-pandemic norm. During this recovery phase, ridesharing trips in suburban areas displayed the increase of weekly periodicity compared to 2020. The pattern of weekly periodicity in 2022 became consistent with the pattern in 2019, while both averages of weekly periodicity are almost same. In 2024, the average of weekly periodicity is higher than 2019, as shown in Figure 5C, demonstrating the increase of weekly periodicity in suburban areas and the decrease of weekly periodicity in downtown areas.

Figure 5B presents the spatial patterns of weekly periodicity of ridesharing trips across 77 spatial areas from 2019 to 2024 in Chicago. As can be seen, ridesharing trip time series of downtown areas are less periodic than suburban areas. A similar mobility regularity

compared to NYC emerged where the average of weekly periodicity of ridesharing trips in 2019 is stronger than 2020. However, in contrast to NYC, the pattern transition of weekly periodicity in Chicago is slower, demonstrating marginal pattern changes between 2020 and 2021. The pattern of weekly periodicity in Chicago during 2023–2024 became consistent with the pattern of weekly periodicity in 2019, suggesting a return to pre-pandemic mobility regularity. In 2024, although the ridesharing trips demonstrated the same level of weekly periodicity as 2019, downtown areas became less periodic, see Figure 5C.

The aforementioned patterns of weekly periodicity and their comparison between NYC and Chicago allow a further discussion about mobility resilience. Comparing the aggregated trips in Figure 5D, NYC showed a quick return of ridesharing trip counts in the post-pandemic years and eventually reached the same level of 2019, while Chicago showed a poorer recovery of ridesharing trips. Figure 5E visualizes the hourly trip count time series averaged across weeks in each year and marks the percentage of trip count recovery compared to the baseline, i.e., the total trip count in 2019. NYC ridesharing trips reduced to 55% in 2020, and then showed a recovery percentage as 68% in 2021. In contrast, Figure 5F shows that Chicago ridesharing trips remained at 44% in 2021, which is close to the percentage 43% in 2020. While the average of weekly periodicity of Chicago (see Figure 5B) in 2023 reached the same level as 2019, returning to the normal mobility regularity, the increase of total ridesharing trips between 2023 and 2024 is marginal. Thus, at least observing the total ridesharing trips and weekly periodicity, it might be difficult for Chicago to be fully recovered. Our analysis so far suggests that the COVID-19 pandemic leads to a significant decline in the weekly periodicity and disruptive temporal regularities of urban human mobility in both NYC and Chicago. Both cities showed signs of recovery in 2021–2023, with the pattern of weekly periodicity returning to pre-pandemic levels. However, the pace of recovery in terms of both weekly periodicity and total trip counts in NYC appeared to be faster than Chicago.

6 Mobility Periodicity of Multiple Travel Modes

In urban systems, human mobility patterns vary with different travel modes. It is meaningful to fairly quantify mobility periodicity of different travel modes and utilize periodicity patterns to allocate resources and make policies. In this study, we consider the trip records of ridesharing, yellow taxi, subway[4], and bikesharing[5] in Manhattan, one of the most densely populated boroughs in NYC, across the whole year of 2024. By using the multidimensional sparse autoregression method, this analysis aims to demonstrate the differences of weekly periodicity across four travel modes and the underlying spatiotemporal patterns.

As shown in Figure 6A, the total trip counts of ridesharing, yellow taxi, subway, and bikesharing are 93.1 million, 36.4 million, 677.5 million and 28.1 million, respectively. The daily time series in Figure 6B indicates that subway ridership exhibits the most remarkable periodicity, with highly regular weekly cycles that were sustained throughout 2024. In contrast, ridesharing, yellow taxi, and bikesharing trips display lower periodic structures. Figure 6C further disaggregates bikesharing trips into membership and causal users. Membership users contribute 23 million rides, i.e., 82% of the total bikesharing trips, while causal trips account for the remaining 18%. Bikesharing trips reach their peak during the warmer months (May–October) and decline sharply in winter (e.g., January, February and December), reflecting the sensitivity of cycling activity to weather conditions.

