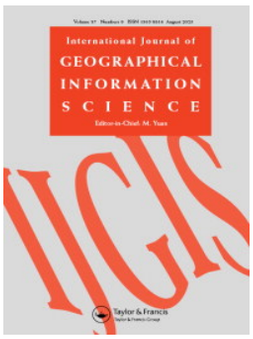


International Journal of Geographical Information Science

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/tgis20

Toward urban traffic scenarios and more: a spatio-temporal analysis empowered low-rank tensor completion method for data imputation

Zilong Zhao, Luliang Tang, Mengyuan Fang, Xue Yang, Chaokui Li & Qingquan Li

To cite this article: Zilong Zhao, Luliang Tang, Mengyuan Fang, Xue Yang, Chaokui Li & Qingquan Li (2023): Toward urban traffic scenarios and more: a spatio-temporal analysis empowered low-rank tensor completion method for data imputation, International Journal of Geographical Information Science, DOI: 10.1080/13658816.2023.2234434

To link to this article: https://doi.org/10.1080/13658816.2023.2234434

Published online: 19 Jul 2023.

INTERNATIONAL JOURNAL OF GEOGRAPHICAL INFORMATION SCIENCE
https://doi.org/10.1080/13658816.2023.2234434

Taylor & Francis
Taylor & Francis Group

RESEARCH ARTICLE

Toward urban traffic scenarios and more: a spatio-temporal analysis empowered low-rank tensor completion method for data imputation

Zilong Zhao[a], Luliang Tang[a], Mengyuan Fang[a], Xue Yang[b], Chaokui Li[c] and Qingquan Li[d]

[a]State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan, China; [b]School of Geography and Information Engineering, China University of Geosciences, Wuhan, China; [c]National-Local Joint Engineering Laboratory of Geo-Spatial Information Technology, Hunan University of Science and Technology, Xiangtan, China; [d]Guangdong Key Laboratory of Urban Informatics, Shenzhen University, Shenzhen, China

ABSTRACT

Existing traffic monitoring approaches cannot completely cover all road segments in real-time, leading to massive amounts of missing traffic data, which limits the implementation of intelligent transportation systems. Most existing methods lack deep mining of the unique spatiotemporal characteristics of traffic flows, resulting in difficulty in application to urban traffic with complex topologies and variable states. In this paper, we propose a novel Spatio-Temporal constrained Low-Rank Tensor Completion (ST-LRTC) method, which adopts a manifold embedding approach to depict the local geometric structure of spatiotemporal domains. Specifically, under the low-rank assumption, the method introduces temporal constraints based on the continuity and periodicity of traffic flow and a spatial constraint matrix reflecting the traffic flow transmission mechanism. We embed low-dimensional spatiotemporal constraint matrices into the low-rank tensor completion solving process to fully utilize the global features and local spatiotemporal characteristics of the traffic tensor. Experiments were performed using traffic data from Xi'an, China, and the results indicated that ST-LRTC outperformed state-of-the-art methods under various missing rates and patterns. Thorough experiments have demonstrated that the incorporation of spatiotemporal analysis can enhance the adaptability of the tensor completion model to complex urban scenarios, which guarantees better monitoring, diagnosis, and optimization of urban traffic states.

ARTICLE HISTORY
Received 12 July 2022
Accepted 4 July 2023

KEYWORDS
Geographical information system for transportation (GIS-T); missing data imputation; low-rank tensor completion; spatial-temporal analysis; urban traffic scenarios

1. Introduction

Accurately monitoring traffic states is the foundation of Intelligent Transportation Systems (ITS), supporting both travelers to plan the optimal routes and policymakers to develop effective strategies to improve traffic operating efficiency (Chen *et al.* 2020,

CONTACT Luliang Tang tll@whu.edu.cn

Wang *et al*. 2019). However, due to the limitations of installation cost, uneven spatial distribution, and unstable equipment operation status, existing traffic monitoring approaches (including loop detectors (Chen *et al*. 2003), microwave sensors (Rodrigues *et al*. 2011), road-side cameras (Douxchamps *et al*. 2006), and floating-cars (Li *et al*. 2011a)) are unable to cover all the road segments in real-time. As a result, there are massive missing values within the obtained traffic states (Yuebiao *et al*. 2014, Wang *et al*. 2019). According to the Texas Transportation Institute, the rate of missing data often reaches more than 50% and even exceeds 90% at times (Smith *et al*. 2003, Tan *et al*. 2013), which seriously restricts the development of smart cities and ITS, such as traffic forecasting (Fang *et al*. 2022), movement pattern analysis (Dodge *et al*. 2020), travel time estimation (Tang *et al*. 2016), and anomaly detection (Sofuoglu and Aviyente 2022).

Therefore, the accurate and efficient reconstruction of missing data in urban road networks to support monitoring, diagnosis, and optimization of the entire urban traffic states has become the focus of urban perception and transportation research.

Theoretically, a high-dimensional tensor is a desirable representation of spatiotemporal field data with multi-dimensionality, such as traffic speed data (Li *et al*. 2022). Based on this, tensor operators (e.g. tensor decomposition and tensor reconstruction) can reveal the intrinsic structure of complex spatiotemporal data and capture interdependencies in multiple dimensions. Accordingly, in this study, we represent traffic states (e.g. traffic speed and traffic flow) with a tensor $\mathcal{X}^{Road \times Time \times Day}$ to highlight different dimensional properties, thus facilitating accurate traffic data imputation.

Furthermore, owing to the cyclical nature of travel behavior and the interconnection of neighboring road segments, traffic states are highly correlated in the temporal and spatial domains (Wang *et al*. 2019, Feng *et al*. 2022). Many scholars have gradually focused on the low-rank characteristics of traffic data and have attempted to estimate missing data using compressed sensing theories, such as low-rank tensor decomposition (Chen *et al*. 2020, Ji *et al*. 2012) and tensor principal component analysis (Feng *et al*. 2022).

However, most existing low-rank tensor-based methods estimate values from a data-driven perspective by exploiting only the implicit regularities of the data. These methods do not consider the spatiotemporal characteristics of geographic elements, which makes their application in real scenarios problematic. Given the complexity of urban scenes and the applicability of low-rank tensors, critical challenges exist in traffic state imputation.

- Accurate estimation. Traffic data are usually disturbed by noise, which affects the mining of the spatiotemporal features of traffic states. However, the development of accurate and robust imputation models remains an unexplored avenue.
- Urban Scenarios. Most existing methods do not consider the structural properties of road networks or the evolution patterns of traffic flows, making their application to complex urban road scenarios difficult. Hence, fully exploring the spatiotemporal correlation characteristics of traffic states and developing a model that is more applicable to urban traffic data imputation is still a research hotspot.

- Extreme absence. In practice, most cities deploy sensors only on a few critical roadways. Even worse, even in this part of the data, there are inevitable sparse or missing data owing to collection imperfections, equipment failures, etc. Accordingly, traffic data imputation in extreme missing cases, that is, high missing rates and nonrandom absences, is crucial for fine-grained urban perception.

Intuitively, the incorporation of geographic methods and thinking endows tensor models with explicit representations of spatiotemporal patterns, thereby facilitating the practical application of high-dimensional tensor theory in real scenarios, such as urban perception and geospatial analysis. Here, we propose a temporally continuous, spatially regularized, and more compact low-rank tensor model, named Spatio-Temporal constrained Low-Rank Tensor Completion (ST-LRTC), for urban traffic speed data imputation. The main contributions of the proposed method are as follows:

- We adopted a low-rank tensor model to mine the global spatiotemporal properties of traffic data. Inspired by CHEN *et al*. (2020), we introduced a truncated nuclear norm as a tighter nonconvex alternative to the tensor low-rank function to achieve more accurate estimates of urban traffic states.
- Based on the low-rank assumption, we adopted a manifold embedding approach to depict the local geometric structure across each mode and propose a K-order proximity-based 'static road topology & dynamic traffic flow' co-driven approach for road network spatial similarity metrics. Spatiotemporal constraints explicitly express the continuity, periodicity, and transitivity of traffic flow, which can provide prior knowledge of the reconstruction process of missing data to enhance reliability.
- Under extreme scenarios with large missing rates or nonrandom missing patterns, the proposed method can still achieve stable and accurate imputation results, illustrating the applicability and robustness of ST-LRTC in complex road networks.

The remainder of this paper is organized as follows. In Section 2, we provide a literature review of traffic data imputation. Preliminaries are provided in Section 3. Section 4 introduces the key aspects of the ST-LRTC method. Section 5 evaluates performance using a real traffic dataset. Finally, Section 6 concludes the study and provides directions for future research.

2. Related works

In recent years, compressed sensing-based sparse representation and its derivative, Low-Rank Matrix Recovery (LRMR), have made significant progress in theoretical studies, making sparse representations more popular (Donoho 2006, Candes and Wakin 2008).

Owing to the daily life patterns and road network settings, traffic data tend to be continuous, periodic, and spatially similar (Zhang *et al*. 2020). Therefore, traffic flow data usually have low rank and sparse properties. Low-Rank Matrix Completion (LRMC) exploits a low-rank structure to recover the missing data and uses convex relaxation

iterations (e.g. singular value thresholding algorithm) to optimize the solution. In addition, some scholars have incorporated Bayesian models into LRMC and developed Bayesian Principal Component Analysis (BPCA) (Li *et al*. 2008), Probabilistic Principal Component Analysis (PPCA) (Qu *et al*. 2009), and Kernel Probabilistic Principal Component Analysis (KPPCA) (Li *et al*. 2013) to recover and reconstruct missing values of matrices, and have obtained more accurate results at low missing rates.

However, the aforementioned matrix-based traffic data imputation methods tend to operate on a two-dimensional traffic matrix with stacked columns, which results in the natural multidimensional nature of traffic data being neglected. Consequently, matrix completion methods always appear powerless when the missing rate is higher than 60% (Tan *et al*. 2013).

To fully exploit the implicit information of multidimensional data, tensor representations for data recovery have recently been investigated. Indeed, a tensor can encompass more global information compared to a matrix such as an additional third dimension representing time (Said and Erradi 2021). In the recent literature, the successful recovery of tensor completion mainly relies on its low-rank assumption (Long *et al*. 2019). These methods fall into two main categories: tensor-decomposition-based methods and minimum-rank-based methods.

2.1. Tensor decomposition-based *methods*

Methods based on tensor decomposition always need to set a predefined rank and then optimize the decomposed feature factors. CANDECOMP/PARAFAC (CP) decomposition (Carroll and Chang 1970, Harshman 1970) and Tucker decomposition (Tucker 1966) are two representative tensor decomposition methods. Acar *et al*. (2011) proposed a CP-weighted optimization algorithm with first-order optimization to solve a weighted least-squares problem. Tan *et al*. (2013) constructed traffic data as a four-dimensional tensor and used the Tucker decomposition method for tensor completion. The superiority of the high-dimensional data-based form is confirmed by comparing its results with those of a two-dimensional matrix. Goulart *et al*. (2017) proposed a tensor completion algorithm based on Tucker decomposition kernel tensor iterative soft thresholding, considering the strong correlation of traffic data.

To better capture the spatiotemporal correlation of traffic data to distinguish it from other types of tensor imputation (e.g. image processing, computer vision, recommender systems), Zhou *et al*. (2015) combined the factor matrices of CP decomposition with spatiotemporal intramodal regularization, and proposed a spatiotemporal tensor completion method to recover missing traffic data. Said and Erradi (2021) computed an urban similarity matrix and incorporated feature vectors for each region into the factors of the CP decomposition, in order to explicitly mine the intrinsic patterns of traffic data.

Nevertheless, because the above models are nonconvex, they are subject to local minima and depend on good initial values to perform well. Furthermore, in practice, the tensor-rank bounds may not be available for certain applications. When only a few observations are available, the selection of inappropriate rank bounds may lead to underfitting or overfitting (Long *et al*. 2019).

2.2. Overall minimum rank-based methods

The latter takes a global perspective and exploits the low-rank nature of traffic data to determine the minimum rank in the overall structure. This avoids the selection and analysis of the rank of individual feature factors and retains the inherent spatiotemporal characteristics of traffic data. However, because the rank function is nonconvex, the rank minimization problem is NP-hard (Hillar and Lim 2013).

Most existing methods (Ji *et al*. 2012, Hu and Work 2021, Sofuoglu and Aviyente 2022, Feng *et al*. 2022) used the nuclear norm (NN) as a convex surrogate for the nonconvex rank function and relax the problem of finding the minimum rank to minimize the tensor nuclear norm when capturing the low-rank properties of spatiotemporal data. Experiments have shown that these methods can achieve better results than CP decomposition (Ji *et al*. 2012).

Although NN-based methods have achieved impressive results in multidimensional data analysis, they have some limitations. Considering that the rank operator is nonconvex and discontinuous, the convex nuclear norm may not serve as a good approximation of the rank operator. Specifically, when seeking tensor rank minimization, the NN explicitly considers each singular value equally and shrinks all singular values using the same parameters. Nevertheless, in practical applications, larger singular values are generally associated with salient spatiotemporal information in urban traffic patterns and are therefore more important. Thus, the larger singular values should shrink less to ensure that the essential traffic patterns are not corrupted (Cao *et al*. 2017).

To address this problem, many recent studies have begun to utilize certain nonconvex functions to approximate the rank function, one of which is the truncated nuclear norm (TNN) (Hu *et al*. 2013). Unlike the nuclear norm, the TNN method minimizes only the smallest (min(m, n) − r) singular values to retain the prominent intrinsic features of the data (Zhang *et al*. 2012). For instance, Chen *et al*. (2020) replaced the original convex nuclear norm with a TNN and proposed an LRTC-TNN model. Experiments demonstrated that the tighter substitution function can better capture the low-rank nature of the traffic tensor, which yields a more desirable imputation accuracy than the convex substitution.

2.3. Research gap

Although current low-rank tensor completion models have achieved ideal results in some cases (e.g. random missing cases and low missing-rate cases), they still face difficulties when applied to real road scenarios.

1. Most methods are limited to extracting information from the tensor itself, and are more inclined to estimate values from a data-driven perspective. They lack deep mining of traffic spatiotemporal features (e.g. road network topology and traffic flow evolution patterns), which are crucial for urban traffic data imputation with complex structures and variable states.
2. In the bad case with large missing rates or non-random missing data (e.g. spatially missing and temporally missing data), the implicit regularity of urban traffic is difficult to obtain accurately and completely from the limited observations by

existing models, leading to unstable results or even failed reconstruction (Tan *et al*. 2013).

Hence, the establishment of data imputation models that can be adapted to complex urban scenarios and extreme missing cases is of great importance for the development of ITS. Motivated by this intuition, the Spatio-Temporal constrained Low-Rank Tensor Completion (ST-LRTC) is proposed in this paper and applied to urban traffic speed data imputation.

