用于高速公路交通动态分析的稀疏促进反循环动态模态分解

Xudong Wang，Lijun Sun*

Department of Civil Engineering, McGill University, Montreal, QC H3A 0C3, Canada

arXiv:2302.07108v1 [stat.AP] 14 Feb 2023

摘要

从传感器网络采集的高速公路交通状态数据可视为一个高维非线性动力系统。本文提出一种新颖的数据驱动方法——稀疏促进反循环动态模态分解（circDMDsp），用于研究高速公路交通速度数据的动态特性。具体而言，circDMDsp解决了阻碍现有动态模态分解模型应用的若干问题：空间维度有限、同时存在重复模式和非重复性模式、噪声水平较高，以及模态稳定性已知。所提出的circDMDsp框架使我们能够以数值方式提取具有物理意义或可解释性的时空相干结构：动态模态反映相干空间基，相应的时间模式则捕捉这些动态模态的时间振荡与演化。基于西雅图高速公路环形线圈检测器数据得到的结果表明，交通速度数据由一组周期性成分支配，例如均值模式、日模式和周模式，且每种成分都具有独特的空间结构。通过外推时间范德蒙德矩阵，这些时空模式还可用于恢复或去噪观测数据，并预测任意时间戳的未来值。实验还表明，与其他基于DMD的模型相比，所提出的circDMDsp框架在数据重建和预测方面更加准确、稳健。

关键词： 高速公路交通动态、时空交通数据、反循环动态模态分解、交通模式

1. 引言

从高速公路网络上的传感器采集的时空交通状态数据，可以建模为高维非线性动力系统，用以在数学上刻画交通变量（例如速度、流量和密度）随时间的演变。深入理解交通感知数据背后的交通动态，有助于洞察其潜在的复杂交通现象；更重要的是，还能实现对未来交通状态的短期和长期预测，这对于支持智能交通系统（ITS）中的诸多应用至关重要。

大规模交通数据的丰富推动研究人员开发数据驱动模型，以揭示交通动态并进行预测。一种很有前景的方法是动态模态分解（DMD），最初是在流体动力学领域提出的（Schmid，2010）。DMD 基于这样一个事实：Koopman算子（Koopman和Neumann，1932）是一个无限维但线性的算子，能够准确描述非线性动力系统，并且可以通过大量有限个观测量进行良好近似。DMD 是一种无方程方法，可以将高维时空数据分解为由动态模态（空间模式）、标量振幅和时间动力学（时间模式）构成的三元组，从而从动力系统中提取相干时空结构（Schmid，2022）。与奇异值分解等其他矩阵分解方法不同，DMD 提供了一种数据驱动的谱分解，

*通讯作者。地址：加拿大魁北克省蒙特利尔市西舍布鲁克街492-817号，麦克唐纳工程大楼，邮编 H3A 0C3

Email addresses: xudong.wang2@mail.mcgill.ca（Xudong Wang），lijun.sun@mcgill.ca（Lijun Sun）

而 DMD 识别出的时空模式具有明确的物理意义——动态模态揭示空间基的结构，时间动力学则描述相应动态模态的振荡模式（Brunton 等人，2016b）。这种相干时空结构提供了一种解析解，可用于数据重建、系统诊断以及长期外推/预测。

动态模态分解及其变体在分析流体动力学之外多个领域的时空数据方面取得了巨大成功，例如神经记录、高分辨率视频和环境数据等（Brunton 等人，2016a；Kutz 等人，2016b；Erichson 等人，2019）。然而，直接将动态模态分解应用于时空交通数据分析仍面临若干重大挑战。首先，动态模态分解通常要求空间维度远大于时间维度，才能找到最佳的线性动力系统近似。然而，鉴于交通数据是在较长时间内从数量有限的传感器（例如，规模为 100 至 1000 个）持续采集的，真实动力学往往高度非线性且过于复杂，难以用线性动力系统近似。其次，现实世界的交通状态数据既包含重复模式（例如早高峰拥堵），也包含非重复模式（例如意外事故/事件引发的拥堵）（Li 等人，2015）。一般而言，重复/周期性数据可由具有长程周期相关性的低频模式捕捉，而异常及非重复数据则可视为仅具有短程相关性的高频变化（即噪声和脉冲信号）。由于动态模态分解主要用于提取相干的时空模式，非重复数据的存在可能会严重污染动态模态分解所估计的模式。第三，由于人类出行固有的节律，交通系统呈现日周期和周周期。交通数据中普遍存在的季节性/周期性提供了很强的先验信息，应当作为一项约束——我们确实期望检测到的模态保持稳定（即既不增长也不衰减）。然而，现有的动态模态分解模型无法保证模态稳定性。

在现有文献中，只有少数研究使用基于DMD的模型分析交通动态。Liu 等人(2016) 首次使用DMD对高速公路交通探测数据进行建模，结果揭示了不同时间窗口（例如早高峰/晚高峰）下交通动态的异质性。Avila 和 Mez´ic (2020) 使用 Hankel-DMD（Brunton 等人，2016a；Arbabi 和 Mezic，2017）分析多车道高速公路交通动力学。Lehmberg 等人(2021) 使用结合时间延迟嵌入和扩散映射算法的扩展动态模态分解（Williams 等人，2015）研究人群动力学。Wang 和 Sun(2022) 使用 Hankel-DMD 和层次聚类研究广州地铁系统的动态出行模式。上述研究均通过“Hankel化”应对空间维度不足这一首要挑战，即利用延迟嵌入扩大数据的空间维度。这些研究表明，Hankel-DMD 确实能够提高可解释性，并提供更准确的数据重构和预测。然而，Hankel-DMD 仍无法将强周期性编码为先验信息，而且容易受到分析数据中非重复观测值和离群值的影响；现场传感器采集的交通状态数据便存在这种情况。

本研究提出了一种新颖的稀疏促进反循环动态模态分解（circDMDsp）框架，旨在提升以重复模式为主、但同时含有大量噪声、离群值和非重复观测值的时空数据集的“稳健性”和“模态稳定性”。circDMDsp框架包含两项关键特征／设计。首先，我们提出在扩展空间维度时使用反循环矩阵，而非Hankel矩阵。这种扩增与离散傅里叶变换（DFT）密切相关——离散傅里叶变换矩阵是一种特定的范德蒙德矩阵，其特征值为单位根，因此循环扩增能够有效促进模态稳定性。其次，我们引入稀疏促进策略，从大量候选模式中选出主导模式；此外，还引入最优硬阈值（Gavish和Donoho，2014），以确定最优矩阵秩，使所提出的模型能够轻松推广至不同的交通网络，而无需调整许多参数。所提出的circDMDsp框架专为季节性强的系统设计，因此适用于分析高速公路交通状态数据。本研究有三个目标：（1）发现动态交通模式

（第4.3节）；（2）基于揭示的时空模式重建历史交通速度并预测未来值（第4.2节）；以及（3）使用DMD量化传感器层级和网络层级交通状态数据的短期可预测性。

本文其余部分安排如下。第2 节介绍符号说明、Koopman谱分析和D的原理，这些是所提出框架的基础。本节还介绍一些相关的DMD变体。第3节介绍所提出的促进稀疏性的反循环动态模态及其实现与应用。第4节展示一项基于西雅图高速公路交通数据的真实案例研究。我们将重构和预测性能与基线模型进行比较，并分析所发现的时空模式。此外，我们还讨论传感器层级和网络层级的短期可预测性。第5节总结本研究并讨论未来的研究方向。

2. 预备知识

2.1. 符号说明

全文中，我们用字母表示标量，例如 $x \in \mathbb{R}$ 或 $X \in \mathbb{R}$；用小写粗体字母表示向量，例如 $\boldsymbol{x} \in \mathbb{R}^n$；用大写粗体字母表示矩阵，例如 $\boldsymbol{X} \in \mathbb{R}^{n_1 \times n_2}$。矩阵的第 (i, j) 个元素记为 $\boldsymbol{X}_{i,j}$。 $\boldsymbol{X}$ 的弗罗贝尼乌斯范数定义为 $\|\boldsymbol{X}\|_F = \sqrt{\sum_{i=1}^{n_1} \sum_{j=1}^{n_2} \boldsymbol{X}_{i,j}^2}$，而 $\|\boldsymbol{x}\|_1 = \sum_{i=1}^{n} |x_i|$的 l_1-范数为。