As shown in Figures 7A–B, subway and bikesharing trip records at the station level are spatially projected onto spatial areas of Manhattan, yielding trip time series with the same spatial resolution as ridesharing and yellow taxi. If we treat each travel mode as a variable, then the hourly trip time series across different spatial areas forms a tensor, referring to

[4]The MTA subway hourly ridership dataset is available at https://data.ny.gov/Transportation/MTA-Subway-Hourly-Ridership-2020-2024/wujg-7c2s/about_data.

[5]The NYC Citi Bike system dataset is available at https://s3.amazonaws.com/tripdata/index.html.

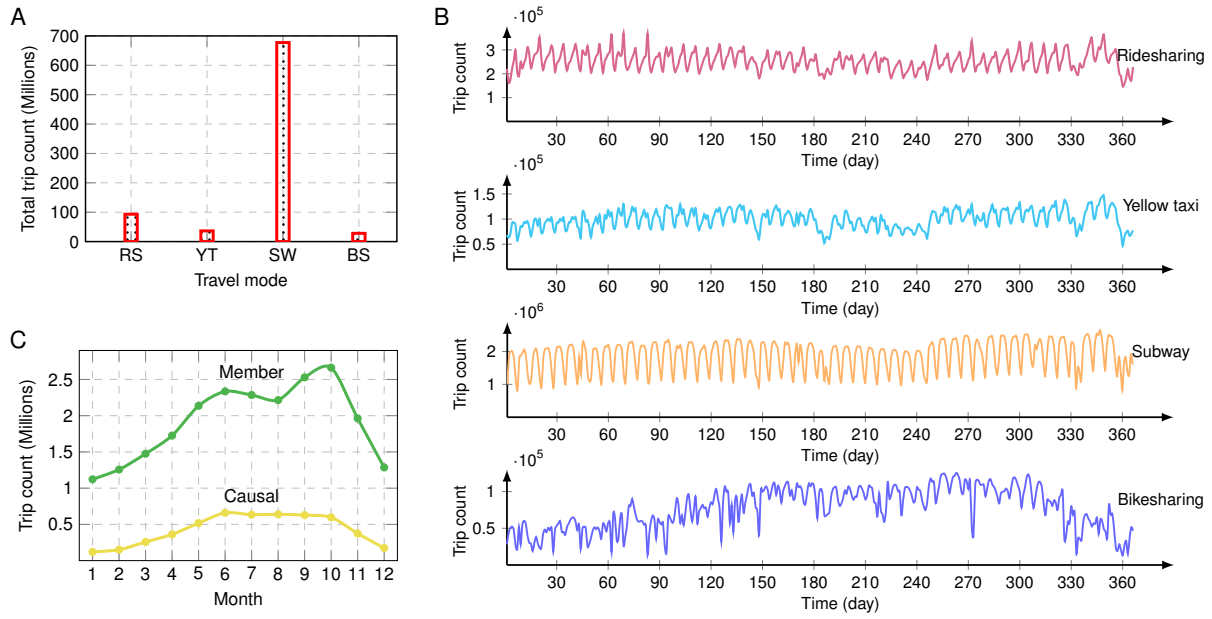

Figure 6: Trip counts of ridesharing (RS), yellow taxi (YT), subway (SW), and bikesharing (BS) during the whole year of 2024 in Manhattan. (A) Bar chart of four travel modes. (B) Trip time series with a daily resolution in Manhattan. (C) Bikesharing trips of membership and causal users across 12 months of 2024.

Figure 1A. Given the distinct travel purposes and different mobility patterns of bikesharing membership and causal travels, the variable dimension of the mobility tensor includes ridesharing, yellow taxi, subway, bikesharing (member), bikesharing (causal), and bikesharing (all). The resulting data is of size $69 \times 6 \times 8784$, with the temporal dimension spanning all 8,784 hours of 2024. By using the multidimensional sparse autoregression method with the sparsity level $\tau = 6$, the index set is optimized as $\Omega = \{1, 23, 24, 144, 167, 168\}$ where the time lag $k = 168$ corresponds to a weekly cycle. The auto-correlations at the time lag $k = 168$ quantifies the weekly periodicity of trip time series across different spatial areas and travel modes. Figures 7C–D bring the following findings and insights:

- We first discuss the weekly periodicity of four travel modes in 2024 (Figure 7C). Subway trips are more periodic than other travel modes. Although the trip time series of

ridesharing and yellow taxi have the same level of weekly periodicity, ridesharing trips are less periodic in the south areas and more periodic in the north areas of Manhattan than yellow taxi trips. In the bikesharing system, the comparison between membership and causal travels clearly demonstrates more periodic patterns of membership travels. Although the total bikesharing trip count is less than ridesharing and yellow taxi trips, the overall weekly periodicity of bikesharing is stronger than the latter.