3. Preliminaries

Throughout this study, we inherited the tensor definition as in Kolda *et al*. (2009). Scalars vectors and matrices are denoted by lowercase letters, bold lowercase and uppercase letters respectively, e.g. x, $\boldsymbol{x} \in \boldsymbol{R}^n$, $\boldsymbol{X} \in \boldsymbol{R}^{m\times n}$. Tensors are denoted by bold Eulerian letters, e.g. $\boldsymbol{\mathcal{X}} \in \boldsymbol{R}^{m\times n\times t}$.

Definition 1 (Tensor construction): Assuming that there are M road segments, D days, and T time intervals per day, the traffic state data can be constructed as a tensor $\boldsymbol{\mathcal{X}} \in \mathbb{R}^{M\times T\times D}$ in three dimensions (M road segments – T time intervals/day – D days), as shown in Figure 1. The element value of $\mathcal{X}_{m,t,d}$ denotes the traffic state of the m-th road segment at time interval t on day d.

Definition 2 (Tensor unfolding and folding): Matricization (unfolding) is the process of reordering the elements of an N-dimensional tensor into a matrix. The unfolding of a tensor $\mathcal{X} \in \mathbb{R}^{I_1\times I_2\times\cdots\times I_N}$ in mode-n can be expressed as $\mathcal{X}_{(n)}$, where the tensor element (i_1, i_2, …, i_N) is mapped to the matrix element (i_n, j). Essentially, the mode-n unfolding operator arranges the n-th mode of $\mathcal{X}$ as a row (i.e. i_n), whereas the remaining modes are the columns of the mode-n unfolding matrix (i.e. j). Mathematically, i_n is the ordinal number of the elements in the unfolded mode, and j satisfies:

$$j = i_1 + \sum_{\substack{k=1\\ k\neq n}}^{N} (i_k - 1)J_k \text{ with } J_k = \prod_{\substack{m=1\\ m\neq n}}^{k-1} I_m, \tag{1}$$

where I_m denotes the maximum index of the m-th dimension. For a more intuitive explanation, we present an example of a three-dimensional tensor unfolding, as shown in Figure 2.

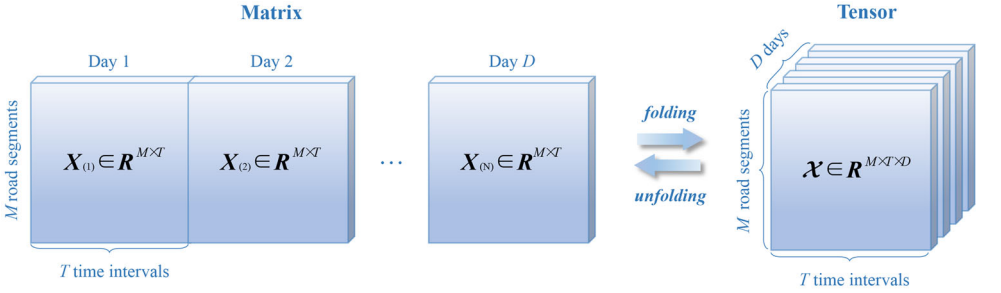


Figure 1. Tensor construction and transformation.

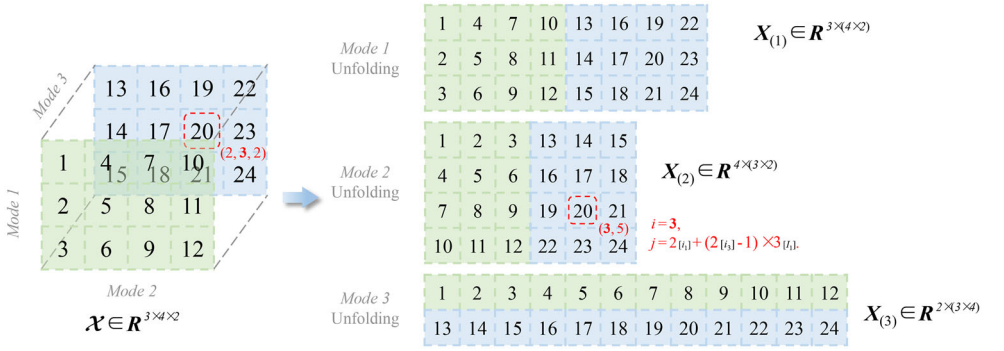


Figure 2. Example of three-dimensional tensor unfolding.

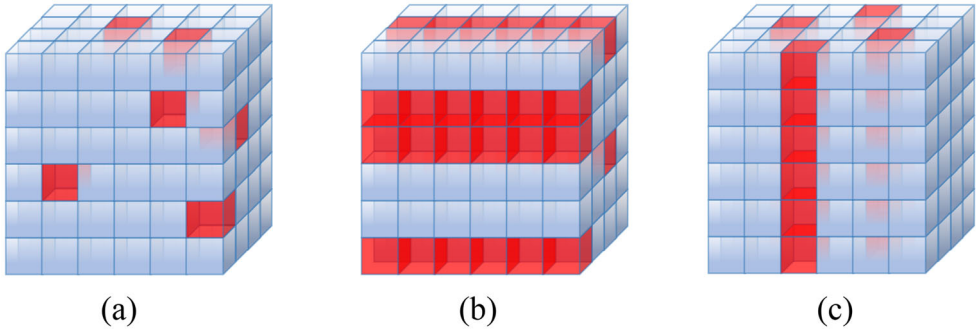


Figure 3. Different missing patterns of traffic data. (a) Random missing pattern. (b) Spatial missing pattern. (c) Temporal missing pattern.

Similarly, we define a folding operator that converts a matrix into a higher-order tensor in the *n*th-mode as $fold_n(\cdot)$. Thus, we have $fold_n(\mathcal{X}_{(n)}) = \mathcal{X}$.

Definition 3 (Data missing patterns): In real-world scenarios, missing data may not strictly adhere to a uniform and random distribution but rather be structurally missing because of sensor malfunctions or communication failures (Deng *et al.* 2021, Qu *et al.* 2009, Wang *et al.* 2019). Therefore, we divide the cases of missing data in complex urban scenarios into three categories, as shown in Figure 3.

- Random Missing Pattern (RM): In most cases, the loss of traffic data is random and uniform. We simulate this situation by randomly dropping certain elements of the traffic tensor.
- Spatial Missing Pattern (SM): Circumstances such as power shortages may result in the loss of traffic data from sensors. This situation is simulated by dropping data from multiple consecutive road segments.
- Temporal Missing Pattern (TM): Traffic data may be lost for some time due to damage to the collection device or the disconnection of the wireless network. Therefore, we randomly select sensors and drop the data from each selected sensor at some consecutive time intervals to simulate this missing pattern.

Definition 4 (Tensor imputation): also known as tensor completion. The purpose of tensor imputation is to reconstruct the missing values using the observed data tensor. Because traffic data are partially missing and only a subset of the tensor provides valid traffic data (Ji *et al*. 2012, Tan *et al*. 2013, Chen *et al*. 2020), we introduce a mask operator to represent the part of the tensor where valid data exist, as in Equation (2):

$$\mathcal{P}_\Omega(\boldsymbol{\mathcal{X}}) = \mathcal{P}_\Omega(\boldsymbol{\mathcal{Z}}), \tag{2}$$

where $\boldsymbol{\mathcal{X}}, \boldsymbol{\mathcal{Z}} \in \mathbb{R}^{M\times T\times D}$ denotes the imputed and the original tensor data, respectively; Ω represents the subset of valid observations. The operator $\mathcal{P}_\Omega : \mathbb{R}^{M\times T\times D} \mapsto \mathbb{R}^{M\times T\times D}$ is an orthogonal projection supported on Ω, which represents the valid observations of $\boldsymbol{\mathcal{X}}$:

$$[P_\Omega(\boldsymbol{\mathcal{X}})]_{mtd} = \begin{cases} \boldsymbol{\mathcal{X}}_{mtd}, & \text{if } (m,t,d) \in \Omega \\ 0, & \text{otherwise} \end{cases}. \tag{3}$$

4. Method

4.1. Low-rank tensor completion

Because of the cyclical nature of travel behavior and the interconnection of neighboring road segments, traffic data are highly correlated in the temporal and spatial domains (Li *et al*. 2011b, Wang *et al*. 2019). Therefore, the tensor tends to be low-rank or approximately low-rank (Chen *et al*. 2020, Ji *et al*. 2012), making the data imputation problem a minimum-rank-seeking problem, as shown in Equation (4),

$$\begin{aligned} &\min \quad rank(\boldsymbol{\mathcal{X}}) \\ &s.t.\ \ \mathcal{P}_\Omega(\boldsymbol{\mathcal{X}}) = \mathcal{P}_\Omega(\boldsymbol{\mathcal{Z}}) \end{aligned}. \tag{4}$$

where $\boldsymbol{\mathcal{X}}, \boldsymbol{\mathcal{Z}} \in \mathbb{R}^{M\times T\times D}$ denotes the imputed and the original tensor data, respectively.

However, the optimization problem in Equation (4) is NP-hard because the function *rank* ($\boldsymbol{\mathcal{X}}$) is nonconvex. To address this issue, researchers have developed several convex relaxation methods and nonconvex functions as alternatives to the rank minimization problem (Hillar and Lim 2013, Chen *et al*. 2020, Ji *et al*. 2012).

The truncated nuclear norm, as a representative nonconvex substitution, can effectively retain the prominent intrinsic features of the data and better capture the low-rank characteristics of the traffic tensor (Chen *et al*. 2020). Thus, in what follows, we use a truncated nuclear norm relaxation method to yield tight estimates of traffic data with low-rank properties (Hu *et al*. 2013, Han *et al*. 2017, Huang *et al*. 2014).

Definition 5 (Truncated nuclear norm, TNN): Given a matrix $\boldsymbol{X} \in \boldsymbol{R}^{m\times n}$, m and n are non-negative integers. Suppose $r \leq min\ (m, n)$, TNN $\|X\|_{r,\ *}$ is defined as the sum of $min\ (m, n) - r$ minimum singular values, that is,

$$\|\boldsymbol{X}\|_{r,*} = \sum_{i=r+1}^{\min\{m,n\}} \sigma_i(\boldsymbol{X}), \tag{5}$$

where $\sigma_i\ (\boldsymbol{X})$ is the i-th singular value of matrix $\boldsymbol{X}$, sorted in descending order.

To this end, the optimization problem in Equation (4) is eventually expressed as a low-rank tensor completion based on minimization (Chen *et al*. 2020), as shown in Equation (6):

$$\begin{aligned} &\min_{\mathcal{X}} \quad \sum_{i=1}^{N} \alpha_i \left\| \mathcal{X}_{(i)} \right\|_{r_i,*}, \\ &s.t.\ \mathcal{P}_{\Omega}(\mathcal{X}) = \mathcal{P}_{\Omega}(\mathcal{Z}) \end{aligned} \tag{6}$$

where i denotes the dimension index of the tensor $\mathcal{X}$, and N denotes the total number of dimensions. α_i s are constants satisfying $\alpha_i \geq 0$ and $\sum_{i=1}^{N} \alpha_i = 1$. The truncation r_i for each tensor mode satisfies:

$$r_i = \left\lceil \theta \cdot \min\left\{ n_i, \prod_{k \neq i} n_k \right\} \right\rceil, \forall i \in \{1, 2, \ldots, N\}, \tag{7}$$

where $\lceil x \rceil$ denotes the smallest integer not less than x, n_i denotes the maximal index of the m-th dimension, and θ controls the proportion of truncation in the different modes of the tensor $\mathcal{X}$.

Essentially, the truncated nuclear norm of a tensor is a combination of the truncated nuclear norm of all matrices expanded in each mode. However, despite the availability of tighter rank substitution functions, these methods fail to yield desirable results, particularly when extreme absences occur. Because low-rank tensor completion methods seek low-rank structures with typical patterns, they struggle to mine intrinsic features from limited data when large missing rates occur, which limits the accuracy of the estimates.

4.2. Spatio-temporal manifold embedding

The spatial and temporal autocorrelative characteristics of traffic flows are crucial for missing data imputation (Zhang *et al*. 2020, Wang *et al*. 2019). Chronologically, traffic flow evolves continuously over time, and the patterns are similar on different days. Spatially, the traffic flow on each road segment is affected by its upstream and downstream segments. Thus, the traffic states of adjacent road segments are interrelated.

Accordingly, based on the low-rank assumption, we utilize the intrinsic relationships of each mode ('road segment', 'time interval', and 'day') of the tensor as auxiliary information to improve the quality of traffic data imputation. This approach is also known as manifold learning, and the main idea is that if two objects are close in the intrinsic geometry of the data manifold, they should be close to each other after dimensionality reduction (Sofuoglu and Aviyente 2022, Li *et al*. 2017).

Specifically, considering that each mode of the tensor corresponds to a different property of traffic data, we retained the relationship between each mode to better mine the intrinsic features. Thus, two similar objects can behave similarly in the projected low-dimensional space by separately introducing spatiotemporal constraint matrices on each mode. In the following, we describe the process of constructing the spatio-temporal constraint matrices.

4.2.1. Temporal constraint matrices

Based on Definition 2, Matrix $\boldsymbol{\mathcal{X}}_{(2)}$ is obtained by unfolding the original tensor $\boldsymbol{\mathcal{X}}$ in the time-interval mode, where the columns represent the traffic states for all moments of the day on a certain road segment. The traffic states at moment t should be similar to the adjacent moments $t-1$ and $t+1$, that is, $\|\boldsymbol{\mathcal{X}}_{(2)}^{t-1}+\boldsymbol{\mathcal{X}}_{(2)}^{t+1}-2\boldsymbol{\mathcal{X}}_{(2)}^{t}\|\rightarrow 0$. In other words, when differentiating the traffic data with two adjacent moments at moment t, the values of the matrix mostly converge to zero, which indicates significant sparsity. Because the L1-norm is a convex approximation of element-wise matrix sparsity and is more robust to noise than the L2-norm (Candès *et al*. 2011), it can be used to better capture sparsity. Thus, we utilize the matrix $\boldsymbol{T_P}$, as shown in Equation (8), to capture the continuity of traffic data in the time interval mode, i.e. $\|\boldsymbol{T}_P\boldsymbol{\mathcal{X}}_{(2)}\|_1$, where $\boldsymbol{T_P}$ = *Toeplitz* (0,1,2,1).

$$\boldsymbol{T}_P=\begin{pmatrix}1 & -2 & 1 & 0 & \cdots\\ 0 & 1 & -2 & 1 & \ddots\\ 0 & 0 & 1 & -2 & \ddots\\ \vdots & \vdots & \ddots & \ddots & \ddots\end{pmatrix}_{(n-2)\times n} \tag{8}$$

Similarly, the traffic flow similarity between adjacent days is described by the matrix $\boldsymbol{T_D}$, i.e. $\|\mathbf{T}_D\boldsymbol{\mathcal{X}}_{(3)}\|_1$, where $\boldsymbol{T_D}$ = *Toeplitz* (0,1,2,1). Note that we utilize the low-rank Toeplitz matrices $\boldsymbol{T_P}$ and $\boldsymbol{T_D}$ to mine the local temporal relations of the urban traffic data. By contrast, employing high-rank Toeplitz matrices tends to make the imputation results oversmooth and lose detailed information of actual traffic states.