*2.2. Koopman*分析

本节简要介绍 Koopman谱分析的基础理论，它构成了动态模态分解(2016a) 的基础。关于Koopman分析理论及其与动态模态分解之间联系的更多细节，请参阅Schmid (2022)。对于一个非线性离散时间动力系统，时刻 $\boldsymbol{x}_t \in \mathbb{R}^N$ 在状态空间 $\mathcal{M}$ 的状态 t 可表示为：

$$\boldsymbol{x}_{t+1} = \boldsymbol{f}(\boldsymbol{x}_t), \tag{1}$$

其中， $\boldsymbol{f}(\cdot)$描述系统的非线性动力学，例如状态的常微分方程，它将状态 $\boldsymbol{x}_t$ 映射到给定时间间隔 Δt 之后的 $\boldsymbol{x}_{t+1}$。

分析（1）的传统方法是研究平衡点以及小扰动的线性动力学，前提是 $\boldsymbol{f}$ 能够通过守恒定律和物理模型表述（Schmid，2022）。然而，在现实系统中，动力学 $\boldsymbol{f}$ 往往难以或无法实际建模。Koopman分析旨在从数据中提取关键的相干结构，而无需了解 $\boldsymbol{f}$。Koopman分析并非基于状态变量 $\boldsymbol{x}_t$，而是依托于状态 $\boldsymbol{g}(\boldsymbol{x}_t)$的观测量： $\mathcal{M} \to \mathbb{C}$ ，其位于可观测空间 $\mathcal{G}$中。一个称为Koopman算子的无限维线性算子 $\mathcal{K}$ 作用于观测函数$\boldsymbol{g}$，使得 $\mathcal{K}\boldsymbol{g} = \boldsymbol{g} \circ \boldsymbol{f}$，其中 $\circ$ 表示 $\boldsymbol{g}$ 与 $\boldsymbol{f}$ 的复合（Brunton 等人，2016b）。借助Koopman算子 $\mathcal{K}$，（1）中 $\boldsymbol{x}_t$ 上的非线性动力学分析等价于 $\boldsymbol{g}(\boldsymbol{x}_t)$上的一个无限维线性问题：

$$\mathcal{K}\boldsymbol{g}(\boldsymbol{x}_t) = \boldsymbol{g}(\boldsymbol{f}(\boldsymbol{x}_t)) = \boldsymbol{g}(\boldsymbol{x}_{t+1}). \tag{2}$$

图1 展示了如何通过观测函数将以状态变量表示的非线性系统转换为以观测量表示的线性系统的过程 $\boldsymbol{g}(\boldsymbol{x}_t)$，并通过一个简单的非线性动力系统阐释了 Koopman分析的关键概念。给定合适的观测量（例如图 1B），非线性动力系统可以等价地描述为一个线性系统。

为确定此类观测量 ，通过对 $\mathcal{K}$进行特征分解来考察 $\mathcal{K}$ 的不变子空间 ：

$$\mathcal{K}\varphi_k(\boldsymbol{x}) = \lambda_k \varphi_k(\boldsymbol{x}), \quad k \in \mathbb{N}, \tag{3}$$

其中， $\lambda_k \in \mathbb{C}$ 是第 k个特征值（*Koopman*特征值）， $\varphi_k : \mathcal{M} \to \mathbb{C}$ 是对应的特征函数（*Koopman*特征函数）。Koopman特征函数可视为一组基

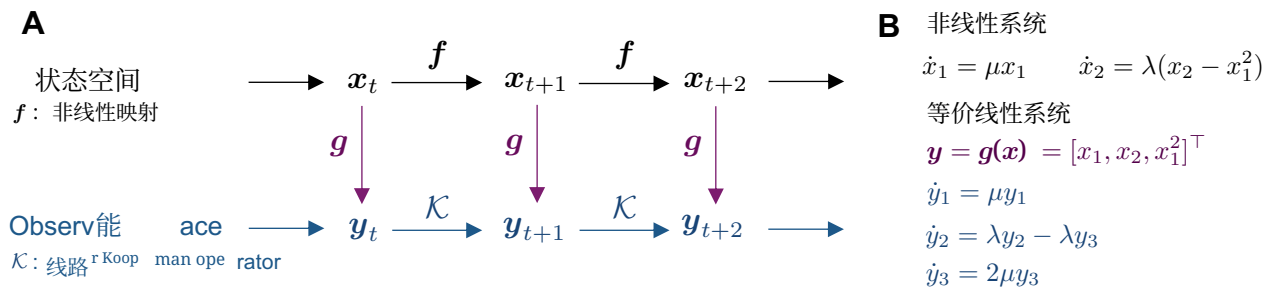

图1：A：非线性动力系统的 Koopman 算子示意图；B：将 Koopman 算子应用于非线性动力系统的示例。

函数（内在测量坐标）；如果 $\boldsymbol{g}$ 位于特征函数张成的空间内，则观测量可以表示为这些特征函数的线性组合，例如，$\boldsymbol{g}(\boldsymbol{x}_t) = \sum_{k=1}^{\infty} c_k \varphi_k(\boldsymbol{x}_t)$。在实践中，难以从无限维Koopman算子$\mathcal{K}$中求得特征函数和特征值。

2.3. 动态模态分解及其变体

动态模态分解是一种数据驱动方法，当观测量（数据的空间维度）趋于无穷时，用于近似无穷维 Koopman 算子的特征值和特征向量（Korda 和 Mezi´c，2018）。动态模态分解旨在通过求解最小二乘问题，找到这些主导特征向量 $\boldsymbol{\Phi}$ 和特征值 $\boldsymbol{\lambda}$，它们属于能够最佳拟合所有数据 $\boldsymbol{A}$的线性算子 $\boldsymbol{X} \in \mathbb{R}^{N\times T}$：

$$\min_{\boldsymbol{A}} \|\boldsymbol{X}_2 - \boldsymbol{A}\boldsymbol{X}_1\|_F^2. \tag{4}$$

其中 $\boldsymbol{X}_1 = [\boldsymbol{x}_1, \cdots, \boldsymbol{x}_{T-1}] \in \mathbb{R}^{N\times(T-1)}$，$\boldsymbol{X}_2 = [\boldsymbol{x}_2, \cdots, \boldsymbol{x}_T] \in \mathbb{R}^{N\times(T-1)}$。在动力系统中，动态矩阵的特征向量（动态模态）反映空间相关性，相应的特征值则决定这些模态的时间动力学，包括振荡频率和衰减/增长率。在实际应用中，状态维度 N 较大，直接求解最优问题（4）难以实现。为克服这一挑战，动态模态分解通过正交分解（POD）投影，对 $\boldsymbol{A}$进行降维表示，得到˜，而非直接对 $\boldsymbol{A}$进行特征分解。算法 1 概述了动态模态分解的流程，更多细节见第 3.1节。

数据： $\boldsymbol{X} \in \mathbb{R}^{N\times T}$**结果：** $\boldsymbol{\Phi} \in \mathbb{C}^{N\times r}$, $\boldsymbol{b} \in \mathbb{C}^{r\times r}$, $\boldsymbol{\Psi} \in \mathbb{C}^{r\times T}$步骤 1：根据数据 $\boldsymbol{X}$定义矩阵 $\boldsymbol{X}_1$ 和 $\boldsymbol{X}_2$；步骤 2：对矩阵 $\boldsymbol{X}_1$进行约化奇异值分解，计算 $\boldsymbol{U}_r$，$\boldsymbol{\Sigma}_r$和 $\boldsymbol{V}_r$，秩为 r，即$\boldsymbol{X}_1 \approx \boldsymbol{U}_r\boldsymbol{\Sigma}_r\boldsymbol{V}_r^{\star}$; ˜步骤 3：将 $\boldsymbol{A}$ 投影到由 $\boldsymbol{U}_r$ 张成的 POD 子空间，得到低秩动力学矩阵 $\boldsymbol{A}$: $\tilde{\boldsymbol{A}} = \boldsymbol{U}_r^{\top}\boldsymbol{A}\boldsymbol{U}_r = \boldsymbol{U}_r^{\top}\boldsymbol{X}\boldsymbol{V}_r\boldsymbol{\Sigma}_r^{-1}$; ˜步骤 4：对 $\boldsymbol{A}$ 进行特征分解，求得特征向量（动态模态）$\boldsymbol{\Phi}$ 和特征值 $\boldsymbol{\lambda}$：$\tilde{\boldsymbol{A}}\boldsymbol{W} = \boldsymbol{W}\mathrm{diag}(\boldsymbol{\lambda})$ 和 $\boldsymbol{\Phi} = \boldsymbol{X}_2\boldsymbol{V}_r\boldsymbol{\Sigma}_r^{-1}\boldsymbol{W}$；步骤 5：根据初始条件 $\boldsymbol{x}_1$ 和动态模态 $\boldsymbol{\Phi}$计算振幅 $\boldsymbol{b}$：$\boldsymbol{b} = \boldsymbol{\Phi}^{\dagger}\boldsymbol{x}_1$。步骤 6：构造范德蒙德矩阵 $\boldsymbol{\Psi}$，其每一行的特征值构成几何级数。**算法1：** 动态模态分解（动态模态分解）