- We then discuss the daily periodicity of four travel modes during weekdays in 2024 (Figure 7D). As we removed the trip data of weekends, the size of mobility tensor becomes $69 \times 6 \times 6288$ with the temporal dimension spanning 262 weekdays in 2024. By setting the sparsity level of our method as $\tau = 6$, then the index set is optimized as $\Omega = \{1, 23, 24, 95, 96, 120\}$ where the time lag $k = 24$ corresponds to a daily cycle. Observing the pattern of daily periodicity, subway trips are still more periodic than other travel modes. The daily periodicity of subway trips in the north areas is stronger than the south areas of Manhattan. In contrast to the weekly periodicity as shown in Figure 7C, yellow taxi trips are more periodic than ridesharing trips at a daily cycle. During weekdays, the difference of daily periodicity between membership trips and all trips in bikesharing becomes marginal.

As mentioned in Section 5, we find that the weekly periodicity of ridesharing trips in NYC and Chicago shows a remarkable variability across different years. To examine time-varying mobility periodicity patterns, we analyze bimonthly trip datasets of 2024 and quantify the weekly periodicity of four different travel modes in Manhattan. As reported in Table 2, weekly periodicity in the early and late months of 2024 is generally lower than other months. In particular, weekly periodicity in November and December declines remarkably, possibly due to winter weather conditions and holiday-related disruptions. In terms of subway trips, the weekly periodicity shows a significant reduction during the early and late months in