4.2.2. Spatial constraint matrix

In fact, the reliance on network topology for the transmission and mobility of traffic elements (e.g. crowds, cabs, and trucks) is one of the most distinctive features that distinguishes urban transportation from other types of data imputation. In previous studies, most scholars adopted the 'node-arc' model to construct the topological relationship of the road network, which usually uses 0/1 to indicate whether the road segments are connected or not (Feng *et al*. 2022, Lin *et al*. 2018). Recently, some researchers propose to construct an undirected local K-connected graph for data $\boldsymbol{X}\in\boldsymbol{R}^{m\times n}$ and encode it using a symmetric affinity matrix $\boldsymbol{A}\in\boldsymbol{R}^{n\times n}$, where $0\leq a_{ij}\leq 1$ reflects the probability that points x_i and x_j are connected (Wang *et al*. 2019). Most affinity matrices $\boldsymbol{A}$ are set by the Euclidean distance between roads (Li *et al*. 2018, Ye *et al*. 2020).

However, such a static representation only reflects the topological structure of the graph, and it is difficult to describe the transmission relationships between traffic flows accurately. Even for geographically adjacent roads, road similarities are influenced by factors such as road grade, convenience, and comfort. Intuitively, people prefer to drive on a spacious and safe road rather than on an isolated narrow path, even though it may be shorter.

To this end, we propose a K-order proximity-based 'static road topology & dynamic traffic flow' co-driven method for spatial similarity measurement of road networks. Laplace regularization is utilized to describe the dynamic correlations between road segments and to sense the effects of abrupt changes in actual traffic conditions

on their upstream and downstream parts. In summary, the introduction of spatial correlation can be divided into three steps: calculation of traffic flow similarity, construction of spatial constraint matrix, and embedding of Laplace regularization (Zhang *et al*. 2020).

4.2.2.1. Calculation of traffic flow similarity. We exploit the dynamic traffic flow data to calculate the transmission coefficient a_{ij} between adjacent roads. First, we normalize the observations into the same range [0,1] using the min-max method to avoid the maximum traffic observation being limited by its physical characteristics (Zhang *et al*. 2020). Thus, the data after min-max normalization can be expressed as:

$$\tilde{x}_{idt} = \frac{x_{idt} - x_i^{\min}}{x_i^{\max} - x_i^{\min}} \tag{9}$$

where x_{idt} denotes the traffic status of road i during time t on day d, and $x_i^{\max}$ and $x_i^{\min}$ are the maximum and minimum observations of road i, respectively.

Furthermore, we adopt a z-score transformation (Yang *et al*. 2017) to remove the daily periodicity of the traffic data and eliminate strong temporal autocorrelation in the time series.

$$\hat{x}_{idt} = \frac{\tilde{x}_{idt} - \bar{x}_{it}}{\sigma_{it}} \tag{10}$$

$$\bar{x}_{it} = \frac{1}{D}\sum_{d=0}^{D-1}\tilde{x}_{idt}, \sigma_{it} = \sqrt{\frac{1}{D-1}\sum_{d=0}^{D-1}\left(\tilde{x}_{idt} - \bar{x}_{it}\right)^2} \tag{11}$$

where $\bar{x}_{it}$ and σ_{it} are the mean and standard deviation of the normalized traffic state of road i at time t, respectively. D is the total number of days.

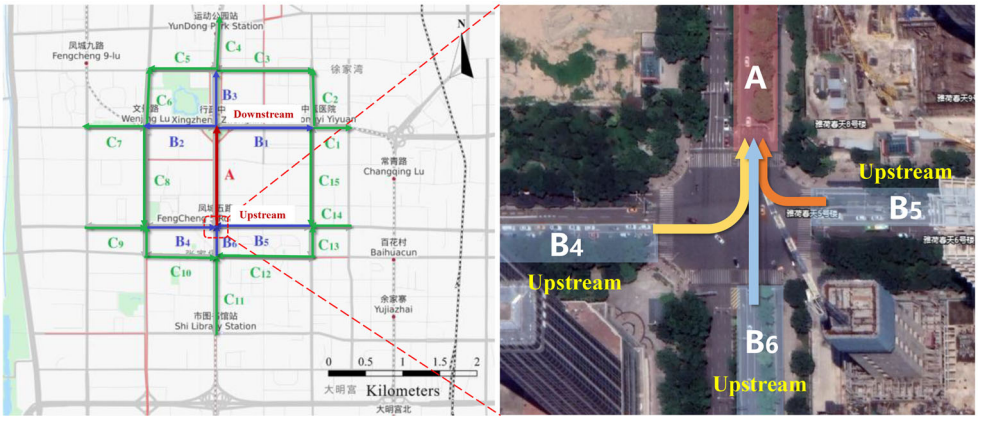
So far, the transmission coefficient a_{ij} between roads i and j can be calculated as follows:

$$a_{ij} = \frac{\sum_{d=0}^{D-1}\sum_{t=0}^{T-1}\left(\hat{x}_{idt} - \bar{x}_i\right)\left(\hat{x}_{jdt} - \bar{x}_j\right)}{\sqrt{\sum_{d=0}^{D-1}\sum_{t=0}^{T-1}\left(\hat{x}_{idt} - \bar{x}_i\right)^2 \sum_{d=0}^{D-1}\sum_{t=0}^{T-1}\left(\hat{x}_{jdt} - \bar{x}_j\right)^2}} \tag{12}$$

where $\bar{x}_i$ denotes the mean of the i-th horizontal slice of tensor $\hat{\mathcal{X}}$. T is the total number of individuals per day. In Equation (12), the numerator denotes the covariance of the observations on roads i and j, and the denominator represents the product of their standard deviations.

4.2.2.2. Construction of spatial constraint matrix. With this, we can obtain the interrelationship of traffic flows on different road segments. In practice, two road segments with a high correlation may not be geographically adjacent, which may cause faulty estimation of traffic states. Instead, the transmission and feedback of traffic flows are primarily based on the road network topology.

We introduce the K-order proximity in graph embedding into the construction of spatial constraints to enhance the adaptability of traffic data imputation under complex topologies and variable states (Yang *et al*. 2018). In Figure 4, for example, the central bi-directional road A is adjacent to the surrounding B1∼B6 roads in the first



(a) (b)

Figure 4. Bi-directional road network topology. (a) Bi-directional road A and its upstream and downstream relationship (b) Turning relationship between road A and its upstream segments (B4, B5, B6).

order, and C1∼C15 roads in the second order. The spatial transferability of traffic flow in bidirectional road networks is based on the turning relationships between roads. As shown in Figure 4(b), roads B4, B5, and B6 are connected to bidirectional road A as its upstream by a left turn, right turn, and straight ahead, respectively. This implies that if traffic congestion occurs on road A, the traffic status of its surrounding roads (including upstream B4∼B6 and downstream B1∼B3) will be affected, and the degree of impact will depend on the similarity of the traffic flow patterns.

Therefore, this study combines the actual road topology with the dynamic traffic flow interrelationship to construct a weighted graph to characterize the spatial relationships, which is expressed as

$$\boldsymbol{A}_{ij} = \begin{cases} a_{ij} \text{ if } j \text{ is a K} - \text{order proximity of } i \\ 0 \text{ otherwise} \end{cases}. \tag{13}$$

4.2.2.3. Embedding of Laplace regularization. By constructing a spatial correlation matrix $\boldsymbol{A}$ based on road topological relationships, we achieve an explicit characterization of the local correlations of traffic states between roads. This enhances the adaptability and robustness of traffic data imputation in complex urban scenarios. Thus, we embed the spatial constraint matrix representing local geographic relationships into the road-mode unfolding matrix $\boldsymbol{\mathcal{X}}_{(1)}$ of low-rank tensor, forming a joint spatial constraint with multi-perspective of 'static & dynamic' and 'global & local'.

Specifically, we utilize the L1-norm to describe the local spatial constraints on the traffic states between roads, i.e.

$$\sum_{i=1}^{N}\sum_{j=1}^{N} a_{ij}\left|x_i - x_j\right| = \left\|\mathbf{L}_{\mathbf{A}}\boldsymbol{\mathcal{X}}_{(1)}\right\|_1 \tag{14}$$

where $\boldsymbol{L}_{\boldsymbol{A}} = \boldsymbol{D} - \boldsymbol{A}$ denotes the Laplace matrix and the degree matrix $\boldsymbol{D}\in\boldsymbol{R}^{N\times N}$ is a diagonal matrix with diagonal elements $d_{ii} = \sum_{j=1}^{N} a_{ij}$, i=0, 1, …, N.

4.3. Spatio-temporal constrained low-rank tensor completion

Although the low-rank tensor-based imputation method can capture some intrinsic characteristics of traffic flow globally, it tends to suffer from information mining biases when handling missing data imputation in complex urban scenarios. Explicitly considering the unique spatiotemporal characteristics of traffic can provide a local perspective on the imputation problem and be more consistent with urban traffic patterns. Meanwhile, the embedding of manifold learning can also provide partial prior knowledge for the imputation model and improve the convergence of iterations and operational efficiency.

To this end, under the assumption of a low-rank tensor, manifold learning is used to improve the quality of data imputation by introducing relationships among data with different mode-unfolding matrices as auxiliary information. In this section, we describe in detail the low-rank tensor completion model under spatiotemporal low-dimensional embedding. The overall framework of the proposed algorithm is given in Figure 5.

In particular, for a given incomplete traffic data $\mathbf{Z}$ we consider the global low-rank property of traffic data and spatiotemporal correlations on the local scale. For the low-rank property of the traffic data tensor, we approximate the rank function using a more compact nonconvex function (TNN). Meanwhile, we adopt the spatiotemporal low-dimensional embedding of manifold learning to capture the local characteristics of traffic flows. Specifically, the temporal constraint matrices $\boldsymbol{T_D}$ and $\boldsymbol{T_P}$ are incorporated into the unfolding matrices to capture the periodicity and continuity, respectively, and the spatial constraint matrix $\boldsymbol{L_A}$ is used to represent the transitivity and feedback of the traffic flow in the bidirectional road network.

In summary, by combining the global low-rank properties (Equation (6)) and the local spatiotemporal correlations, the goal of data imputation is to consider both tensor nuclear norm minimization and spatiotemporal constraints minimization, as in Equation (15),

$$\begin{aligned} \min_{\mathcal{X}} \quad & \sum_{k=1}^{3} \alpha_k \|\boldsymbol{\mathcal{X}}_{(k)}\|_{r_k} + \beta_1 \|\mathbf{L_A}\boldsymbol{\mathcal{X}}_{(1)}\|_1 + \beta_2 \|\mathbf{T_P}\boldsymbol{\mathcal{X}}_{(2)}\|_1 + \beta_3 \|\mathbf{T_D}\boldsymbol{\mathcal{X}}_{(3)}\|_1 \\ & s.t.\ \mathcal{P}_{\Omega}(\boldsymbol{\mathcal{X}}) = \mathcal{P}_{\Omega}(\boldsymbol{\mathcal{Z}}) \end{aligned}, \tag{15}$$

where β_1, β_2, β_3 are non-negative weights balancing the corresponding terms. For brevity, we use $\boldsymbol{S}_k$ ($k = 1,2,3$) to denote the spatial and temporal constraint matrices $\boldsymbol{L_A}$, $\boldsymbol{T_P}$, $\boldsymbol{T_D}$, respectively,

$$\begin{aligned} \min_{\mathcal{X}} \quad & \sum_{k=1}^{3} \alpha_k \|\boldsymbol{\mathcal{X}}_{(k)}\|_{r_k} + \sum_{k=1}^{3} \beta_k \|\mathbf{S}_k\boldsymbol{\mathcal{X}}_{(k)}\|_1 \\ & s.t.\ \mathcal{P}_{\Omega}(\boldsymbol{\mathcal{X}}) = \mathcal{P}_{\Omega}(\boldsymbol{\mathcal{Z}}) \end{aligned}. \tag{16}$$

Furthermore, considering the independence of nuclear norm optimization and the convenience of L1-norm optimization, we introduce auxiliary tensor variables $\mathbf{M}$ and $\tilde{\mathcal{X}}_k$ to avoid the dependence of the TNN solution on spatiotemporally constrained optimization, and express the optimization model as follows:

$$\begin{aligned} \min_{\boldsymbol{\mathcal{X}}_k, \tilde{\mathcal{X}}_k, \mathcal{Q}_k, \boldsymbol{\mathcal{M}}} \quad & \sum_{k=1}^{3} \alpha_k \|\boldsymbol{\mathcal{X}}_{k(k)}\|_{r_k} + \sum_{k=1}^{3} \beta_k \|\boldsymbol{Q}_k\|_1 \\ s.t. \quad & \boldsymbol{\mathcal{X}}_k = \boldsymbol{\mathcal{M}},\ \tilde{\mathcal{X}}_k = \boldsymbol{\mathcal{M}},\ \boldsymbol{Q}_k = \boldsymbol{\mathcal{S}}_k\tilde{\mathcal{X}}_{k(k)}\ \ (k = 1,2,3) \\ & \mathcal{P}_{\Omega}(\boldsymbol{\mathcal{M}}) = \mathcal{P}_{\Omega}(\boldsymbol{\mathcal{Z}}) \end{aligned}. \tag{17}$$