与奇异值分解（SVD）类似，其中 $\boldsymbol{X} = \boldsymbol{U}\boldsymbol{S}\boldsymbol{V}^{\top}$，数据 $\boldsymbol{X}$ 也通过动态模态分解分解为三个低秩矩阵，使得 $\boldsymbol{X} = \boldsymbol{\Phi}\mathrm{diag}(\boldsymbol{b})\boldsymbol{\Psi}$，其中 $\boldsymbol{\Psi}$ 是包含特征值的范德蒙德矩阵。动态模态分解与奇异值分解的主要区别在于，在奇异值分解中，低秩矩阵 $\boldsymbol{U}$ 和 $\boldsymbol{V}$

是正交的，因此每个独立模式都是用于表示数据的混合模式。相比之下，动态模态分解引入了动态约束，促使数据分解为一组振荡频率和模态。模态按照各自的振荡频率演化。然而，在动态模态分解中，范德蒙德矩阵 $\boldsymbol{\Psi}$ 的元素 $\boldsymbol{\lambda}$ 不一定是单位根，这可能导致演化不稳定。具体而言，具有增长演化特征的模态会随着窗口大小的增加而发散至无穷大，而具有衰减演化特征的动态模态最终会收敛至零。

动态模态分解最初应用于流体流动领域，在该领域中，流体流动的空间维度通常远大于时间维度，即 $N \gg T$。然而，在交通网络的背景下，数百个传感器或站点每周采集数千个交通变量的情况十分常见。因此，在研究时间窗口内得到的数据矩阵 $\boldsymbol{X}$ 通常具有 $N < T$。在这种情况下，真实动力学具有高度非线性，而仅由 $\boldsymbol{\Psi}$ 和 r 个参数表示的时间演化低秩近似，难以充分表征复杂数据。因此，使用简单的线性动力系统来准确近似 $\boldsymbol{X}$ 中的动力学颇具挑战。实现线性近似最常见的方法是进行数据增强。汉克尔动态模态分解（HankelDMD）对数据的汉克尔变换结果应用动态模态分解（Brunton 等人，2016a；Avila 和 Mezi´c，2020；Lehmberg 等人，2021；Wang 等人，2021）。给定一个 $N \times T$ 矩阵 $\boldsymbol{X}$，延迟嵌入长度为 τ 的 Hankel矩阵 $\boldsymbol{H} \in \mathbb{R}^{N\tau \times (T-\tau+1)}$ 定义如下：

$$\boldsymbol{H} = \begin{bmatrix} \boldsymbol{x}_1 & \boldsymbol{x}_2 & \dots & \boldsymbol{x}_{T-\tau+1} \\ \boldsymbol{x}_2 & \boldsymbol{x}_3 & \dots & \boldsymbol{x}_{T-\tau+2} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{x}_\tau & \boldsymbol{x}_{\tau+1} & \cdots & \boldsymbol{x}_T \end{bmatrix} = \begin{bmatrix} \boldsymbol{h}_1 & \dots & \boldsymbol{h}_{T-\tau+1} \end{bmatrix}. \tag{5}$$

其中，空间维度从 N增加到 $N\tau$。随后，可以对 $\boldsymbol{H}$ 而非原始数据 $\boldsymbol{X}$进行默认的动态模态分解分析。

应用动态模态分解时的另一个关键问题是，现场传感器网络生成的真实世界数据往往包含大量噪声、离群值以及非重复模式（例如，由轻微事件/事故引起的短期交通拥堵），这些现象很难用主导动力学原理解释。默认的动态模态分解算法对噪声和离群值非常敏感；更重要的是，非重复时间数据的存在会导致对振幅和周期性方面的时间演化产生有偏估计。˜为了解决这一问题，研究者提出了去偏算法，通过定义 $\boldsymbol{A}$ 为 $\boldsymbol{A}$ 在 POD子空间中。例如，前向-后向动态模态分解（fbDMD）（Dawson 等人，2016）分别对前向动力学 $\boldsymbol{X}_2 \approx \boldsymbol{A}_f \boldsymbol{X}_1$ 和反向动力学 $\boldsymbol{X}_1 \approx \boldsymbol{A}_b \boldsymbol{X}_2$进行建模，并采用两个不同的传播算子：

$$\tilde{\boldsymbol{A}}_f = \boldsymbol{U}_1^\top \boldsymbol{X}_2 \boldsymbol{V}_1 \boldsymbol{\Sigma}_1^{-1}, \quad \tilde{\boldsymbol{A}}_b = \boldsymbol{U}_2^\top \boldsymbol{X}_1 \boldsymbol{V}_2 \boldsymbol{\Sigma}_2^{-1}, \tag{6}$$

其中，$\boldsymbol{X}_1 \approx \boldsymbol{U}_1 \boldsymbol{\Sigma}_1 \boldsymbol{V}_1$ 和 $\boldsymbol{X}_2 \approx \boldsymbol{U}_2 \boldsymbol{\Sigma}_2 \boldsymbol{V}_2$ 分别是 $\boldsymbol{X}_1$ 和$\boldsymbol{X}_2$的约化奇异值分解（SVD）。那么，前向传播算子的去偏估计为

$$\tilde{\boldsymbol{A}} = \left(\tilde{\boldsymbol{A}}_f \tilde{\boldsymbol{A}}_b^{-1} \right)^{1/2}. \tag{7}$$

˜由于平方根不唯一，$\boldsymbol{A}$ 的特征值由平方根确定，存在一个 ± 1 的因子（Askham 和 Kutz，2018）。另一种广为人知的去偏 DMD 是全最小二乘DMD（tlsDMD）（Hemati 等人，2017），其旨在寻找噪声-free 数据在 $\boldsymbol{X}_1$ 和 $\boldsymbol{X}_2$之间的精确线性关系：

$$\begin{aligned} \min_{\boldsymbol{A}, \Delta\boldsymbol{X}_1, \Delta\boldsymbol{X}_2} & \left\| \begin{bmatrix} \Delta\boldsymbol{X}_1 \\ \Delta\boldsymbol{X}_2 \end{bmatrix} \right\|_F \\ \text{s.t.} \quad & \boldsymbol{X}_2 - \Delta\boldsymbol{X}_2 = \boldsymbol{A}(\boldsymbol{X}_1 - \Delta\boldsymbol{X}_1). \end{aligned} \tag{8}$$

优化问题（8）可通过将 $\boldsymbol{X}_1$ 和 $\boldsymbol{X}_2$ 投影到由增广矩阵张成的 POD 模态上来求解

$$\boldsymbol{Z} = \begin{pmatrix} \boldsymbol{X}_1 \\ \boldsymbol{X}_2 \end{pmatrix}. \tag{9}$$

对 $\boldsymbol{Z}$, $\boldsymbol{Z} \approx \boldsymbol{U}_z \boldsymbol{\Sigma}_z \boldsymbol{V}_z^*$ 进行约化奇异值分解，并取前 z 个特征值。投影后的矩阵

$$\bar{\boldsymbol{X}}_1 = \boldsymbol{X}_1 \mathbb{P}_{\boldsymbol{Z}^*}, \quad \bar{\boldsymbol{X}}_2 = \boldsymbol{X}_2 \mathbb{P}_{\boldsymbol{Z}^*}, \tag{10}$$

其中，$\mathbb{P}_{\boldsymbol{Z}^*} = \boldsymbol{V}_z \boldsymbol{V}_z^*$用于替代动态模态分解过程中的原始数据。