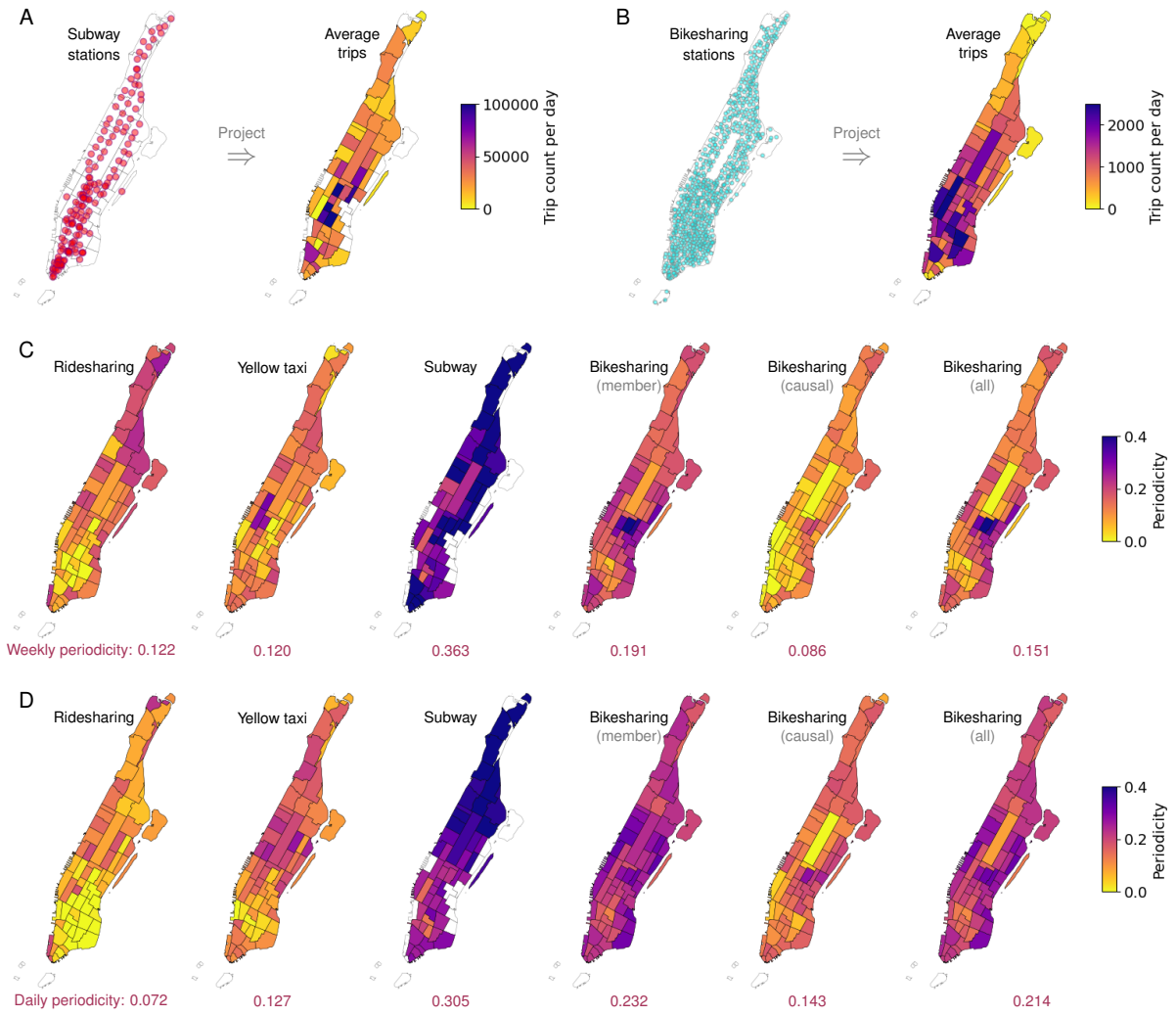

Figure 7: Periodicity in multimodal mobility trip data of 2024 in Manhattan with the multidimensional sparse autoregression method. (A-B) Subway and bikesharing stations are projected onto spatial areas in Manhattan. (C) Weekly periodicity of trip time series across different travel modes. (D) Daily periodicity in weekday's multimodal trip data.

2024. In terms of bikesharing trips, the time series demonstrates the greatest seasonal sensitivity, with notably smaller weekly periodicity in colder months such as January–April and November–December than other months. From May to October, the weekly periodicity of bikesharing trips is even stronger than ridesharing and yellow taxi trips.

Table 2: Weekly periodicity of bimonthly trip data of four travel modes in Manhattan.

	Ridesharing	Yellow taxi	Subway	Bikesharing (member)
January & February	0.142	0.144	0.231	0.131
March & April	0.191	0.198	0.449	0.134
May & June	0.152	0.143	0.289	0.216
July & August	0.183	0.133	0.408	0.219
September & October	0.205	0.195	0.415	0.290
November & December	0.046	0.055	0.136	0.120

7 Discussion

Understanding the daily and weekly periodicity of human mobility is critical for revealing the underlying structure of individual and collective movement patterns in urban areas. The periodicity patterns underlying human mobility ranging from daily commutes to weekly rituals are manifestations of deeply rooted behavioral rhythms. These patterns are essential components for making informed decisions, developing effective urban policies, and building resilient and responsive urban systems. In literature, such rhythms and temporal regularities have been studied in behavioral sciences and temporal network theory (Piccardi et al. 2024, Holme & Saramäki 2012). Drawing on these insights, we develop a multidimensional sparse autoregression method for quantifying and leveraging periodicity in large-scale mobility data. This enables a deeper understanding of urban rhythms and enhances predictive modeling for transportation planning, epidemic spread (Balcan et al. 2009, Belik et al. 2011, Castillo-Chavez et al. 2016), and urban resilience assessment (Xu et al. 2025).

Our findings discovered from extensive human mobility datasets advance people's understanding of real-world urban systems and show the capability of multidimensional sparse autoregression for quantifying periodicity from complicated trip time series. Our findings

concerning the weekly periodicity of ridesharing trip data in NYC and Chicago from 2019 to 2024 demonstrate the disruptive impact of the pandemic on mobility regularity and subsequent recovery trends. In Manhattan, the patterns of daily and weekly periodicity across four different travel modes, including ridesharing, yellow taxi, subway, and bikesharing, demonstrate the behavioral rhythm and temporal regularities across different spatial areas and bimonthly time phases. The interpretability offered by the multidimensional sparse autoregression method is a key advantage, as sparse and non-negative auto-correlations provide insights into the automatic learning of dominant time lags driving the observed periodicity. The method we present allows substantial analysis of time series data collected from real-world systems with inherent periodicity. In addition, the dominant auto-correlations in the index set Ω enable us to introduce proper differencing operations and address the non-stationary time series issue (Chen, Digalakis Jr, Ding, Zhuang & Zhao 2025). To summarize, this work provides a robust foundation for quantifying temporal regularities and their patterns in complicated data and systems beyond human mobility data.