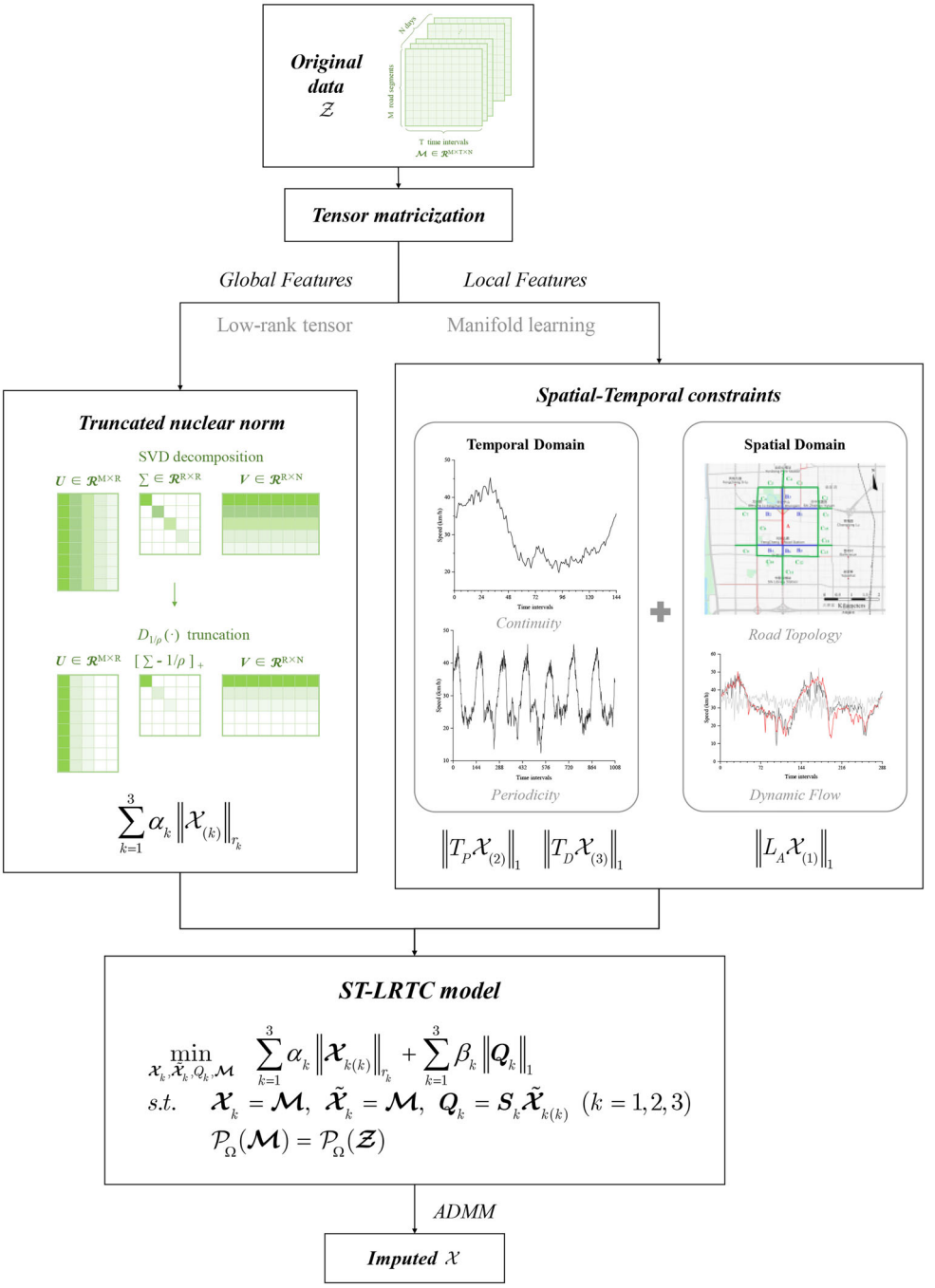


Figure 5. The overall framework of ST-LRTC model for traffic data imputation.

4.4. Solving ST-LRTC for traffic data imputation

A widely used approach to solve this optimization problem of Equation (17) is the Alternating Direction Method of Multipliers (ADMM) framework (Boyd 2010). Specifically, the Lagrangian function of problem (17) is

$$\begin{aligned}\mathcal{L} = &\sum_{k=1}^{3}\alpha_k\|\boldsymbol{\mathcal{X}}_{k(k)}\|_{r_k} + \sum_{k=1}^{3}\beta_k\|\boldsymbol{Q}_k\|_1 \\ &+ \sum_{k=1}^{3}\left(\langle\boldsymbol{\mathcal{X}}_{k(k)} - \boldsymbol{\mathcal{M}}_{(k)}, \boldsymbol{\mathcal{T}}_{k(k)}\rangle + \langle\tilde{\mathcal{X}}_{k(k)} - \boldsymbol{\mathcal{M}}_{(k)}, \boldsymbol{\mathcal{G}}_{k(k)}\rangle + \langle\boldsymbol{Q}_k - \boldsymbol{\mathcal{S}}_k\tilde{\mathcal{X}}_{k(k)}, \ \boldsymbol{Y}_k\rangle\right)\end{aligned} \tag{18}$$

with the definition $\langle\boldsymbol{X}, \boldsymbol{Y}\rangle = tr(\boldsymbol{X}^T\boldsymbol{Y})$. $\boldsymbol{\mathcal{T}}_k, \boldsymbol{\mathcal{G}}_k, \boldsymbol{Y}_k$ $(k = 1,2,3)$ are dual variables.

Based on Equation (18), the following augmented Lagrangian function is provided to facilitate the establishment and solution of the subsequent ADMM, as in Equation (19).

$$\begin{aligned}\mathcal{L}\left(\left[\boldsymbol{\mathcal{X}}_k, \tilde{\mathcal{X}}_k, \mathbf{Q}_k\right]_{k=1}^{3}, \boldsymbol{\mathcal{M}}\right) = &\sum_{k=1}^{3}\alpha_k\|\boldsymbol{\mathcal{X}}_{k(k)}\|_{r_k} + \sum_{k=1}^{3}\beta_k\|\mathbf{Q}_k\|_1 + \sum_{k=1}^{3}\left(\frac{\rho_k}{2}\|\boldsymbol{\mathcal{X}}_{k(k)} - \boldsymbol{\mathcal{M}}_{(k)}\|_F^2\right. \\ &+ \langle\boldsymbol{\mathcal{X}}_{k(k)} - \boldsymbol{\mathcal{M}}_{(k)}, \boldsymbol{\mathcal{T}}_{k(k)}\rangle + \frac{\lambda_k}{2}\|\tilde{\mathcal{X}}_{k(k)} - \boldsymbol{\mathcal{M}}_{(k)}\|_F^2 + \langle\tilde{\mathcal{X}}_{k(k)} - \boldsymbol{\mathcal{M}}_{(k)}, \boldsymbol{\mathcal{G}}_{k(k)}\rangle \\ &\left.+ \frac{\mu_k}{2}\|\mathbf{Q}_k - \mathbf{S}_k\tilde{\mathcal{X}}_{k(k)}\|_F^2 + \langle\mathbf{Q}_k - \mathbf{S}_k\tilde{\mathcal{X}}_{k(k)}, \ \mathbf{Y}_k\rangle\right)\end{aligned} \tag{19}$$

where $\|\cdot\|_F^2$ is the penalty term and ρ_k, λ_k, $\mu_k > 0$ are the penalty parameters. Essentially, the augmented Lagrangian adds penalty terms to the Lagrangian multiplier method to ensure the convergence of the function and the robustness of the dual ascent method.

So far, we have transformed the large-scale tensor completion problem in Equation (19) into the following subproblems, which can be solved iteratively and respectively.

$$\begin{aligned}\boldsymbol{\mathcal{X}}_k^{l+1} &:= \arg\ \min_{\boldsymbol{\mathcal{X}}}\mathcal{L}\left(\left[\boldsymbol{\mathcal{X}}_k^{l+1}, \tilde{\mathcal{X}}_k^{l}, \mathbf{Q}_k^{l}\right]_{k=1}^{3}, \boldsymbol{\mathcal{M}}^{l}\right), \\ \mathbf{Q}_k^{l+1} &:= \arg\ \min_{\mathbf{Q}}\mathcal{L}\left(\left[\boldsymbol{\mathcal{X}}_k^{l+1}, \tilde{\mathcal{X}}_k^{l}, \mathbf{Q}_k^{l+1}\right]_{k=1}^{3}, \boldsymbol{\mathcal{M}}^{l}\right), \\ \tilde{\mathcal{X}}_k^{l+1} &:= \arg\ \min_{\tilde{\mathcal{X}}}\mathcal{L}\left(\left[\boldsymbol{\mathcal{X}}_k^{l+1}, \tilde{\mathcal{X}}_k^{l+1}, \mathbf{Q}_k^{l+1}\right]_{k=1}^{3}, \boldsymbol{\mathcal{M}}^{l}\right), \\ \boldsymbol{\mathcal{M}}^{l+1} &:= \arg\ \min_{\boldsymbol{\mathcal{M}}}\mathcal{L}\left(\left[\boldsymbol{\mathcal{X}}_k^{l+1}, \tilde{\mathcal{X}}_k^{l+1}, \mathbf{Q}_k^{l+1}\right]_{k=1}^{3}, \boldsymbol{\mathcal{M}}^{l+1}\right)\end{aligned} \tag{20}$$

where we sequentially perform recurrent updates for each k $(k = 1,2,3)$ in the order from top to bottom. When the lth and $l+1$th updates satisfy $\|\boldsymbol{\mathcal{X}}^{l+1} - \boldsymbol{\mathcal{X}}^{l}\|_F^2 / \|\mathcal{P}_\Omega(\boldsymbol{\mathcal{Z}})\|_F^2 < \varepsilon$, the algorithm is considered to have converged, and the tensor $\boldsymbol{\mathcal{X}}$ is output as the final result. Algorithm 1 presents the specific steps for solving the ST-LRTC sub-problem in

Equation (20) using the ADMM. For better flow of the paper, we have moved the detailed solution to the Appendix.

Algorithm 1 Solving ST-LRTC for Traffic Data Imputation

Input: Traffic data tensor $\mathcal{Z} \in \mathbb{R}^{M\times T\times D}$, Parameters α, β, ρ, λ, μ, θ, Spatio – Temporal constraint matrices $\boldsymbol{S}_k$ ($k=1,2,3$);
Initialize: $\mathrm{P}_\Omega(\mathbf{X}_k^0)=\mathrm{P}_\Omega(\tilde{\mathbf{X}}_k^0)=\mathrm{P}_\Omega(\mathcal{M}_k^0)=\mathrm{P}_\Omega(\mathcal{Z})$, $\mathcal{T}_k^0=\mathcal{G}_k^0=\mathbf{0}\in\mathbb{R}^{M\times T\times D}$ ($k=1,2,3$), $\boldsymbol{Y}_1^0=\mathbf{0}\in\mathbb{R}^{M\times TD}$, $\boldsymbol{Y}_2^0=\mathbf{0}\in\mathbb{R}^{T\times MD}$, $\boldsymbol{Y}_3^0=\mathbf{0}\in\mathbb{R}^{D\times MT}$, δ, ρ, λ, μ, ε, *MaxIter*, *Iter* = 0;
While *not converged* and *Iter* $\leq$ *MaxIter* **do**
  Update penalty parameters via (A.16);
  For $k=1$ to 3 **do**
    Update tensors $\mathbf{X}_k$ via (A.7);
  End For
    Update matrices $\boldsymbol{Q}_k$ via (A.10);
  For $k=1$ to 3 **do**
    Update auxiliary tensors $\tilde{\mathbf{X}}_k$ via (A.12);
  End For
  Update tensors **M** via (A.14);
  Update dual variables via (A.15);
  Iter = *Iter* + 1;
End While
Output: Traffic tensor data **X** after imputation.

5. Experiments

We select an urban traffic speed dataset and conduct experiments with different missing cases to demonstrate the superior performance of the proposed ST-LRTC method in complex urban scenarios.

5.1. Dataset

Xi'an urban traffic speed data (XA data): The experiments adopt the average driving speed of 500 roads in the central city of Xi'an over 30 days in April[1], as shown in Figure 6. The time interval is 10 minutes, and 144 observations are collected per day, which can be constructed as a matrix (500 road segments $\times$ 4320 time intervals) or tensor (500 road segments $\times$ 144 time intervals per day $\times$ 30 days).

5.2. Experimental setting

5.2.1. Baseline algorithms

We select classic and recently proposed matrix/tensor completion methods as references, which have been shown to significantly outperform traditional methods. These baseline methods can be divided into three categories: matrix decomposition, tensor decomposition, and LRTC imputation models. The details of the baseline methods are listed in Table 1. Note that two methods that incorporate spatio-temporal constraints, TAS-LR (based on matrix decomposition) and GTC (based on tensor decomposition),

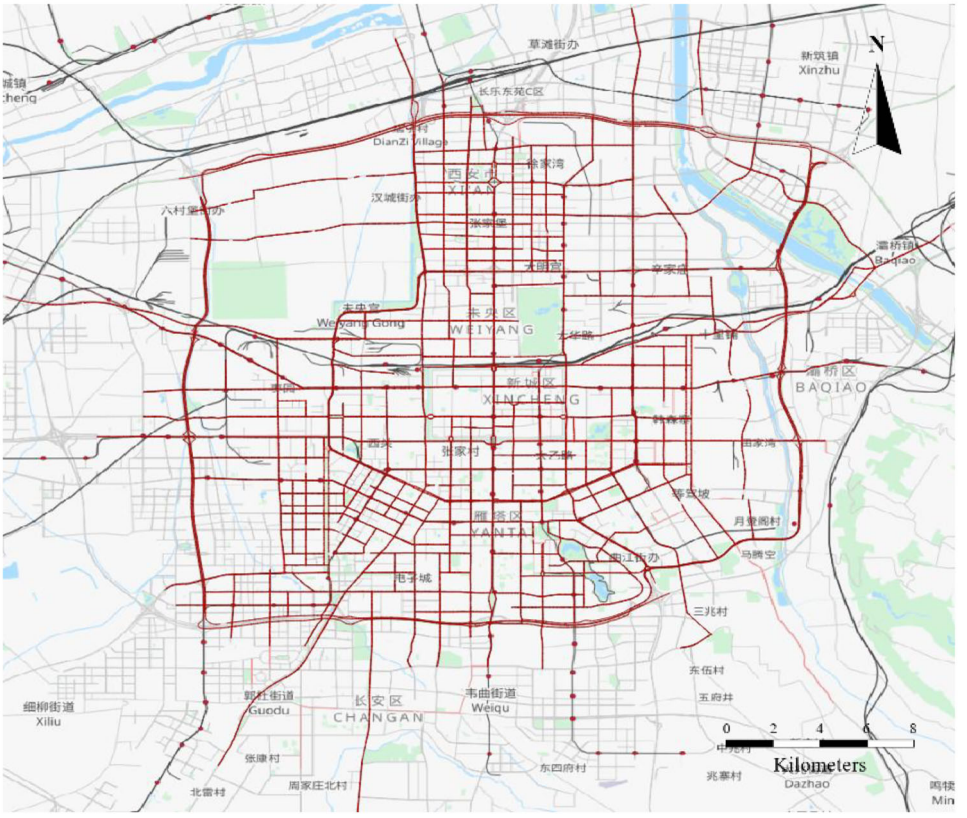


Figure 6. The urban transportation road network in Xi'an, China. Brown lines indicate major road segments in the study area.

are selected for comparison with ST-LRTC to demonstrate the superiority of the proposed method.