3. 稀疏促进反循环动态模态分解（circDMDsp）

如前所述，动态模态分解中的范德蒙德矩阵可能产生演化发散（增长）或收敛（衰减）的模态，从而对长期交通预测的准确性产生负面影响。为缓解这一问题，我们提出了一种反循环矩阵，其设计旨在与离散傅里叶变换（DFT）建立理想的联系；其中，离散傅里叶变换矩阵是一种特殊的范德蒙德矩阵，其元素取为单位根。所提出的反循环动态模态分解（circDMD）利用了反循环矩阵可由离散傅里叶变换对角化这一性质。该框架包含四个关键步骤，如图2所示。首先，我们将原始交通数据整理为矩阵 $\boldsymbol{X}$ （传感器 $\times$ 时间点），然后使用循环数据增强算子 $\mathcal{C}_\tau$将其转换为反循环矩阵 $\boldsymbol{C}$ 。其次，我们假设增强后的数据在任意时间戳 t 均遵循局部线性系统，使得 $\boldsymbol{c}_{t+1} \approx \boldsymbol{\mathcal{A}} \boldsymbol{c}_t$，其中$\boldsymbol{\mathcal{A}}$ 是一个 $N\tau \times N\tau$ 动态矩阵。我们使用动态模态分解估计最优动态矩阵$\boldsymbol{\mathcal{A}}$ ，由此得到动态模态 $\boldsymbol{\Phi}$、时间演化 $\boldsymbol{\Psi}$ 及相应的振幅 $\boldsymbol{D}_b$。第三，我们应用稀疏促进动态模态分解方法（DMDSP）（Jovanović 等人，2014），通过将不重要模态的振幅设为零，获得简洁的表示。基于优化后的振幅，我们随后选取主导模态，并利用这些模态分析交通网络的动态特性、重构历史数据以及预测未来值。最后，我们使用反循环逆算子 $\mathcal{C}_\tau^{-1}$将反循环空间中的估计结果（重构和预测）还原到原始空间。由于我们的模型确保了单位根，因此它为周期性数据提供了一种自然的解决方案，并且对异常值数据更加稳健。

3.1. 反循环动态模态分解（*circDMD*）

首先，定义算子 circshift($\boldsymbol{X}$，L)，该算子将 $\boldsymbol{X}$ 的每一列向右循环移位 L 个位置（L 为负）或向左循环移位 L 个位置（L 为正）。随后，通过反循环算子 $\mathcal{C}_\tau(\boldsymbol{X})$ 构造反循环矩阵 $\boldsymbol{C}$： $\mathbb{R}^{N\times T} \Rightarrow \mathbb{R}^{N\tau\times T}$ 作用于 $\boldsymbol{X}$，其中 $i = 1$, ..., τ，其形式为:

$$\mathcal{C}_\tau(\boldsymbol{X}) := \boldsymbol{C}_{(i-1)N+1:iN,\,:} = \operatorname{circshift}(\boldsymbol{X}, -i) = \begin{bmatrix} \boldsymbol{x}_2 & \boldsymbol{x}_3 & \cdots & \boldsymbol{x}_T & \boldsymbol{x}_1 \\ \boldsymbol{x}_3 & \boldsymbol{x}_4 & \cdots & \boldsymbol{x}_1 & \boldsymbol{x}_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \boldsymbol{x}_{\tau+1} & \boldsymbol{x}_{\tau+2} & \cdots & \boldsymbol{x}_{\tau-1} & \boldsymbol{x}_\tau \end{bmatrix} = \left[\boldsymbol{c}_2, \boldsymbol{c}_3, \ldots, \boldsymbol{c}_T, \boldsymbol{c}_1\right]. \tag{11}$$

L设一个 $T\times T$ 的置换矩阵为 $\boldsymbol{P}_T = \boldsymbol{I}_T$（: p） $p = [T,\ 1:\ T-1]$，则有

$$\boldsymbol{C}\boldsymbol{P}_T = \begin{bmatrix} \boldsymbol{x}_1 & \boldsymbol{x}_2 & \cdots & \boldsymbol{x}_{T-1} & \boldsymbol{x}_T \\ \boldsymbol{x}_2 & \boldsymbol{x}_3 & \cdots & \boldsymbol{x}_T & \boldsymbol{x}_1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \boldsymbol{x}_\tau & \boldsymbol{x}_{\tau+1} & \cdots & \boldsymbol{x}_{\tau-2} & \boldsymbol{x}_{\tau-1} \end{bmatrix} = \begin{bmatrix} \boldsymbol{c}_1 & \ldots & \boldsymbol{c}_T \end{bmatrix}. \tag{12}$$

如我们所见，反循环矩阵的构造与式（5）中的“Hankel化”非常相似。然而，值得注意的是，式（12）所描述的反循环矩阵是一种循环数据结构，可以使用离散傅里叶变换（DFT）进行对角化。当 $\tau = T$ 时，$\mathcal{C}_T(\boldsymbol{X})\boldsymbol{P}_T$的结果变为一个完整的分块反循环矩阵，因此有（Yamamoto 等人，2022）

$$\mathcal{C}_T(\boldsymbol{X})\boldsymbol{P}_T = (\boldsymbol{I}_N \otimes \boldsymbol{F}_T) \begin{bmatrix} \boldsymbol{\Theta}_1 \\ \boldsymbol{\Theta}_2 \\ \vdots \\ \boldsymbol{\Theta}_N \end{bmatrix} \boldsymbol{F}_T, \tag{13}$$

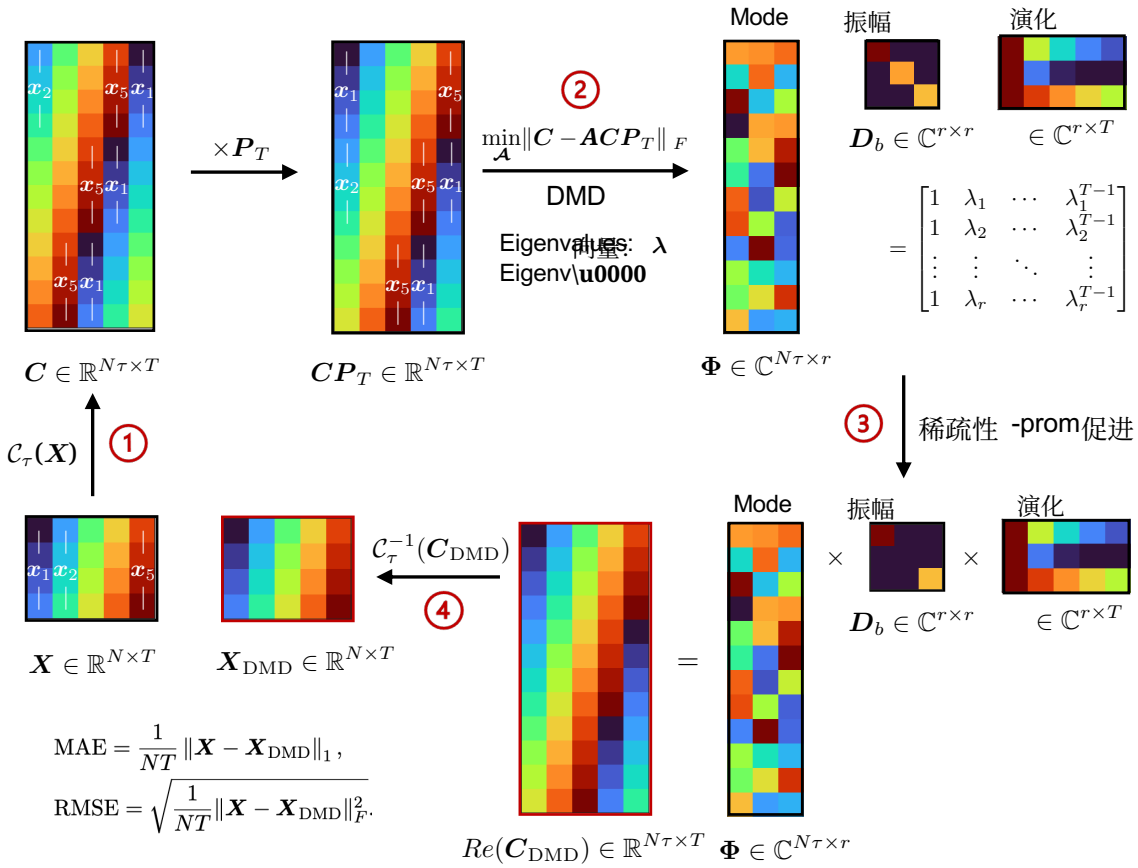

图2：当 $\tau = 3$ 时，对于 4×5 矩阵 $\boldsymbol{X}$，所提出的 circDMDsp 的示意图。该模型包括四个步骤：(1) 应用反循环算子以增加观测量的数量；(2) 使用动态模态分解将反循环矩阵分解为动态模态、振幅和时间演化；(3) 采用稀疏性促进方法（DMDSP）以获得简洁模型；(4) 利用反循环结构进行历史数据重构和未来值预测。