References

Alessandretti, L., Aslak, U. & Lehmann, S. (2020), 'The scales of human mobility', *Nature* **587**(7834), 402–407.

Balcan, D., Colizza, V., Gonçalves, B., Hu, H., Ramasco, J. J. & Vespignani, A. (2009), 'Multiscale mobility networks and the spatial spreading of infectious diseases', *Proceedings of the National Academy of Sciences* **106**(51), 21484–21489.

Basu, S. & Michailidis, G. (2015), 'Regularized estimation in sparse high-dimensional time series models', *The Annals of Statistics* pp. 1535–1567.

Basu, S., Shojaie, A. & Michailidis, G. (2015), 'Network granger causality with inherent

grouping structure', *The Journal of Machine Learning Research* **16**(1), 417–453.

Belik, V., Geisel, T. & Brockmann, D. (2011), 'Natural human mobility patterns and spatial spread of infectious diseases', *Physical Review X* **1**(1), 011001.

Bertsimas, D., Digalakis Jr, V., Li, M. L. & Lami, O. S. (2024), 'Slowly varying regression under sparsity', *Operations Research* .

Bertsimas, D., King, A. & Mazumder, R. (2016), 'Best subset selection via a modern optimization lens', *The Annals of Statistics* **44**(2), 813–852.

Bertsimas, D., Pauphilet, J. & Van Parys, B. (2020), 'Sparse regression', *Statistical Science* **35**(4), 555–578.

Brunton, S. L. & Kutz, J. N. (2022), *Data-driven science and engineering: Machine learning, dynamical systems, and control*, Cambridge University Press.

Brunton, S. L., Proctor, J. L. & Kutz, J. N. (2016), 'Discovering governing equations from data by sparse identification of nonlinear dynamical systems', *Proceedings of the National Academy of Sciences* **113**(15), 3932–3937.

Castillo-Chavez, C., Bichara, D. & Morin, B. R. (2016), 'Perspectives on the role of mobility, behavior, and time scales in the spread of diseases', *Proceedings of the National Academy of Sciences* **113**(51), 14582–14588.

Chen, X., Cai, H., Liu, F. & Zhao, J. (2025), 'Correlating time series with interpretable convolutional kernels', *IEEE Transactions on Knowledge and Data Engineering* **37**(6), 3272–3283.

Chen, X., Digalakis Jr, V., Ding, L., Zhuang, D. & Zhao, J. (2025), 'Interpretable time series autoregression for periodicity quantification', *arXiv preprint arXiv:2506.22895* .

Chen, X. & Sun, L. (2022), 'Bayesian temporal factorization for multidimensional time

series prediction', *IEEE Transactions on Pattern Analysis and Machine Intelligence* **44**(9), 4659–4673.

Çolak, S., Lima, A. & González, M. C. (2016), 'Understanding congested travel in urban areas', *Nature Communications* **7**(1), 10793.

Davis, R. A., Zang, P. & Zheng, T. (2016), 'Sparse vector autoregressive modeling', *Journal of Computational and Graphical Statistics* **25**(4), 1077–1096.

Du, Z., Fox, S. J., Holme, P., Liu, J., Galvani, A. P. & Meyers, L. A. (2018), 'Periodicity in movement patterns shapes epidemic risk in urban environments', *arXiv preprint arXiv:1809.05203* .

Gonzalez, M. C., Hidalgo, C. A. & Barabasi, A.-L. (2008), 'Understanding individual human mobility patterns', *Nature* **453**(7196), 779–782.

Goulet-Langlois, G., Koutsopoulos, H. N., Zhao, Z. & Zhao, J. (2017), 'Measuring regularity of individual travel patterns', *IEEE Transactions on Intelligent Transportation Systems* **19**(5), 1583–1592.

Guihaire, V. & Hao, J.-K. (2008), 'Transit network design and scheduling: A global review', *Transportation Research Part A: Policy and Practice* **42**(10), 1251–1273.

Hamilton, J. D. (2020), *Time series analysis*, Princeton university press.