5.2.2. Error measures and configurations

To evaluate the imputation accuracy of the model, we simulate three missing patterns on the complete dataset with missing rates ranging from 10 to 90% in steps of 10%. By comparing the difference between the recovered and actual values, we can evaluate the effectiveness of the imputation algorithms using the RMSE and MAPE:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{N}}, MAPE = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \times 100\%$$

where y_i is the actual value of the simulated missing locations, $\hat{y}_i$ is the estimated value of the completion algorithm, and N is the total number of missing values.

In general, when the RMSE and MAPE are smaller, the data imputation accuracy of the completion algorithm is higher.

Table 1. Different types of baseline models.

Data structure		Method	Description
Matrix (Road segments × time-intervals)		BTMF (Chen and Sun 2022)	This method integrates a vector auto-regression model into the latent temporal factors, and develops a Bayesian matrix decomposition model.
		TAS-LR (Wang *et al*. 2019)	Based on matrix decomposition, the method incorporates adaptive spatio-temporal constraints into the latent factors to characterize local features.
Tensor (Road segments × time-intervals × days)	Based on tensor decomposition	TF-ALS (Jain and Oh 2014)	Based on the CP decomposition, the method employs the alternating least squares method for data imputation.
		BGCP (Chen *et al*. 2019)	The method uses Markov chain Monte Carlo to learn the latent factors, and constructs a Bayesian-based tensor factorization model.
		GTC (Deng *et al*. 2021)	Constructing temporal regularization constraints and using a graph-tensor singular value decomposition framework to capture the spatio-temporal characteristics of traffic data.
	Based on nuclear norm relaxation	HaLRTC (Ji *et al*. 2012)	This method utilizes the nuclear norm as a convex relaxation alternative to the rank function and captures the low-rank properties of traffic data globally.
		LRTC-TNN (Chen *et al*. 2020)	This method uses the truncated nuclear norm as a more compact nonconvex substitution for the rank function.

5.2.3. Parameters setting

For the BTMF, TF-ALS, BGCP, HaLRTC, and LRTC-TNN models, we follow the method of setting ranks described by Chen *et al*. (2020). For the spatiotemporally constrained models, we use cross-validation in TAS-LR to select the optimal parameters, whereas in GTC, we follow the original parameters of the model. To ensure the reliability of the experimental results, we set the maximum number of iterations to 200, and adopt the following convergence criterion (Chen *et al*. 2020).

$$\frac{\|\mathcal{X}^{l+1}-\mathcal{X}^{l}\|_F^2}{\|\mathcal{P}_\Omega(\mathcal{Z})\|_F^2}<\varepsilon \text{ or } \frac{\|\boldsymbol{X}^{l+1}-\boldsymbol{X}^{l}\|_F^2}{\|P_\Omega(\boldsymbol{Z})\|_F^2}<\varepsilon,$$

where $\mathcal{X}^l, \boldsymbol{X}^l$ represents the imputed tensor and matrix after the lth iteration, respectively; $\boldsymbol{\mathcal{Z}}$, $\boldsymbol{Z}$ denote the original missing tensor and matrix data.

Similarly, in the ST-LRTC, some parameters need to be set. We chunk the dataset for seven days, and perform parameter selection and cross-validation on each subdata set to ensure that the parameters are selected reasonably and effectively. We finally set the truncated nuclear norm ratio $\theta = 0.15$, the low-rank tensor completion parameter $\alpha = [0.3,0.4,0.3]$, the parameter used to regulate the spatiotemporal constraints $\beta_k = 10^{-5}$, $\mu_k = 10^{-6}$ ($k = 1,2,3$), and other penalty parameters $\rho_k = 10^{-5}$, $\lambda_k = 10^{-5}$ ($k = 1,2,3$).

5.3. Comparisons with random missing pattern

Firstly, we verify the superiority of our algorithm for a random missing pattern. As presented in Table 2, we evaluate the accuracy of the imputation results for the different algorithms with missing rates ranging from 10 to 90% in steps of 10%.

Compared to the baseline models, ST-LRTC yields the best completion results with the lowest error. Specifically, the imputation accuracy is relatively improved by 6.0% on average compared to the LRTC-TNN model, and by more than 30% compared to other state-of-the-art methods.

Furthermore, when comparing the performance of the three types of models under different missing rates, it is observed that the LRTC models (HaLRTC, LRTC-TNN, and ST-LRTC) outperform other baseline models when encountering low missing rates. This indicates that the imputation methods based on tensor nuclear norm relaxation can better capture the low-rank properties of traffic data.

However, the RMSE/MAPE values of HaLRTC and LRTC-TNN fail to achieve the desired results when the missing rate is greater than 60%, whereas our ST-LRTC method is able to maintain a low RMSE/MAPE, even in extreme missing cases. These results illustrate that the combination of tight low-rank estimation and low-dimensional spatiotemporal embedding makes the model more robust and better able to cope with traffic data imputation in different missing scenarios.

For a more intuitive representation of the reliability of the proposed method under extreme absences, Figure 7 shows the errors between the imputation results and observations for the morning peak (8:00 a.m.), off-peak (2:00 p.m.), and evening peak (6:00 p.m.) periods in Xi'an on April 1, 2018.

Table 2. Performance comparison (RMSE/MAPE) under random missing pattern on XA-data.

Missing rate	BTMF	TAS-LR	TF-ALS	BGCP	GTC	HaLRTC	LRTC-TNN	ST-LRTC
0.1	3.257/9.04%	3.228/8.63%	3.343/9.28%	3.329/9.25%	3.613/10.26%	2.557/6.54%	2.438/5.77%	**2.350**/**5.74**%
0.2	3.264/9.03%	3.268/8.71%	3.336/9.25%	3.328/9.22%	3.616/10.20%	2.655/6.89%	2.524/6.03%	**2.424**/**5.98**%
0.3	3.267/9.07%	3.342/8.86%	3.343/9.30%	3.333/9.28%	3.705/10.58%	2.752/7.30%	2.610/6.35%	**2.493**/**6.30**%
0.4	3.273/9.05%	3.368/8.87%	3.343/9.27%	3.340/9.25%	3.758/10.75%	2.850/7.62%	2.707/6.65%	**2.568**/**6.54**%
0.5	3.296/9.07%	3.467/9.01%	3.367/9.33%	3.344/9.26%	3.924/11.21%	2.968/8.01%	2.824/7.02%	**2.659**/**6.82**%
0.6	3.306/9.12%	3.630/9.28%	3.342/9.28%	3.337/9.26%	4.158/11.88%	3.084/8.44%	2.938/7.43%	**2.737**/**7.11**%
0.7	3.330/9.16%	3.861/9.69%	3.343/9.25%	3.335/9.22%	4.570/13.02%	3.232/8.93%	3.078/7.91%	**2.837**/**7.42**%
0.8	3.363/9.24%	4.217/10.54%	3.360/9.31%	3.360/9.31%	5.415/15.25%	3.450/9.67%	3.294/8.65%	**2.985**/**7.87**%
0.9	3.476/9.50%	5.043/13.05%	3.406/9.43%	3.401/9.44%	7.394/19.73%	33.296/100.00%	4.134/10.82%	**3.219**/**8.57**%

Bold values indicate the best performance among different models.

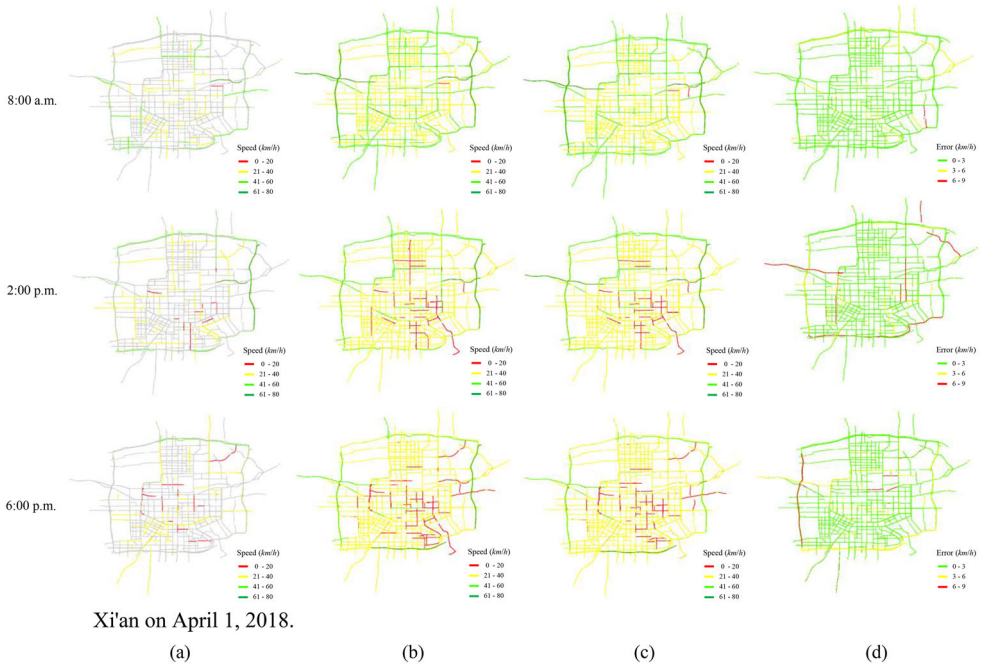


Figure 7. Imputation results of XA-data under extreme RM cases (80%) at different time periods. (a) Map of missing road states, with gray roads indicating the presence of missing data. (b) Observed traffic states. (c) Traffic states after imputation. (d) Error distribution.

Experimental results show that the ST-LRTC model based on spatiotemporal low-dimensional embedding can accurately reconstruct missing data while maintaining the original structure and characteristics of the traffic data. Despite the missing rate reaching 80%, more than 86% of the road segments still have errors of less than 3 km/h (RMSE) and approximately 97% had errors of less than 6 km/h (2RMSE).

In summary, the results indicate that the proposed method can effectively reconstruct missing data with strong stability, albeit with large missing rates.

5.4. Comparisons with non-random missing patterns

Furthermore, we investigate the recovery performance of the proposed algorithm in non-random missing (NM) scenarios, i.e. spatial and temporal missing patterns. Figure 8 shows the relative errors of the different imputation algorithms at different missing rates.

Overall, ST-LRTC has better and more stable imputation results than the baseline models. Specifically, ST-LRTC achieves a performance similar to that of LRTC-TNN at low missing rates. However, owing to the full utilization of the topological structure of road networks and the spatiotemporal characteristics of traffic flow, ST-LRTC achieves much better performance under high absences.

In addition, by combining the RM results in Section 5.3, it can be seen that the MAPE in the RM case is lower than that in the NM case with respect to different missing patterns. All baseline models can achieve the desired imputation accuracy in the RM cases. However, under NM, particularly in extreme cases (missing rate $\geq$ 70%),

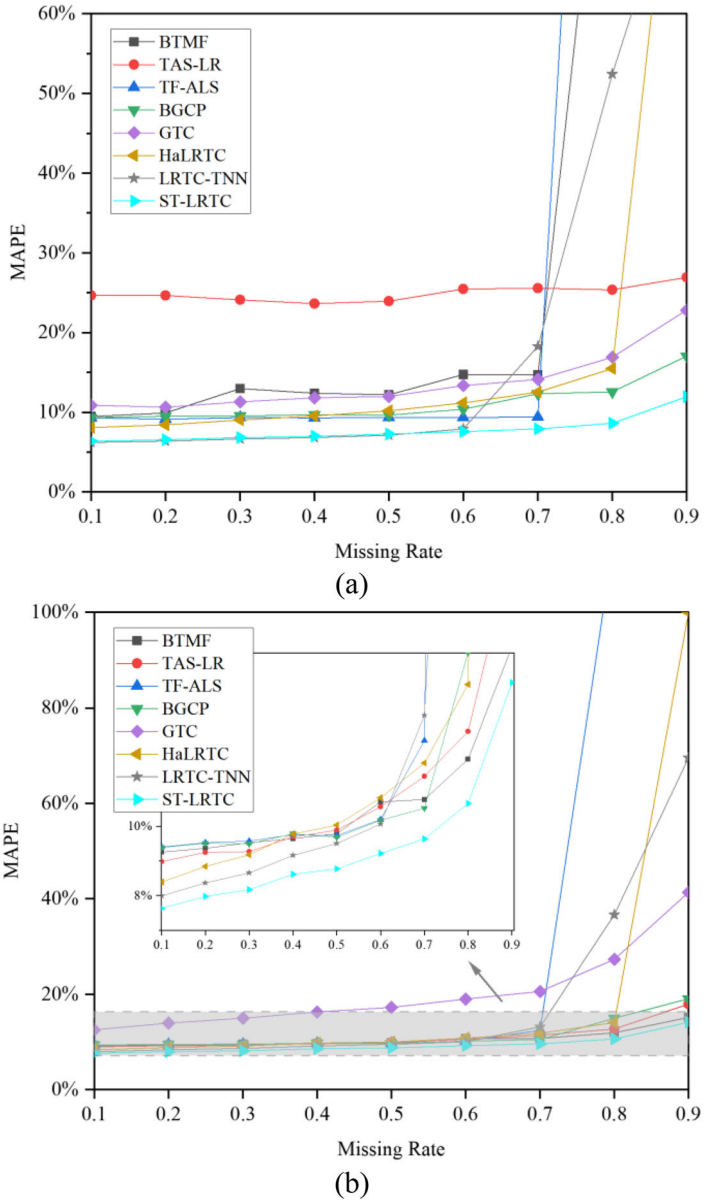


Figure 8. Relative errors at different sampling rates under non-random missing patterns. (a) Spatial Missing Pattern (SM). (b) Temporal Missing Pattern (TM).

most baseline models are less accurate (MAPE reaches more than 60%). This indicates that data imputation in NM cases tends to be more challenging than in RM cases, and the baseline models do not achieve desirable results in NM cases. In contrast, our ST-LRTC method can achieve higher stability in different types of extreme absences, and the imputation relative errors (MAPE) are all guaranteed to be less than 15%.

In-depth, we compare two baseline methods that incorporate spatiotemporal constraints with ST-LRTC. TAS-LR is a matrix-based imputation algorithm vulnerable to

structural deficiencies. Notably, TAS-LR is particularly ineffective for the imputation of the SM pattern because structural absences destroy the original intrinsic structure of the data matrix, leading to the non-convergence of iterations. For GTC, on the one hand, the graph tensor singular value decomposition requires a predefined definition of rank. However, when only a few observations are available, an improper ranking may lead to model divergence. On the other hand, because urban road states are always variable and feature-rich, the 14th-order Toeplitz matrix used in the GTC may smooth out detailed information. Thus, it is not suitable for data imputation of urban roads.

Thorough experiments demonstrate that the proposed ST-LRTC model achieves better imputation results for all missing patterns, and is more applicable to urban traffic data imputation with complex topologies and variable states.