其中，$\boldsymbol{F}_T \in \mathbb{C}^{T\times T}$是离散傅里叶矩阵，$\boldsymbol{F}_T^* \in \mathbb{C}^{T\times T}$是共轭离散傅里叶矩阵，$\boldsymbol{\Theta}_i = \mathrm{diag}(\frac{1}{T}\boldsymbol{F}_T^*\boldsymbol{X}_{i,:}) \in \mathbb{C}^{T\times T}$表示 $\boldsymbol{X}$的第 i行。设 $\boldsymbol{Q}_\tau = [\boldsymbol{I}_\tau, \boldsymbol{0}_{\tau,T-\tau}] \in \mathbb{R}^{\tau\times T}$为截断矩阵，则对于延迟嵌入长度 τ，有

$$\mathcal{C}_\tau(\boldsymbol{X})\boldsymbol{P}_T = (\boldsymbol{I}_N \otimes (\boldsymbol{Q}_\tau\boldsymbol{F}_T)) \begin{bmatrix} \boldsymbol{\Theta}_1 \\ \boldsymbol{\Theta}_2 \\ \vdots \\ \boldsymbol{\Theta}_N \end{bmatrix} \boldsymbol{F}_T, \tag{14}$$

式（14）表明，$\boldsymbol{C}\boldsymbol{P}_T$ 的右半部分是一个离散傅里叶矩阵，其特征值为单位根，从而消除了变化模态。因此，额外的数据使模型能够更稳健地刻画关键动态模态；这些模态产生稳定的振荡，与交通状态数据潜在的日/周周期性相匹配。

在反循环动态模态分解（cricDMD）中，我们假设交通系统是一个局部线性动力系统，因此 $\boldsymbol{c}_{t+1} \approx \boldsymbol{\mathcal{A}}\boldsymbol{c}_t$，其中 $\boldsymbol{c}_t \in \mathbb{R}^{N\tau}$ 是 $\boldsymbol{C}\boldsymbol{P}_T$ 的 t 列，而 $\boldsymbol{\mathcal{A}} \in \mathbb{R}^{N\tau\times N\tau}$ 是动态矩阵。于是，circDMD 的目标转变为通过求解，寻找与$\boldsymbol{C}\boldsymbol{P}_T$ 和 $\boldsymbol{C}$ 相关的$\boldsymbol{\mathcal{A}}$ 的主导特征分解

$$\min_{\boldsymbol{\mathcal{A}}} \|\boldsymbol{C} - \boldsymbol{\mathcal{A}}\boldsymbol{C}\boldsymbol{P}_T\|_F. \tag{15}$$

而解为：

$$\boldsymbol{\mathcal{A}} = \boldsymbol{C}(\boldsymbol{C}\boldsymbol{P}_T)^\dagger, \tag{16}$$

其中 $\dagger$ 表示摩尔–彭若斯逆，可通过奇异值分解计算，即 $(\boldsymbol{C}\boldsymbol{P}_T)^\dagger \approx \boldsymbol{V}\boldsymbol{\Sigma}^{-1}\boldsymbol{U}^\top$。然而，当维度 $\boldsymbol{\mathcal{A}}$ 从式 (16) 中直接分析时，如果 $N\tau$ 和/或 T 较大，则难以处理。提出动态模态分解是为了解决维度问题：将 $\boldsymbol{\mathcal{A}}$ 投影到由前 r个左奇异向量 $\boldsymbol{U}_r \in \mathbb{R}^{N\tau\times r}$ (Tu, 2013) 张成的低维空间，即 POD子空间：

$$\tilde{\boldsymbol{\mathcal{A}}} = \boldsymbol{U}_r^\top \boldsymbol{\mathcal{A}} \boldsymbol{U}_r = \boldsymbol{U}_r^\top \boldsymbol{C}\boldsymbol{V}_r \boldsymbol{\Sigma}_r^{-1}, \tag{17}$$

其中 $\boldsymbol{\Sigma}_r \in \mathbb{R}^{r\times r}$ 和 $\boldsymbol{V}_r \in \mathbb{R}^{T\times r}$， r 表示$\boldsymbol{\mathcal{A}}$的秩。

矩阵$\tilde{\boldsymbol{\mathcal{A}}}$和$\boldsymbol{\mathcal{A}}$相似变换相关联，并且具有相同的非零主导个特征值，因此$\boldsymbol{\mathcal{A}}$ 的特征值 $\boldsymbol{\lambda} \in \mathbb{C}^r$ 可以由$\tilde{\boldsymbol{\mathcal{A}}}$计算得到：

$$\tilde{\boldsymbol{\mathcal{A}}}\boldsymbol{W} = \boldsymbol{W}\,\mathrm{diag}(\boldsymbol{\lambda}), \tag{18}$$

其中， $\boldsymbol{W} \in \mathbb{C}^{r\times r}$ 的列是 $\tilde{\boldsymbol{\mathcal{A}}}$的特征向量，则 $\boldsymbol{\Phi} \in \mathbb{C}^{N\tau\times r}$ A 的特征向量可表示为（Kutz 等，2016a）：

$$\boldsymbol{\Phi} := \boldsymbol{C}\boldsymbol{V}_r\boldsymbol{\Sigma}_r^{-1}\boldsymbol{W} \quad \text{(exact dynamic modes)} \quad \text{or} \quad \boldsymbol{\Phi} := \boldsymbol{U}_r\boldsymbol{W} \quad \text{(projected dynamic modes)}, \tag{19}$$

如果 $\boldsymbol{C}$ 和 $\boldsymbol{C}\boldsymbol{P}_T$ 具有相同的列空间（Kutz 等，2016a），则这两种定义将趋于一致。

动态模态分解的优势在于，它能够通过动态矩阵的特征值和特征向量来解读数据背后的物理现象。动态模态（即特征向量） $\boldsymbol{\Phi}$ 描绘了数据中的空间相干结构，而相应的特征值则定义了每个模态的增长/衰减率和振荡频率（Tu，2013）。对于每个特征对（ϕ_i， λ_i），特征值的幅值 $|\lambda_i|$ 决定了相应的动态模态 ϕ_i 是稳态（当 $|\lambda_i| = 1$）还是变化的（当 $|\lambda_i| < 1$时衰减，当 $|\lambda_i| > 1$时增长），而 λ_i 的虚部决定了 ϕ_i的振荡周期。因此，动态模态分解将空间降维技术与时域中的傅里叶变换相结合（Kutz 等，2016a）。

利用检测到的特征值 λ_i 和$\boldsymbol{\mathcal{A}}$的特征向量 ϕ_i，任意时间戳 t处的状态变量 $\boldsymbol{c}_t$ 可近似表示为

$$\boldsymbol{c}_t \approx \sum_{i=1}^{r} b_i \boldsymbol{\phi}_i \lambda_i^{t-1} \tag{20}$$

其中， $\boldsymbol{b}$是利用初始条件 $\boldsymbol{c}_1 = \boldsymbol{C}(:T)$计算得到的模态振幅 $\boldsymbol{b} = \boldsymbol{\Phi}^\dagger \boldsymbol{c}_1$。此外，也可以使用连续时间公式外推任意给定时间 t 处的值

$$\boldsymbol{c}(t) \approx \sum_{i=1}^{r} b_i \boldsymbol{\phi}_i e^{\log(\lambda_i)(t-1)/\Delta t}. \tag{21}$$

记 $\boldsymbol{C}_{\mathrm{DMD}} \in \mathbb{C}^{N\tau\times T}$ 为通过 circDMD 恢复的矩阵，则表达式 (20)可写成矩阵形式

$$\boldsymbol{C}_{\mathrm{DMD}} = \begin{bmatrix}\boldsymbol{\phi}_1 & \boldsymbol{\phi}_2 & \dots & \boldsymbol{\phi}_r\end{bmatrix} \begin{bmatrix} b_1 & & & \\ & b_2 & & \\ & & \ddots & \\ & & & b_r \end{bmatrix} \begin{bmatrix} 1 & \lambda_1 & \cdots & \lambda_1^{T-1} \\ 1 & \lambda_2 & \cdots & \lambda_2^{T-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \lambda_r & \cdots & \lambda_r^{T-1} \end{bmatrix} = \boldsymbol{\Phi}\boldsymbol{D}_b\boldsymbol{\Psi}, \tag{22}$$

其中 $\boldsymbol{D}_b$ 是振幅的对角矩阵 $\boldsymbol{b}$，而 $\boldsymbol{\Psi} \in \mathbb{C}^{r\times T}$ 是由表示时间动力学的特征值构成的范德蒙德矩阵。通过调整范德蒙德矩阵$T_\xi > T$中的时间窗口长度T_ξ，即可实现系统预测（ $\boldsymbol{\Psi}$）。实际应用中，我们将 $\boldsymbol{C}_{\mathrm{DMD}}$ 的实部作为估计值。由于式（22）中的重构是在延迟嵌入空间中进行的，因此还需要将其变换回原始空间。我们不使用每个模态（Brunton 等人，2016a；Avila 和的前 N 个元素 $\boldsymbol{D}_b$