Holme, P. & Saramäki, J. (2012), 'Temporal networks', *Physics Reports* **519**(3), 97–125.

Hyndman, R. J. & Athanasopoulos, G. (2018), *Forecasting: principles and practice*, OTexts.

Jenatton, R., Audibert, J.-Y. & Bach, F. (2011), 'Structured variable selection with sparsity-inducing norms', *The Journal of Machine Learning Research* **12**, 2777–2824.

Jiang, S., Ferreira, J. & Gonzalez, M. C. (2017), 'Activity-based human mobility patterns

inferred from mobile phone data: A case study of singapore', *IEEE Transactions on Big Data* **3**(2), 208–219.

Li, W., Wang, Q., Liu, Y., Small, M. L. & Gao, J. (2022), 'A spatiotemporal decay model of human mobility when facing large-scale crises', *Proceedings of the National Academy of Sciences* **119**(33), e2203042119.

Li, Z., Ding, B., Han, J., Kays, R. & Nye, P. (2010), Mining periodic behaviors for moving objects, *in* 'Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining', pp. 1099–1108.

Manley, E., Zhong, C. & Batty, M. (2018), 'Spatiotemporal variation in travel regularity through transit user profiling', *Transportation* **45**, 703–732.

Murdoch, W. J., Singh, C., Kumbier, K., Abbasi-Asl, R. & Yu, B. (2019), 'Definitions, methods, and applications in interpretable machine learning', *Proceedings of the National Academy of Sciences* **116**(44), 22071–22080.

Piccardi, T., Gerlach, M. & West, R. (2024), Curious rhythms: Temporal regularities of wikipedia consumption, *in* 'Proceedings of the International AAAI Conference on Web and Social Media', Vol. 18, pp. 1249–1261.

Prabhala, B., Wang, J., Deb, B., La Porta, T. & Han, J. (2014), Leveraging periodicity in human mobility for next place prediction, *in* '2014 IEEE Wireless Communications and Networking Conference (WCNC)', IEEE, pp. 2665–2670.

Rudin, C., Chen, C., Chen, Z., Huang, H., Semenova, L. & Zhong, C. (2022), 'Interpretable machine learning: Fundamental principles and 10 grand challenges', *Statistic Surveys* **16**, 1–85.

Simini, F., González, M. C., Maritan, A. & Barabási, A.-L. (2012), 'A universal model for mobility and migration patterns', *Nature* **484**(7392), 96–100.

Song, C., Qu, Z., Blumm, N. & Barabási, A.-L. (2010), 'Limits of predictability in human mobility', *Science* **327**(5968), 1018–1021.

Teixeira, D. D. C., Viana, A. C., Almeida, J. M. & Alvim, M. S. (2021), 'The impact of stationarity, regularity, and context on the predictability of individual human mobility', *ACM Transactions on Spatial Algorithms and Systems* **7**(4), 1–24.

Tibshirani, R. (1996), 'Regression shrinkage and selection via the lasso', *Journal of the Royal Statistical Society Series B: Statistical Methodology* **58**(1), 267–288.

Tillmann, A. M., Bienstock, D., Lodi, A. & Schwartz, A. (2024), 'Cardinality minimization, constraints, and regularization: a survey', *SIAM Review* **66**(3), 403–477.

Tirachini, A. (2020), 'Ride-hailing, travel behaviour and sustainable mobility: an international review', *Transportation* **47**(4), 2011–2047.

Valdés-Sosa, P. A., Sánchez-Bornot, J. M., Lage-Castellanos, A., Vega-Hernández, M., Bosch-Bayard, J., Melie-García, L. & Canales-Rodríguez, E. (2005), 'Estimating brain functional connectivity with sparse multivariate autoregression', *Philosophical Transactions of the Royal Society B: Biological Sciences* **360**(1457), 969–981.

Xu, F., Wang, Q., Moro, E., Chen, L., Salazar Miranda, A., González, M. C., Tizzoni, M., Song, C., Ratti, C., Bettencourt, L. et al. (2025), 'Using human mobility data to quantify experienced urban inequalities', *Nature Human Behaviour* pp. 1–11.

Zhang, L. & Song, J. (2022), 'The periodicity and initial evolution of micro-mobility systems: A case study of the docked bike-sharing system in new york city, usa', *European Transport Research Review* **14**(1), 27.