5.5. Ablation studies

In what follows, we illustrate the positive effect of spatio-temporal manifold embedding on low-rank tensor completion from the perspective of error distribution, and demonstrate the superiority of the K-order proximity-based 'static road topology & dynamic traffic flow' co-driven approach.

5.5.1. The advantages of spatio-temporal manifold embedding

To further evaluate the effect of introducing spatial and temporal manifold embedding, we compare the distributions of imputation errors under extreme missing rates (80%) before and after incorporating the constraints, as shown in Figure 9.

As can be seen, compared with the LRTC-TNN model, the distributions of ST-LRTC imputation error are all 'leaner and taller' and more clustered around 0. In other words, the imputation results are more accurate and stable with the spatiotemporal constraints.

Furthermore, as shown in Figure 10, we compare the changes in the imputation errors on each segment before and after introducing the spatiotemporal constraints under RM (the same results under NM) from different time scales. It can be seen that the accuracy of data imputation is improved for both the medium (one week) and long (one month) time periods. For daily data imputation, we also obtain a higher accuracy for the majority of road segments.

5.5.2. The superiority of 'static road topology & dynamic traffic flow' co-driven method

In ST-LRTC, we propose a 'static & dynamic' co-driven weighted adjacency (WA) matrix based on K-order proximity, which combines traffic flow similarity with the topology of the real road network. Here, we analyze the superiority of doing so by comparing the accuracy with that of a static adjacency (SA) matrix (consisting of 0/1) using only topology, and a traffic correlation (TC) matrix (calculated by Pearson coefficient p_{ij}) using only data similarity. Taking the SM pattern as an example, Table 3 lists the imputation comparison errors using TC, SA, and WA.

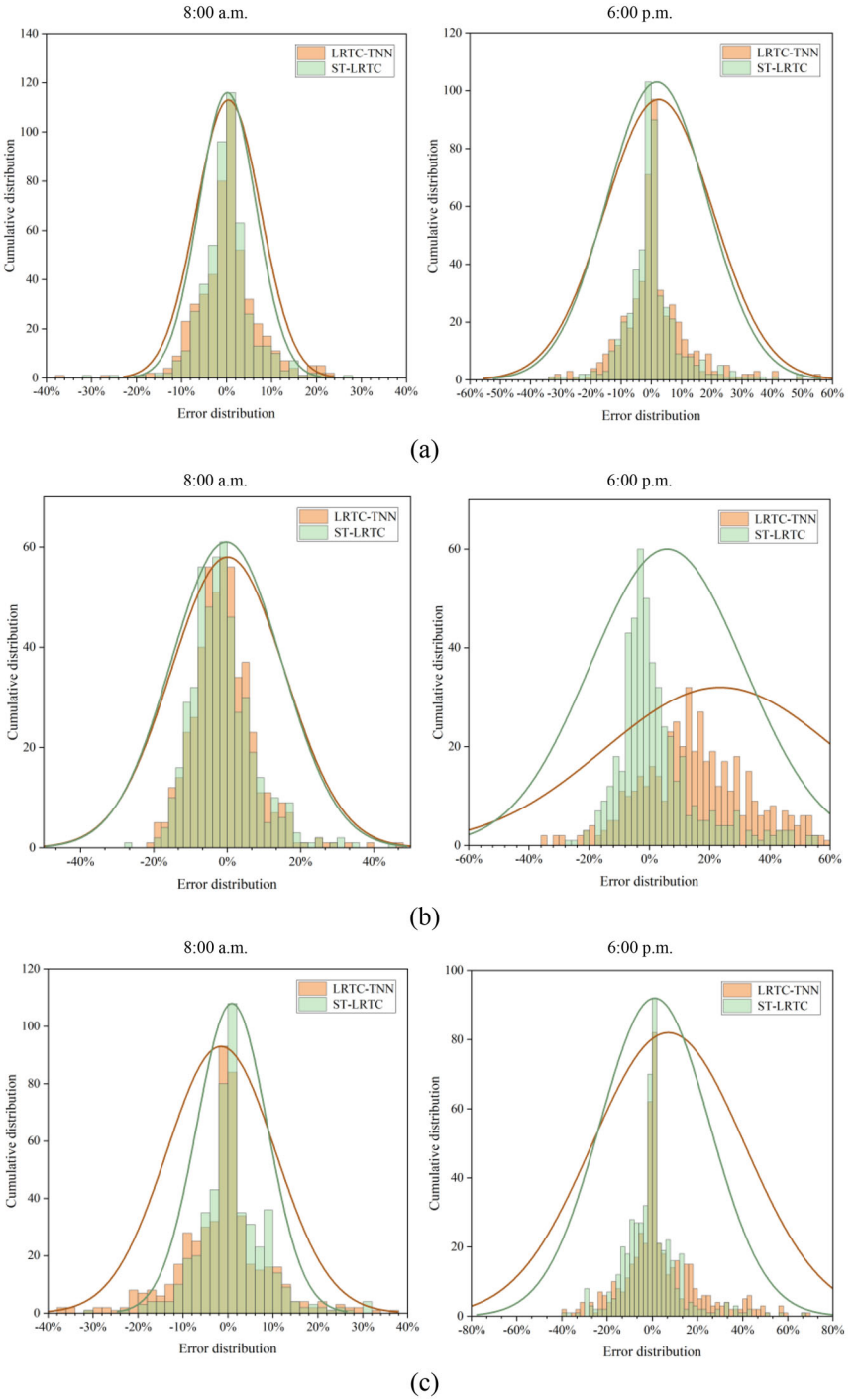


Figure 9. Distribution of the imputation error before and after introducing spatiotemporal constraints (8:00 a.m. and 6:00 p.m.) (a) RM. (b) SM. (c) TM.

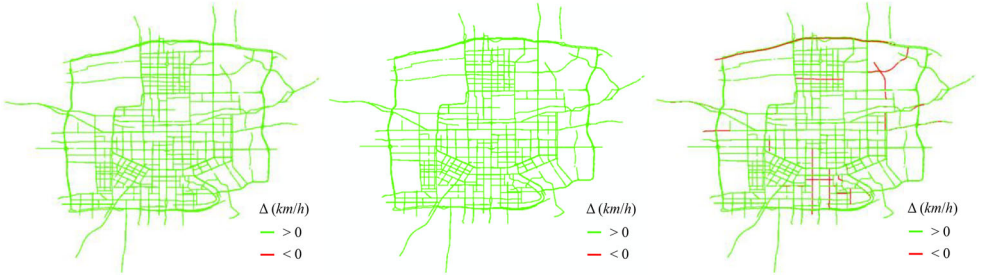


Figure 10. Comparison of the imputation errors of each road segment before and after introducing spatiotemporal correlations under the missing rate of 80%. (a) 1-month. (b) 1-week. (c) 1-day. Green indicates that the accuracy of data imputation has improved, and red indicates the opposite.

The experimental results indicate that the ST-LRTC constructed using WA can achieve the best overall imputation accuracy. Specifically, the spatial matrix construction method, which considers the actual road topology, has better imputation results than methods using only the data correlation between road segments. The WA, in turn, is more flexible than the SA and better reflects the different degrees of traffic flow transmission and feedback between adjacent roads.

The results reveal that the improvement in WA accuracy becomes more obvious, particularly when the missing rates are $\geq$ 70% (underlined in the table). This is due to the fact that high missing rates imply limited valid data for road segments, and traffic data imputation is more dependent on surrounding adjacent roads. The results indicate that the ST-LRTC model, which incorporates a weighted adjacency matrix, can better reconstruct traffic data with high missing data rates.

5.6. Discussion

5.6.1. Computation time

Admittedly, in terms of algorithmic efficiency, the introduction of spatiotemporal constraints increases the running time of individual loops. However, the prior knowledge it provides can facilitate the convergence of the algorithm, thus reducing the number of iterations and shortening its overall running time.

In summary, the proposed algorithm can obtain excellent results both in terms of imputation effect and operation efficiency, and can be more applicable to traffic data imputation in complex urban scenarios.

5.6.2. Extensibility

Traffic speed data imputation, considering its representativeness, is selected as the scenario for data completion and spatial analysis in this study. Thorough experiments demonstrated that the incorporation of spatiotemporal analysis can enhance the adaptability of the tensor completion model to complex topologies and variable states.

Indeed, traffic speed data imputation is a typical application scenario of urban sensing. These can be replaced by other types of geographic data. For example, we can denote origin-destination (OD) flow by a tensor $\mathcal{X}^{Origin \times Destination \times Time}$ to mine crowd

Table 3. Effect of whether to use a weighted adjacency matrix on imputation accuracy (RMSE/MAPE).

Missing rate	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Traffic Correlation matrix (TC, p_{ij})	3.02/7.63%	3.10/7.97%	3.18/8.18%	3.31/8.62%	3.32/8.79%	3.40/9.19%	3.53/9.58%	3.72/10.23%	5.53/13.48%
Static Adjacency matrix (SA, 0/1)	2.44/6.39%	2.54/6.56%	2.64/6.85%	2.66/7.01%	2.74/7.30%	2.87/7.65%	3.01/8.08%	3.40/9.18%	5.30/13.29%
Weighted Adjacency matrix (WA, w_{ij})	**2.43/6.38%**	**2.53/6.54%**	**2.63/6.83%**	**2.65/6.98%**	**2.73/7.26%**	**2.85/7.58%**	**2.94/7.91%**	**3.15/8.63%**	**4.46/11.96%**

Bold values indicate the best performance among different models. Underlined values indicate that the improvement in WA accuracy is more obvious when the missing rates are ≥ 0.7.

movement patterns (Cao *et al*. 2020), or represent traffic flow data with a tensor $\mathcal{X}^{Road\times Time\times Day}$ to facilitate accurate traffic state imputation and forecasting (Zhou *et al*. 2019).

In terms of theoretical methods, embedding local spatiotemporal manifold features into global low-rank tensor models can facilitate multi-scale and fine-grained perceptions of urban scenes. The multi-perspective analysis framework proposed in this study can also be applied to other aspects of GIS, such as travel pattern mining (Cao *et al*. 2020, Gao *et al*. 2019), social hotspots detection (Li *et al*. 2018), and air quality forecasting (Xu *et al*. 2019, Zhang *et al*. 2022).

Overall, incorporating tensor models into geospatial analysis provides a better way to deal with spatiotemporal field data, reveal the intrinsic structure of complex spatiotemporal data, and capture interdependencies in multiple dimensions. It also enables a unified framework for irregular spatiotemporal fields from data organization, management, and imputation to data analysis and application (Li *et al*. 2022). On the other hand, the incorporation of geographic thinking endows tensor models with explicit representations of spatiotemporal patterns, thereby facilitating the practical application of high-dimensional analytical methods in real scenarios, such as urban perception and pattern analysis.

5.6.3. Threats to validity

As stated in Section 4, our ST-LRTC method can achieve higher stability in different types of extreme absences. However, the degree of regularity of urban traffic states is an influential indicator of the model accuracy. The more volatile the traffic states, the more difficult it is to recover missing data using the observed data, and the less accurate the model. Therefore, assessing whether missing data can be accurately completed (i.e. recoverability) is also an essential issue.

Moreover, the incorporation of spatiotemporal features inevitably increases the number of model parameters. The selection of the parameters may also affect the generality of the results and their application in other contexts.

6. Conclusion and future work

Many existing approaches for geographic analysis are based on the assumptions that 1) data are complete, and 2) data are uniformly distributed in space and time. However, the incomplete nature of geographic data significantly affects the accuracy of many GIS models. Therefore, data imputation, particularly for complex urban scenarios, is a critical and essential part between geographic big data collection and application.

In this study, we considered urban traffic as a representative research scenario for data imputation, considering its consistent data dimensions, significant time-series characteristics, and complex spatial association features. To recover missing traffic data, the Spatio-Temporal constrained Low-Rank Tensor Completion (ST-LRTC) was proposed in this paper and applied to urban traffic speed data imputation. Specifically, we adopted a manifold embedding approach to depict the local geometric structure across each mode and proposed a ‘static road topology & dynamic traffic

flow' co-driven approach for road network spatial similarity metrics. Further, we embedded the 'local' spatio-temporal manifold regularization terms into the 'global' low-rank tensor model, and developed a joint imputation method with 'static and dynamic, global and local' multiple views for urban traffic data.

Urban traffic speed data (XA data) were selected for the experiment, and the results indicated that our ST-LRTC method achieved more accurate imputation results under various missing rates and patterns, and the imputation relative errors (MAPE) were guaranteed to be less than 15%. Thorough experiments demonstrated that the spatio-temporal feature embedding approach and the K-order proximity-based 'static-dynamic' co-driven spatial similarity metrics developed in this paper can enhance the adaptability of the imputation model to complex urban topology and variable states.

However, because of the limited information considered in this study, the processing of a single type of dataset is no longer sufficient to achieve accurate data recovery under sudden circumstances. Therefore, in future studies, we will attempt to incorporate more data (e.g. event and POI information) to make the imputation results more realistic.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Note

1. https://sts.didiglobal.com

Funding

This research was supported by the Fundamental Research Funds for the Central Universities, and the National Natural Science Foundation of China [grant numbers 41971405, 41671442, 42271449, 41901394].

Notes on contributors

Zilong Zhao received his B.E. degree from Wuhan University, Wuhan, China, in 2021, where he is currently pursuing an M.S. degree at the State Key Laboratory of Information Engineering in Surveying, Mapping, and Remote Sensing. His research interests include spatiotemporal data mining, tensor imputation, and intelligent transportation systems. Contributions: Methodology, Software, Writing - original draft.

Luliang Tang received his Ph.D. degree from Wuhan University, Wuhan, China, in 2007. He is currently a Professor at Wuhan University. His research interests include spatiotemporal GIS, GIS for transportation, and change detection. Contributions: Writing - Review & Editing, Funding acquisition.

Mengyuan Fang received his M.E. degree from Wuhan University, Wuhan, China, in 2020. His research interests include spatiotemporal data mining and geographic information science. Contributions: Conceptualization, Writing - Review & Editing.

Xue Yang received her Ph.D. degree from Wuhan University, Wuhan, China, in 2018. She is currently an associate professor at China University of Geosciences, Wuhan. Her research interests

include intelligent transportation systems, spatiotemporal data analysis, and information mining. Contributions: Conceptualization, Funding acquisition.

Chaokui Li received his Ph.D. degree from Central South University, Changsha, China, in 2001. He is currently a Professor at Hunan University of Science and Technology. His research interests include 3D Geographic Modeling and Geographic Information Systems. Contributions: Formal analysis, Resources.

Qingquan Li received his Ph.D. degree in geographic information system (GIS) and photogrammetry from Wuhan Technical University of Surveying and Mapping, Wuhan, China, in 1998. He is currently a Professor at Shenzhen University, Guangdong, China. His research interests include dynamic data modeling in GIS, surveying engineering, and intelligent transportation system. Contributions: Resources, Supervision.