Mezi´c，2020），而是利用所有 $N\tau$ 模态进行变换。我们定义反循环逆算子 $\mathcal{C}_\tau^{-1}(\boldsymbol{C})$： $\mathbb{R}^{N\tau\times T}\Rightarrow\mathbb{R}^{N\times T}$，即 $\mathcal{C}_\tau^{-1}(\mathcal{C}_\tau(\boldsymbol{X}))=\boldsymbol{X}$:

$$\mathcal{C}_\tau^{-1}(\boldsymbol{C}):=\sum_{i=1}^{\tau}\text{circshift}\left(\boldsymbol{C}_{(i-1)N+1:iN,\,:},i-1\right)/\tau. \tag{23}$$

3.2. 通过促进稀疏性改进动态模态分解

在许多情况下，秩 r 可能较大，从而导致模型表示冗长。我们采用促进稀疏性的动态模态分解（DMDSP）方法（Jovanovi´c 等人，2014），基于 circDMD 的结果进一步得到简洁的表达式。我们的目标是仅选取最重要的模式，这可以通过以下优化问题实现：

$$\min_{\boldsymbol{b}}\|\boldsymbol{C}\boldsymbol{P}_T-\boldsymbol{\Phi}\boldsymbol{D}_b\boldsymbol{\Psi}\|_F+\gamma\|\boldsymbol{b}\|_1, \tag{24}$$

其中， γ 是一个用于控制模态振幅 $\boldsymbol{b}$稀疏程度的参数—— γ 的值越大， $\boldsymbol{b}$中的零项就越多。DMDSP 算法包括两个步骤：

(1) 确定稀疏结构，通过交替方向乘子法（ADMM）框架（即问题（24））平衡近似误差和模态振幅；

(2) 在给定稀疏结构的条件下，通过求解带等式约束的二次规划问题，确定所选模态的振幅。

请注意，DMDSP 中使用的是投影动态模态。关于 DMDSP 的详细信息，参见Jovanovi´等人（2014）。

3.3. 实现

所提出的模型需要确定三个超参数：反循环矩阵中的延迟嵌入长度 τ、DMD 中约化矩阵的秩 r，以及 DMDSP 中的权衡参数 γ。回顾一下，数据 $\boldsymbol{X}$ 的延迟嵌入矩阵 $\boldsymbol{C}$ 是一个 $N\tau\times T$ 矩阵，其中空间维度 $N\tau$ 应远大于时间维度 T，以满足 Koopman 条件。我们将延迟嵌入长度设为数据的周期长度，例如交通数据的每日时间戳，以捕捉周期性。特别地，当 $\tau=T$ 时，矩阵 $\boldsymbol{C}$ 将成为完全分块反循环矩阵。较大的 τ 可以提升 circDMDsp 的性能，但由于需要使用 $\boldsymbol{C}\boldsymbol{P}_T(\boldsymbol{C}\boldsymbol{P}_T)^\top\in\mathbb{R}^{N\tau\times N\tau}$ 来计算奇异值分解中的左奇异向量 $\boldsymbol{U}$，计算效率会降低。为克服维数灾难，我们采用快照法，先计算一个小得多的矩阵 $(\boldsymbol{C}\boldsymbol{P}_T)^\top\boldsymbol{C}\boldsymbol{P}_T\in\mathbb{R}^{T\times T}$，以获得右奇异向量$\boldsymbol{V}$。较小矩阵的大小与延迟嵌入长度 τ 无关。获得 $\boldsymbol{V}$ 和 $\boldsymbol{\Sigma}$ 后，可用 $\boldsymbol{C}\boldsymbol{P}_T\boldsymbol{V}\boldsymbol{\Sigma}^{-1}$ 近似式（17）中的左奇异向量 $\boldsymbol{U}$。对于 DMD 中的矩阵秩 r，我们采用最优硬阈值（Gavish和Donoho，2014） δ 来确定其取值：

$$\delta=(0.56\beta^3-0.95\beta^2+1.82\beta+1.43)\sigma_{\text{med}}, \tag{25}$$

其中 $\beta=T/N\tau$和 σ_{med}是奇异值 diag(**Σ**) 的中位数，仅保留大于阈值的奇异值。稀疏参数 γ 控制最终的模态数量。在第4.2节中，我们设置了不同的 γ 值，以考察重建结果。

为量化估计性能，采用平均绝对误差（MAE）和均方根误差（RMSE）进行整体评估：

$$\begin{aligned}\text{MAE}&=\frac{1}{NT_\xi}\|\boldsymbol{X}-\boldsymbol{X}_{\text{DMD}}\|_1,\\ \text{RMSE}&=\sqrt{\frac{1}{NT_\xi}\|\boldsymbol{X}-\boldsymbol{X}_{\text{DMD}}\|_F^2}.\end{aligned} \tag{26}$$

对于第 4.4 节中的可预测性分析，我们使用平均绝对百分比误差（MAPE）来评估第 n 个传感器：

$$\text{MAPE}(n)=\frac{100\%}{T}\left\|\frac{\boldsymbol{X}(n,:)-\boldsymbol{X}_{\text{DMD}}(n,:)}{\boldsymbol{X}(n,:)}\right\|_1. \tag{27}$$

4. 案例研究

本研究采用所提出的circDMDsp框架分析真实世界的高速公路交通速度数据。本研究有三个目标：(1) 探究高速公路交通状态的时空特征；(2) 重建历史交通速度；(3) 基于识别出的动态交通模式预测未来值。

4.1. 高速公路交通速度数据集

我们使用从西雅图（美国华盛顿州）高速公路网络上的环形线圈检测器采集的交通速度数据（Cui 等人，2019）评估所提出的框架。传感器安装在四条高速公路上：I-5、I-90、I-405 和 SR-520。数据集包含南向（SB）的 157 个传感器和北向（NB）的 165 个传感器，时间分辨率为 5分钟（每天 288 个时间点）。为进行评估，我们使用前两周的数据进行训练，并利用随后一周的数据评估预测性能。

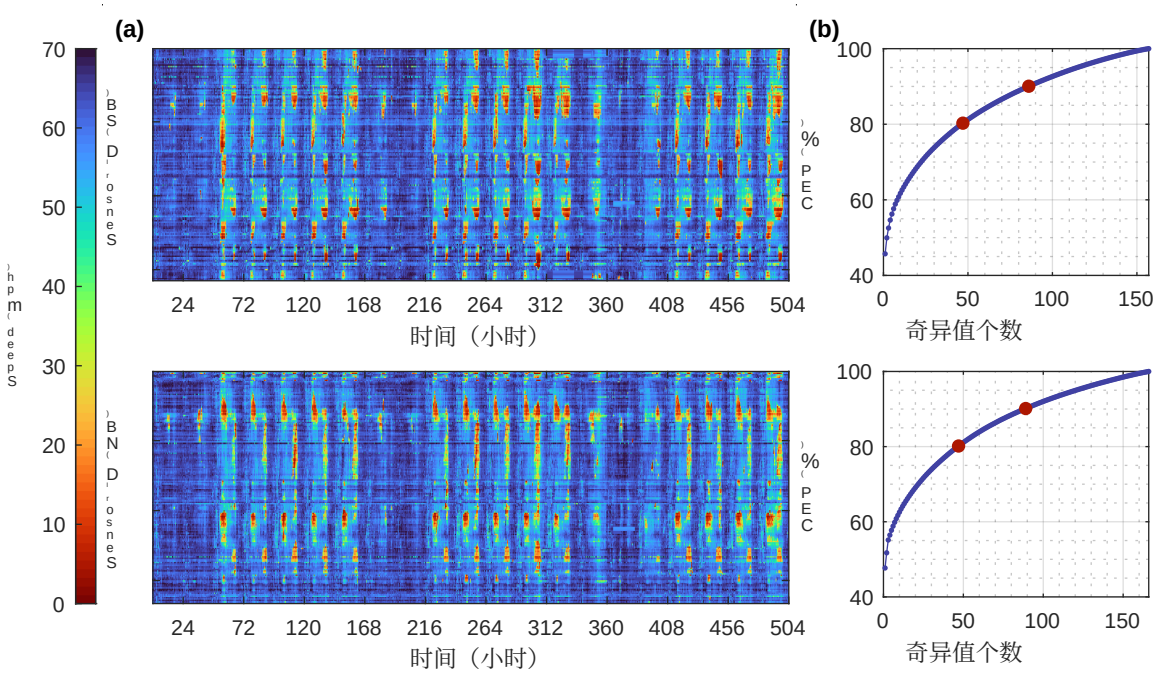

图 3：（a）：从美国西雅图采集的高速公路交通速度数据，时间分辨率为 5分钟；（b）：通过奇异值分解计算得到的训练数据累积特征值百分比（CEP）。红点分别表示累积特征值百分比达到 80% 和 90% 时对应的首个特征值。（上：南向；下：北向）。

我们在图 3 (a) 中展示训练数据（前 336 小时）和测试数据（第 336 至 504 小时）。可视化结果清楚地显示，交通数据两个方向都存在日/周周期性。为考察数据的低秩特性，我们在图 3 (b) 中绘制了累积特征值百分比（CEP）——$\sum_{i=1}^{r} a_i / \sum_{i=1}^{\min(N,T)} a_i$，$a_i$ 是 i-th largest singular values。如图所示，前 50 个奇异值占 CEP 的 80%，这表明交通数据可由少数潜在因子表示。低秩特性有助于降低模型复杂度，同时保持良好的准确性。然而，受交通事故、节假日/活动或传感器故障等多种因素影响，采集到的数据容易出现异常/噪声。例如，训练数据中南向（SB）前几个传感器连续缺失的数值（第 312 至 336 小时之间）是通过复制先前数值进行插补的。另一个例子是假日的影响，例如马丁·路德·金纪念日，这是美国的联邦假日，会改变正常的交通规律。这可以从测试数据第一天（星期六）和第三天（星期一）与前几周对应日期之间的差异中看出

以及前几周的对应日期。此类离群值的存在可能会显著影响动态模态分解模型的性能和应用（Schmid，2022）。

图 4 展示了从时间和空间两个角度观察训练数据中平均交通速度的变化情况。从时间角度来看，工作日两个方向的平均交通速度均呈现典型的“M”形模式（即早晚高峰时段），而周末的速度变化则较为平缓。从空间角度来看，同一位置两个方向的平均速度存在显著差异，呈现出较强的异质性。南向方向平均速度最低（拥堵区域）的位置位于 I-90 和 SR-520 交叉口附近，而北向方向的拥堵区域则位于西雅图西部大桥附近。结合图 3 (a)，我们发现，在高峰时段（6:00 至 10:00 以及 15:00 至 19:00），交叉口附近通常会出现明显的冲击波。这一现象可能是高交通拥堵期间车辆汇流和变道操作造成的。

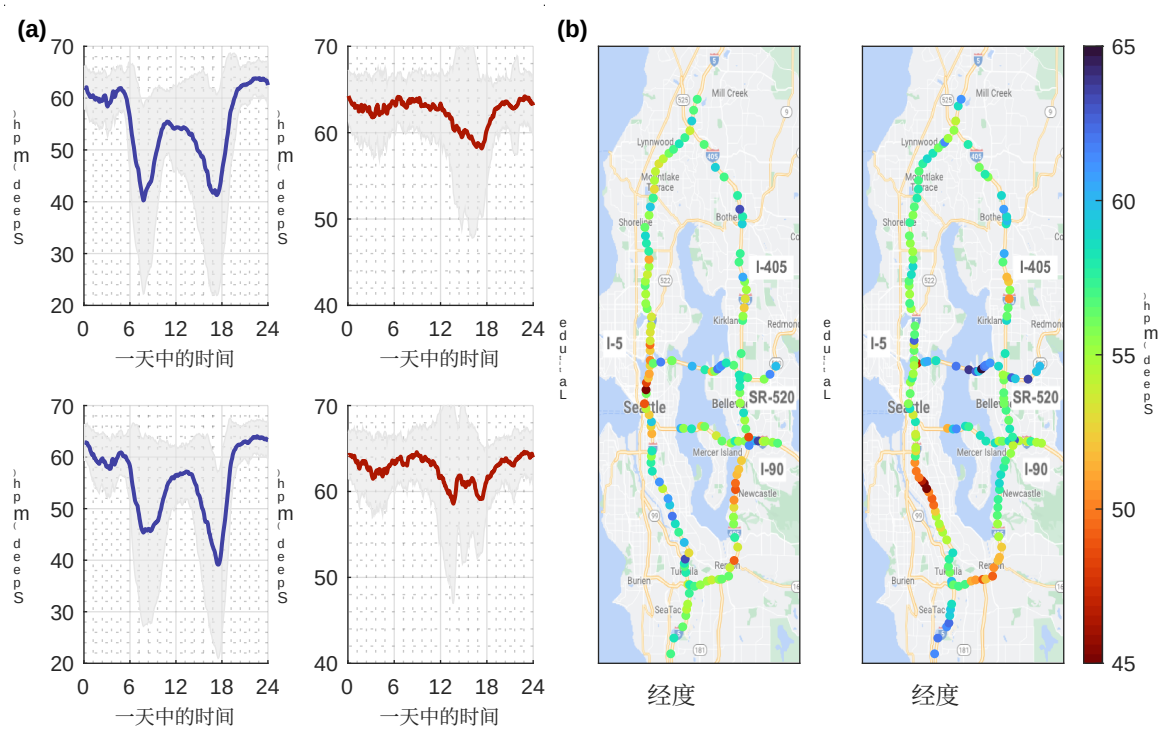

图 4：（a）：SB（上）和 NB（下）的平均交通速度及标准差。蓝线表示工作日，红线表示周末，灰色区域表示平均交通速度 ± 标准差；（b）：SB（左）和 NB（右）中每个传感器在训练期间的平均交通速度。传感器安装在四条高速公路上：I-5、I-90、I-405 和 SR-520（北向上）。

4.2. 交通速度历史重构与长期预测

我们首先在历史重构和长期预测中，将所提出的方法在不同稀疏性促进参数 λ 下与以下基线模型进行比较：

- DMD：标准动态模态分解模型（Kutz 等人，2016a）
- HDMD：汉克尔动态模态分解（Brunton 等人，2016a；Avila 和 Mezi´c，2020）。
- fbHDMD：正向与反向汉克尔动态模态分解（Dawson 等人，2016）
- tlsHDMD：全最小二乘汉克尔动态模态分解（Hemati 等人，2017）
- circDMD（本文提出）：反循环动态模态分解
- circDMDsp（本文提出）：采用稀疏促进策略的反循环动态模态分解

本质上，这些分析模型在三个方面存在差异：观测量（输入）、去噪能力和模态选择。默认的动态模态分解作用于数据矩阵 $\boldsymbol{X} \in \mathbb{R}^{N\times T}$，而 HDMD、fbHDMD 和 tlsHDMD 使用式（5）中定义的 Hankel矩阵 $\boldsymbol{H} \in \mathbb{R}^{(N\tau)\times(T-\tau+1)}$；本文提出的 circDMD 和 circDMDsp 则使用式（11）中定义的反循环矩阵 $\boldsymbol{C} \in \mathbb{R}^{N\tau\times T}$ 作为输入。使用 Hankel矩阵和反循环矩阵可显著增加观测量的数量，从而能够近似 Koopman算子。值得注意的是，当延迟嵌入长度设为 $\tau = 1$时，基于 Hankel矩阵和反循环矩阵的动态模态分解模型会退化为标准动态模态分解模型。在去噪能力方面，前向-后向动态模态分解（fbDMD）（Dawson 等人，2016）和全最小二乘动态模态分解（tlsDMD）（Hemati 等人，2017）旨在提高分析对数据中离群值的鲁棒性。为评估去偏性能，我们将这些模型与 HDMD 结合，构建了 fbHDMD 和 tlsHDMD。最后，circDMDsp模型仅选择主导模态，从而提供简洁的表示。从这个意义上说，circ DMD 是 circDMDsp 的一个特例，其中稀疏性促进参数 $\gamma = 0$。因此，circDMDsp 可视为一种通用模型，能够平衡准确性、鲁棒性和有意义的模式选择。

在历史重构实验中，我们使用式（25中的最优硬阈值来确定约化奇异值分解的秩，并将动态矩阵投影到低维子空间中，适用于所有模型。对于基于 Hankel 矩阵和基于反循环矩阵的模型，延迟嵌入长度 τ均设为 3 天（$\tau = 3\times 288 = 864$）。对于 circDMDsp 模型，稀疏参数 γ设为 500 或 1000。利用空间模态 $\boldsymbol{\Phi}$、振幅 $\boldsymbol{b}$和范德蒙德矩阵 $\boldsymbol{\Psi}$计算重建结果，其中 $T_\xi = T$ 取自式（22），并使用矩阵逆算子（例如，$\mathcal{C}\tau^{-1}$（$\boldsymbol{C}_{\mathrm{DMD}}$））。在长期预测中，空间模态和振幅保持不变，但范德蒙德矩阵中的时间戳长度会扩展至预测时间窗口，即 $T_\xi = T + 288\times 7$，以进行提前一周预测。