Data and codes availability statement

The data and codes that support the findings of this study are available in 'figshare.com' with the identifier https://doi.org/10.6084/m9.figshare.20289078.

References

Acar, E., *et al.*, 2011. Scalable tensor factorizations for incomplete data. *Chemometrics and Intelligent Laboratory Systems*, 106 (1), 41–56.

Boyd, S., 2010. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends® in Machine Learning*, 3 (1), 1–122.

Cai, J.-F., Candès, E.J., and Shen, Z., 2010. A singular value thresholding algorithm for matrix completion. *SIAM Journal on Optimization*, 20 (4), 1956–1982.

Candès, E.J., *et al.*, 2011. Robust principal component analysis? *Journal of the ACM*, 58 (3), 1–37.

Candes, E.J., and Wakin, M.B., 2008. An introduction to compressive sampling. *IEEE Signal Processing Magazine*, 25 (2), 21–30.

Cao, F., *et al.*, 2017. Recovering low-rank and sparse matrix based on the truncated nuclear norm. *Neural Networks*, 85, 10–20.

Cao, M., *et al.*, 2020. Analysis of the spatiotemporal riding modes of dockless shared bicycles based on tensor decomposition. *International Journal of Geographical Information Science*, 34 (11), 2225–2242.

Carroll, J.D., and Chang, J.J., 1970. Analysis of individual differences in multidimensional scaling via an n-way generalization of "Eckart-Young" decomposition. *Psychometrika*, 35 (3), 283–319.

Chen, C., *et al.*, 2003. Detecting errors and imputing missing data for single-loop surveillance systems. *Transportation Research Record*, 1855 (1), 160–167.

Chen, X., He, Z., and Sun, L., 2019. A Bayesian tensor decomposition approach for spatiotemporal traffic data imputation. *Transportation Research Part C: Emerging Technologies*, 98, 73–84.

Chen, X., and Sun, L., 2022. Bayesian temporal factorization for multidimensional time series prediction. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44 (9), 4659–4673.

Chen, X., Yang, J., and Sun, L., 2020. A nonconvex low-rank tensor completion model for spatiotemporal traffic data imputation. *Transportation Research Part C: Emerging Technologies*, 117, 102673.

Deng, L., *et al.*, 2021. Graph spectral regularized tensor completion for traffic data imputation. *IEEE Transactions on Intelligent Transportation Systems*, 23 (8), 10996–11010.

Dodge, S., *et al.*, 2020. Progress in computational movement analysis–towards movement data science. *International Journal of Geographical Information Science*, 34 (12), 2395–2400.

Donoho, D.L., 2006. Compressed sensing. *IEEE Transactions on Information Theory*, 52 (4), 1289–1306.

Douxchamps, D., Macq, B., and Chihara, K., 2006. High accuracy traffic monitoring using roadside line-scan cameras. *IEEE Intelligent Transportation Systems Conference*, 17–20 September 2006, Toronto, ON, Canada, 875–878.

Fang, M., *et al.*, 2022. FTPG: A fine-grained traffic prediction method with graph attention network using big trace data. *IEEE Transactions on Intelligent Transportation Systems*, 23 (6), 5163–5175.

Feng, X., *et al.*, 2022. Traffic data recovery from corrupted and incomplete observations via spatial-temporal TRPCA. *IEEE Transactions on Intelligent Transportation Systems*, 23 (10), 17835–17848.

Gao, Y., *et al.*, 2019. A spatiotemporal constraint non-negative matrix factorization model to discover intra-urban mobility patterns from taxi trips. *Sustainability*, 11 (15), 4214.

Goulart, J.H.D.M., Kibangou, A.Y., and Favier, G., 2017. Traffic data imputation via tensor completion based on soft thresholding of Tucker core. *Transportation Research Part C: Emerging Technologies*, 85, 348–362.

Han, Z.F., *et al.*, 2017. Sparse and truncated nuclear norm based tensor completion. *Neural Processing Letters*, 45 (3), 729–743.

Harshman, R.A., 1970. Foundations of the parafac procedure: models and conditions for an "explanatory" multimodal factor analysis. *Ucla Working Papers in Phonetics*, 16, 1–84.

Hillar, C.J., and Lim, L.H., 2013. Most tensor problems are NP-hard. *Journal of the ACM*, 60 (6), 1–39.

Hu, Y., and Work, D.B., 2021. Robust tensor recovery with fiber outliers for traffic events. *ACM Transactions on Knowledge Discovery from Data*, 15 (1), 1–27.

Hu, Y., *et al.*, 2013. Fast and accurate matrix completion via truncated nuclear norm regularization. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35 (9), 2117–2130.

Huang, L.T., *et al.*, 2014. Truncated nuclear norm minimization for tensor completion. *8th IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM)*, 22–25 June 2014 A Coruna, Spain, 417–420.

Jain, P., and Oh, S., 2014. Provable tensor factorization with missing data. *28th Conference on Neural Information Processing Systems (NIPS)*, 8–13 December 2014 Montreal, Canada, 1431–1439.

Ji, L., *et al.*, 2012. Tensor completion for estimating missing values in visual data. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35 (1), 208–220.

Kolda., *et al.*, 2009. Tensor decompositions and applications. *SIAM Review*, 51 (3), 455–500.

Li, D., *et al.*, 2022. A tensor-based approach to unify organization and operation of data for irregular spatio-temporal fields. *International Journal of Geographical Information Science*, 36 (9), 1885–1904.

Li, L., Li, Y., and Li, Z., 2013. Efficient missing data imputing for traffic flow by considering temporal and spatial dependence. *Transportation Research Part C: Emerging Technologies*, 34 (Sep), 108–120.

Li, Q.-Q., Zhang, T., and Yu, Y., 2011a. Using cloud computing to process intensive floating car data for urban traffic surveillance. *International Journal of Geographical Information Science*, 25 (8), 1303–1322.

Li, Q., *et al.*, 2018. Retweeting prediction based on social hotspots and dynamic tensor decomposition. *IEICE Transactions on Information and Systems*, E101.D (5), 1380–1392.

Li, Q., *et al.*, 2008. A BPCA based missing value imputing method for traffic flow volume data. *IEEE Intelligent Vehicles Symposium*, 4–6 June 2008 Eindhoven, Netherlands, 171–176.

Li, X., *et al.*, 2017. MR-NTD: Manifold regularization nonnegative tucker decomposition for tensor data dimension reduction and representation. *IEEE Transactions on Neural Networks and Learning Systems*, 28 (8), 1787–1800.

Li, Y., *et al.*, 2018. Diffusion convolutional recurrent neural network: data-driven traffic forecasting. *6th International Conference on Learning Representations (ICLR)*, April 30–May 3 2018 Vancouver, BC, Canada, 1–16.

Li, Z., *et al.*, 2011b. Compressive sensing approach to urban traffic sensing. *31st International Conference on Distributed Computing Systems (ICDCS)*, 20–24 June 2011 Minneapolis, MN, 889–898.

Lin, K., *et al.*, 2018. A novel spatial-temporal regularized tensor completion algorithm for traffic data imputation. *10th International Conference on Wireless Communications and Signal Processing (WCSP)*, 18–20 October 2018 Hangzhou, Peoples R China, 1–6.

Liu, J., Ji, S., and Ye, J., 2009. SLEP: Sparse learning with efficient projections. *Arizona State University*, 6 (491), 7.

Long, Z., *et al.*, 2019. Low rank tensor completion for multiway visual data. *Signal Processing*, 155, 301–316.

Peng, X., *et al.*, 2016. Connections between nuclear norm and Frobenius norm based representation. *IEEE Transactions on Neural Networks & Learning Systems*, 29 (1), 218–224.

Qu, L., *et al.*, 2009. PPCA-based missing data imputation for traffic flow volume: a systematical approach. *IEEE Transactions on Intelligent Transportation Systems*, 10 (3), 512–522.

Rodrigues, J.G.P., *et al.*, 2011. A mobile sensing architecture for massive urban scanning.*14th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, 5–7 October 2011 Campus George Washington Univ (GWU), Washington, DC, 1132–1137.

Said, A.B., and Erradi, A., 2021. Spatiotemporal tensor completion for improved urban traffic imputation. *IEEE Transactions on Intelligent Transportation Systems*, 23 (7), 6836–6849.

Smith, B.L., Scherer, W.T., and Conklin, J.H., 2003. Exploring imputation techniques for missing data in transportation management systems. *Transportation Research Record*, 1836 (1), 132–142.

Sofuoglu, S.E., and Aviyente, S., 2022. GLOSS: tensor-based anomaly detection in spatiotemporal urban traffic data. *Signal Processing*, 192, 108370.

Tan, H., *et al.*, 2013. A tensor-based method for missing traffic data completion. *Transportation Research Part C: Emerging Technologies*, 28 (Mar), 15–27.

Tang, L., *et al.*, 2016. Travel time estimation at intersections based on low-frequency spatial-temporal GPS trajectory big data. *Cartography and Geographic Information Science*, 43 (5), 417–426.

Tucker, L.R., 1966. Some mathematical note on three-mode factor analysis. *Psychometrika*, 31 (3), 279–311.

Wang, Y., *et al.*, 2019. Traffic data reconstruction via adaptive spatial-temporal correlations. *IEEE Transactions on Intelligent Transportation Systems*, 20 (4), 1531–1543.

Xu, Y., *et al.*, 2019. Fine-grained air quality inference with remote sensing data and ubiquitous urban data. *ACM Transactions on Knowledge Discovery from Data*, 13 (5), 1–27.

Yang, J.-H., *et al.*, 2018. HOP-rec: high-order proximity for implicit recommendation. *12th ACM Conference on Recommender Systems (RecSys)*, 2–7 October 2018 Vancouver, Canada, 140–144.

Yang, Y., *et al.*, 2017. Understanding structure of urban traffic network based on spatial-temporal correlation analysis. *Modern Physics Letters B*, 31 (22), 1750230.

Ye, J., *et al.*, 2020. How to build a graph-based deep learning architecture in traffic domain: a survey. *IEEE Transactions on Intelligent Transportation Systems*, 23 (5), 3904–3924.

Yuebiao, *et al.*, 2014. Missing traffic data: comparison of imputation methods. *IET Intelligent Transport Systems*, 8 (1), 51–57.

Zhang, D., *et al.*, 2012. Matrix completion by truncated nuclear norm regularization. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 16–21 June 2012 Providence, RI, 2192–2199.

Zhang, H., *et al.*, 2022. Spatio-temporal changes in air quality of the urban area of Chongqing from 2015 to 2021 based on a missing-data-filled dataset. *Atmosphere*, 13 (9), 1473.

Zhang, Y., *et al.*, 2020. A novel residual graph convolution deep learning model for short-term network-based traffic forecasting. *International Journal of Geographical Information Science*, 34 (5), 969–995.

Zhang, Z., *et al.*, 2020. Network-wide traffic flow estimation with insufficient volume detection and crowdsourcing data. *Transportation Research Part C: Emerging Technologies*, 121, 102870.

Zhou, H., *et al.*, 2015. Spatio-temporal tensor completion for estimating missing internet traffic data. *Computing and Communications Conference (IPCCC), 2015 IEEE 34th International Performance*, 14–16 December 2015 Nanjing, Peoples R China, 1–7.

Zhou, Y., *et al.*, 2019. A reliable traffic prediction approach for bike-sharing system by exploiting rich information with temporal link prediction strategy. *Transactions in GIS*, 23 (5), 1125–1151.

Appendix A. Solving ST-LRTC for traffic data imputation

For ST-LRTC model, the variables optimization problem in Equation (20) can be solved by alternative iterations, in which the superscript *l* denotes the current iteration step.

a. Updating tensors $\mathcal{X}_k$ (k = 1,2,3)

For each k, the update of tensor $\mathcal{X}_k$ can be expressed as:

$$\begin{aligned}\mathcal{X}_k^{l+1} &:= \arg\ \min_{\mathcal{X}_k}\ \alpha_k\|\mathcal{X}_{k(k)}^{l+1}\|_{r_k} + \frac{\rho_k}{2}\|\mathcal{X}_{k(k)}^{l+1} - \mathcal{M}_{(k)}^{l}\|_F^2 + \langle \mathcal{X}_{k(k)}^{l+1} - \mathcal{M}_{(k)}^{l}, \mathcal{T}_{k(k)}^{l}\rangle \\ &= \arg\ \min_{\mathcal{X}_k}\ \alpha_k\|\mathcal{X}_{k(k)}^{l+1}\|_{r_k} + \frac{\rho_k}{2}\|\mathcal{X}_{k(k)}^{l+1}\|_F^2 - \rho_k\langle \mathcal{X}_{k(k)}^{l+1}, \mathcal{M}_{(k)}^{l}\rangle + \frac{\rho_k}{2}\|\mathcal{M}_{(k)}^{l}\|_F^2 + \langle \mathcal{X}_{k(k)}^{l+1}, \mathcal{T}_{k(k)}^{l}\rangle - \langle \mathcal{M}_{(k)}^{l}, \mathcal{T}_{k(k)}^{l}\rangle\end{aligned} \tag{A.1}$$

Further, when minimizing tensor $\mathcal{X}_k$, tensor $\mathbf{M}$ can be considered as an invariant, i.e.

$$\begin{aligned}\mathcal{X}_k^{l+1} &:= \arg\ \min_{\mathcal{X}_k}\ \alpha_k\|\mathcal{X}_{k(k)}^{l+1}\|_{r_k} + \frac{\rho_k}{2}\|\mathcal{X}_{k(k)}^{l+1}\|_F^2 - \rho_k\langle \mathcal{X}_{k(k)}^{l+1}, \mathcal{M}_{(k)}^{l}\rangle + \langle \mathcal{X}_{k(k)}^{l+1}, \mathcal{T}_{k(k)}^{l}\rangle \\ &= \arg\ \min_{\mathcal{X}_k}\ \alpha_k\left\|\|\mathcal{X}_{k(k)}^{l+1}\|_{r_k} + \frac{\rho_k}{2}\right\|\mathcal{X}_{k(k)}^{l+1}\|_F^2 - \rho_k\left\langle \mathcal{X}_{k(k)}^{l+1}, \mathcal{M}_{(k)}^{l} - \frac{\mathcal{T}_{k(k)}^{l}}{\rho_k}\right\rangle \\ &= \arg\ \min_{\mathcal{X}_k}\ \alpha_k\|\mathcal{X}_{k(k)}^{l+1}\|_{r_k} + \frac{\rho_k}{2}\|\mathcal{X}_{k(k)}^{l+1} - (\mathcal{M}_{(k)}^{l} - \frac{\mathcal{T}_{k(k)}^{l}}{\rho_k})\|_F^2\end{aligned} \tag{A.2}$$

Then the above optimization problem can be solved by following Lemma 1.

Lemma 1 (Chen *et al.* 2020, Cai *et al.* 2010) For any α, $\rho > 0$, $\mathbf{Z} \in \mathbf{R}^{m\times n}$, and $r \in \mathbf{N}_+$ where $r < \min\{m,n\}$, an optimal solution to the problem

$$\min_{\mathcal{X}}\ \alpha\|\mathbf{X}\|_r + \frac{\rho}{2}\|\mathbf{X} - \mathbf{Z}\|_F^2 \tag{A.3}$$

is given by the generalized singular value thresholding

$$\hat{\mathbf{X}} = \mathbf{U}\mathbf{\Sigma}_{\alpha/\rho}\mathbf{V}^T \tag{A.4}$$

where $\boldsymbol{U\Sigma V}^T$ is the singular value decomposition of $\boldsymbol{Z}$. Therefore, the shrinkage of singular values can be defined as

$$\mathbf{\Sigma}_{\alpha/\rho} = \mathrm{diag}\left((\sigma_1, \sigma_2, ..., \sigma_r, [\sigma_{r+1} - \alpha/\rho]_+, ..., [\sigma_{\min\{m,n\}} - \alpha/\rho]_+)^T\right) \tag{A.5}$$

where $\sigma_1, \sigma_2, ..., \sigma_{\min\{m,n\}}$ are diagonal entries of $\mathbf{\Sigma}$, and $[\cdot]_+$ denotes the positive truncation at *0* satisfying $[x]_+ = max\ (x, 0)$.

According to Lemma 1, $\boldsymbol{Z}$ corresponds to $\mathcal{M}_{(k)}^{l} - \mathcal{T}_{k(k)}^{l}/\rho_k$ in Equation (A.2). Thus, the shrinkage of singular values for the problem (A.2) can be expressed as

$$\sigma_i(\mathcal{X}_{(k)}) = \begin{cases}\left[\sigma_i(\mathcal{M}_{(k)}^{l} - \mathcal{T}_{k(k)}^{l}/\rho_k) - \alpha_k/\rho_k\right]_+, & \text{if } i > r_k; \\ \sigma_i(\mathcal{M}_{(k)}^{l} - \mathcal{T}_{k(k)}^{l}/\rho_k), & \text{otherwise,}\end{cases} \tag{A.6}$$

where $\boldsymbol{U\Sigma V}^T$ corresponds to the singular value decomposition of $\mathcal{M}_{(k)}^{l} - \mathcal{T}_{k(k)}^{l}/\rho_k$, $\sigma_i(\mathcal{M}_{(k)}^{l} - \mathcal{T}_{k(k)}^{l}/\rho_k)$ denotes the i-th diagonal entry of $\mathbf{\Sigma}$. To this end, the truncated nuclear norm minimization problem in Eq. (A.2) has a closed-form solution:

$$\mathbf{X}_k^{l+1} = fold_k(\mathbf{U}diag(\sigma(\mathbf{X}_{(k)}))\mathbf{V}^T). \tag{A.7}$$

b. Updating matrices $\mathbf{Q}_k$ (k = 1,2,3)

The optimization of matrices $\boldsymbol{Q}_k$s is based on the effect of spatiotemporal constraint matrices $\boldsymbol{S}_k$s in the auxiliary tensor. The minimization of $\boldsymbol{Q}_k$s enables similar traffic states between neighboring moments, adjacent days, and topologically connected segments, and its update equation can be expressed as:

$$\begin{aligned}\mathbf{Q}_k^{l+1} &:= \arg\ \min_{\mathbf{Q}_k} \beta_k\|\mathbf{Q}_k^{l+1}\|_1 + \frac{\mu_k}{2}\|\mathbf{Q}_k^{l+1} - \mathbf{S}_k\tilde{\mathcal{X}}_{k(k)}^{l}\|_F^2 + \langle \mathbf{Q}_k^{l+1} - \mathbf{S}_k\tilde{\mathcal{X}}_{k(k)}^{l},\ \mathbf{Y}_k^{l}\rangle \\ &= \arg\ \min_{\mathbf{Q}_k} \beta_k\|\mathbf{Q}_k^{l+1}\|_1 + \frac{\mu_k}{2}\|\mathbf{Q}_k^{l+1}\|_F^2 - \mu_k\langle \mathbf{Q}_k^{l+1}, \mathbf{S}_k\tilde{\mathcal{X}}_{k(k)}^{l}\rangle + \langle \mathbf{Q}_k^{l+1}, \mathbf{Y}_k^{l}\rangle \\ &= \arg\ \min_{\mathbf{Q}_k} \beta_k\|\mathbf{Q}_k^{l+1}\|_1 + \frac{\mu_k}{2}\|\mathbf{Q}_k^{l+1}\|_F^2 - \mu_k\left\langle \mathbf{Q}_k^{l+1}, \mathbf{S}_k\tilde{\mathcal{X}}_{k(k)}^{l} - \frac{\mathbf{Y}_k^{l}}{\mu_k}\right\rangle \\ &= \arg\ \min_{\mathbf{Q}_k} \beta_k\|\mathbf{Q}_k^{l+1}\|_1 + \frac{\mu_k}{2}\left\|\mathbf{Q}_k^{l+1} - \mathbf{S}_k\tilde{\mathcal{X}}_{k(k)}^{l} + \frac{\mathbf{Y}_k^{l}}{\mu_k}\right\|_F^2\end{aligned} \quad . \tag{A.8}$$

For the joint optimization of the L1-norm and the Frobenius norm (the induced matrix norm of the L2-norm (PENG *et al.* 2016)), we utilize the method in Liu *et al.* (2009) to obtain a closed-form solution of Equation (A.8), as Lemma 2.

Lemma 2 (Liu *et al.* 2009) For a given matrix $V \in \boldsymbol{R}^{d\times n}$, the optimal solution $\boldsymbol{X}^*$ to the problem:

$$\min_{\mathbf{X}}\ \lambda\|\mathbf{X}\|_1 + \frac{1}{2}\|\mathbf{X} - \mathbf{V}\|_F^2 \tag{A.9}$$

can be expressed as $\boldsymbol{X}^* = S_\lambda[v]$, where v is an element of the matrix $\boldsymbol{V}$, and $S_\lambda\ [x]$ is a shrinkage operator defined as $\mathcal{S}_\lambda[x] = \mathrm{sgn}(x)\odot\max\{|x| - \lambda, 0\}$.

$\boldsymbol{X}$, $\boldsymbol{V}$ in Lemma 2 correspond to $\mathbf{Q}_k^{l+1}$ and $\mathbf{S}_k\tilde{\mathcal{X}}_{k(k)}^{l} - \mathbf{Y}_k^{l}/\mu_k$ in Equation (A.8), respectively. Thus, a closed-form solution to the optimization problem Equation (A.8) can be given by

$$\begin{aligned}\mathbf{Q}_k^{l+1} &:= \arg\ \min_{\mathbf{Q}_k} \frac{\beta_k}{\mu_k}\|\mathbf{Q}_k^{l+1}\|_1 + \frac{1}{2}\|\mathbf{Q}_k^{l+1} - \mathbf{J}_k^{l}\|_F^2 \\ &= \mathrm{sgn}(\mathbf{J}_k^{l})\odot\max\left\{|\mathbf{J}_k^{l}| - \frac{\beta_k}{\mu_k}, 0\right\}\end{aligned} \tag{A.10}$$

where $\mathbf{J}_k^{l} = \mathbf{S}_k\tilde{\mathcal{X}}_{k(k)}^{l} - \frac{\mathbf{Y}_k^{l}}{\mu_k}$.

c. Updating tensors $\tilde{\mathcal{X}}_k$ (k = 1,2,3)

$\tilde{\mathcal{X}}_k$s are mainly used to perform auxiliary updates to the incorporated spatiotemporal constraints, so the optimization requires a balance between the auxiliary tensor **M** (connected to the original data $\mathcal{X}_k$) and each spatiotemporal constraint matrix $\boldsymbol{S}_k$, i.e.

$$\begin{aligned}\tilde{\mathcal{X}}_k^{l+1} &:= \arg\ \min_{\tilde{\mathcal{X}}_k} \frac{\lambda_k}{2}\|\tilde{\mathcal{X}}_{k(k)}^{l+1} - \mathcal{M}_{(k)}^{l}\|_F^2 + \langle \tilde{\mathcal{X}}_{k(k)}^{l+1} - \mathcal{M}_{(k)}^{l}, \mathcal{G}_{k(k)}^{l}\rangle + \frac{\mu_k}{2}\|\mathbf{Q}_k^{l+1} - \mathbf{S}_k\tilde{\mathcal{X}}_{k(k)}^{l+1}\|_F^2 + \langle \mathbf{Q}_k^{l+1} - \mathbf{S}_k\tilde{\mathcal{X}}_{k(k)}^{l+1},\ \mathbf{Y}_k^{l}\rangle \\ &= \arg\ \min_{\tilde{\mathcal{X}}_k} \frac{\lambda_k}{2}\|\tilde{\mathcal{X}}_{k(k)}^{l+1} - (\mathcal{M}_{(k)}^{l} - \frac{\mathcal{G}_{k(k)}^{l}}{\lambda_k})\|_F^2 + \frac{\mu_k}{2}\|\mathbf{S}_k\tilde{\mathcal{X}}_{k(k)}^{l+1} - \left(\mathbf{Q}_k^{l+1} + \frac{\mathbf{Y}_k^{l}}{\mu_k}\right)\|_F^2\end{aligned} \tag{A.11}$$

Then, the optimization update for each $\tilde{\mathcal{X}}_k^{l+1}$ is

$$\begin{aligned}&\lambda_k\left(\tilde{\mathcal{X}}_{k(k)}^{l+1} - \mathcal{M}_{(k)}^{l} + \frac{\mathcal{G}_{k(k)}^{l}}{\lambda_k}\right) + \mu_k\mathbf{S}_k^T\left(\mathbf{S}_k\tilde{\mathcal{X}}_{k(k)}^{l+1} - \left(\mathbf{Q}_k^{l+1} + \frac{\mathbf{Y}_k^{l}}{\mu_k}\right)\right) = 0 \\ &\Rightarrow \tilde{\mathcal{X}}_k^{l+1} = fold_k\left((\lambda_k + \mu_k\mathbf{S}_k^T\mathbf{S}_k)^{-1}(\lambda_k\mathcal{M}_{(k)}^{l} - \mathcal{G}_{k(k)}^{l} + \mu_k\mathbf{S}_k^T\mathbf{Q}_k^{l+1} + \mathbf{S}_k^T\mathbf{Y}_k^{l})\right)\end{aligned} \tag{A.12}$$

d. Updating the auxiliary tensor M

The auxiliary tensor **M** optimization equation can be obtained from Equation (20), connecting the original data tensor $\mathcal{X}_k$ and the auxiliary tensor $\tilde{\mathcal{X}}_k$ ($k = 1,2,3$),

$$\begin{aligned}\mathcal{M}^{l+1} &:= \arg\ \min_{\mathcal{M}} \sum_{k=1}^{3}\left(\frac{\rho_k}{2}\|\mathcal{X}_{k(k)}^{l+1}-\mathcal{M}_{(k)}^{l+1}\|_F^2+\langle\mathcal{X}_{k(k)}^{l+1}-\mathcal{M}_{(k)}^{l+1},\mathcal{T}_{k(k)}^{l}\rangle+\frac{\lambda_k}{2}\|\tilde{\mathcal{X}}_{k(k)}^{l+1}-\mathcal{M}_{(k)}^{l+1}\|_F^2+\langle\tilde{\mathcal{X}}_{k(k)}^{l+1}-\mathcal{M}_{(k)}^{l+1},\mathcal{G}_{k(k)}^{l}\rangle\right)\\ &= \arg\ \min_{\mathcal{M}} \sum_{k=1}^{3}\left(\frac{\rho_k}{2}\|\mathcal{M}_{(k)}^{l+1}-(\mathcal{X}_{k(k)}^{l+1}+\frac{\mathcal{T}_{k(k)}^{l}}{\rho_k})\|_F^2+\frac{\lambda_k}{2}\|\mathcal{M}_{(k)}^{l+1}-(\tilde{\mathcal{X}}_{k(k)}^{l+1}+\frac{\mathcal{G}_{k(k)}^{l}}{\lambda_k})\|_F^2\right)\end{aligned} \tag{A.13}$$

Then the optimization update of the tensor $\mathbf{M}$ is

$$\begin{aligned}&\sum_{k=1}^{3}\left(\rho_k(\mathcal{M}_{(k)}^{l+1}-\mathcal{X}_{k(k)}^{l+1}-\frac{\mathcal{T}_{k(k)}^{l}}{\rho_k})+\lambda_k(\mathcal{M}_{(k)}^{l+1}-\tilde{\mathcal{X}}_{k(k)}^{l+1}-\frac{\mathcal{G}_{k(k)}^{l}}{\lambda_k})\right)=0\\ &\Rightarrow \mathcal{M}_{(k)}^{l+1}=\frac{1}{\sum_{k=1}^{3}(\rho_k+\lambda_k)}\sum_{k=1}^{3}\left(\rho_k\mathcal{X}_{k(k)}^{l+1}+\lambda_k\tilde{\mathcal{X}}_{k(k)}^{l+1}+\mathcal{T}_{k(k)}^{l}+\mathcal{G}_{k(k)}^{l}\right)\end{aligned} \tag{A.14}$$

e. Updating dual variables

Finally, we compute the dual variables and parameters as follows:

$$\begin{aligned}\mathcal{T}_k^{l+1} &:= \mathcal{T}_k^{l}+\rho_k\left(\mathcal{X}_k^{l+1}-\mathcal{M}^{l+1}\right),\\ \mathcal{G}_k^{l+1} &:= \mathcal{G}_k^{l}+\lambda_k\left(\tilde{\mathcal{X}}_k^{l+1}-\mathcal{M}^{l+1}\right),\\ \mathbf{Y}_k^{l+1} &:= \mathbf{Y}_k^{l}+\mu_k\left(\mathbf{Q}_k^{l+1}-\mathbf{S}_k\tilde{\mathcal{X}}_k^{l+1}\right),\end{aligned} \tag{A.15}$$

$$\begin{aligned}\rho_k^{l+1} &= \min(\delta\rho_k^{l},\rho^{\max}),\\ \lambda_k^{l+1} &= \min(\delta\lambda_k^{l},\lambda^{\max}),\\ \mu_k^{l+1} &= \min(\delta\mu_k^{l},\mu^{\max}).\end{aligned} \tag{A.16}$$

where $\delta>1$ is a constant.

So far, we have completed the step $l+1$ update of all variables. When the convergence condition in Algorithm 1 is satisfied, the iteration will terminate.