方法	南行方向			北行方向		
	重构	Pred-F3	Pred-L4	重构	Pred-F3	Pred-L4
DMD	11.83/14.27	22.38/23.00	25.48/25.77	10.24/12.72	19.77/20.46	23.82/24.10
HDMD	2.44/3.71	7.32/10.56	4.27/5.36	2.41/3.66	6.55/9.59	5.35/6.83
fbHDMD	8.65/13.27	24.57/31.73	37.49/43.19	7.31/11.20	17.43/22.50	28.86/34.04
tlsHDMD	2.41/3.69	6.49/9.59	3.24/4.15	2.40/3.64	6.63/9.68	5.20/6.75
circDMD（$\gamma = 0$）	**2.14/3.20**	**5.15**/9.45	2.76/3.62	**2.12/3.17**	**4.77**/8.39	2.46/3.17
circDMDsp ($\gamma = 500$)	2.31/3.49	5.15/9.43	2.52/3.31	2.24/3.34	4.79/8.40	2.41/3.09
circDMDsp ($\gamma = 10^3$)	3.05/4.67	5.23/**9.15**	**2.51/3.28**	2.94/4.52	4.84/**8.16**	**2.26/2.94**

表1：重构与预测实验：基线模型和所提出的模型在 $\tau = 3\times 288$ 下的平均绝对误差（英里/小时）/均方根误差（英里/小时）。

表1 展示了交通速度重构和预测的平均绝对误差（MAE）与均方根误差（RMSE）。针对两种不同情景评估预测准确率：第一种为包含假日的前三天，记为 Pred-F3；第二种为其余四天，记为 Pred-L4。标准 DMD 在所有任务中的 MAE 和 RMSE 均最高，未能准确估计交通速度。这是因为 DMD 的观测量不足以捕捉整体交通动态。总体而言，基于汉克尔矩阵的动态模态分解和基于循环矩阵的动态模态分解均优于标准 DMD 模型，fbHDMD 除外。fbHDMD 假设原始系统的动态是可逆的（Dawson 等人，2016）；然而，˜而且 $\boldsymbol{A}_f$和 $\tilde{\boldsymbol{A}}_b^{-1}$ 在交通网络中通常被投影到不同的子空间。另一种用于去噪的 DMD——tlsDMD，基于总体最小二乘来处理噪声，在 SB 中的精度较 HDMD 略有提升，但在 NB 中的性能则有所下降。所提出的 circDMD 和 circDMDsp 在重构和预测任务中均表现最佳。作为 circDMD 的扩展，circD MDsp 采用稀疏促进策略来降低秩并优化模态振幅，从而得到更简约的表示。随着 γ 的增大，重构性能会下降，因为模型中使用的时空模式更少。然而，

随着 γ 增大，预测性能略有提升。这表明，少数主导模式即可捕捉交通速度的主要特征。根据这两项任务的表现，我们在后续章节中将 circDMDsp 中的 $\gamma = 500$ 设为用于分析的值。

估计性能取决于动态模态 $\boldsymbol{\Phi}$ （动态矩阵的特征向量）以及由 $\boldsymbol{\lambda}$ （动态矩阵的相应特征值）计算得到的时间动态。具体而言，特征值 $|\lambda_i|$ 的幅值决定了相应的动态模态是稳态（$|\lambda_i| = 1$）还是变化的（衰减 $|\lambda_i| < 1$、增长 $|\lambda_i| > 1$）。也就是说，衰减的动态模态最终会收敛至零，而增长的动态模态则会发散至无穷大。二者都会对长期预测性能产生负面影响。相反，稳态动态模态保持不变或仅有轻微变化，有助于预测未来值。图5将南向基于DMD的模型得到的特征值绘制在复平面上，并标出单位圆。稳态特征值定义为 $(1-10^{-3}) \leq |\lambda_i| \leq (1+10^{-3})$，以蓝色标示；超出此范围的变化特征值则以红色标示。每个子图标题中还显示了 $|\boldsymbol{\lambda}|$ 与$\mathbf{1}$ 之间绝对差值的总和。图5表明，DMD中的所有特征值均位于单位圆内，导致相应空间模式迅速衰减，估计性能较差。相比之下，其他方法的大多数特征值都位于单位圆附近，尤其是所提出的circDMDsp模型，其绝对值总和最小。

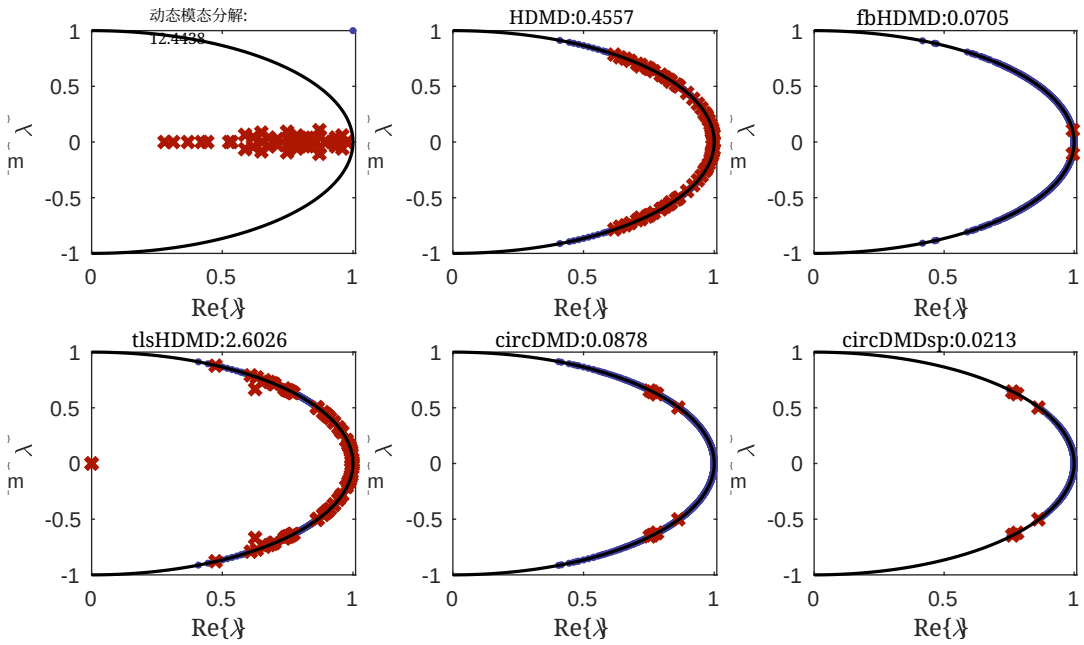

图5：南向复平面上的特征值 $\boldsymbol{\lambda}$ 及单位圆。蓝色圆点表示稳态动态模态，红色叉号表示变化动态模态。标题中显示了 $|\boldsymbol{\lambda}|$ 与$\mathbf{1}$ 之间绝对差值的总和。

4.3. 时空动态模式分析

基于DMD的模型的独特优势在于，它们能够通过对动态矩阵进行特征分解，得到具有可解释性的结果。本小节采用所提出的circDMDsp，从动态视角分析时空模式。

首先，特征值 $\angle\lambda_i$ 的相位决定了动态模态的振荡周期（Brunton 等，2016a）：

$$\text{period}_i = \frac{2\pi\Delta t}{\text{imag}(\log(\lambda_i))}, \tag{28}$$

其中，imag() 表示特征值的虚部，$\Delta t = 1/12$ 小时（5分钟）为本研究中的采样时间。振荡周期及其对应的振幅 $\boldsymbol{b}$ 如图所示

6。持续时间无限的振荡周期（imag(λ_i) = 0）和负周期（正周期对应的共轭特征值）已被排除。图中显示，振荡周期越长，通常振幅越大。换言之，交通速度主要受低频模式控制。无限周期的振幅最大，为 $Re(b) = -2.099 \times 10^4$，这表明非周期动态模态是数据中最主导的模式。第二主导模式的周期为24小时，这证实了交通速度具有日间相似性。模型还检测到了周模式（168小时）。除了日模式和周模式外，日内模式（周期为12小时、8小时、6小时、4.8小时）和周内模式（周期为84小时、28小时和11.2小时）也被识别为主导模式。所有这些周期均从数据中发现，未使用任何先验信息，反映了交通速度的内在演化。它们可用作交通特征，以支持下游任务，例如上一节提到的长期交通预测。

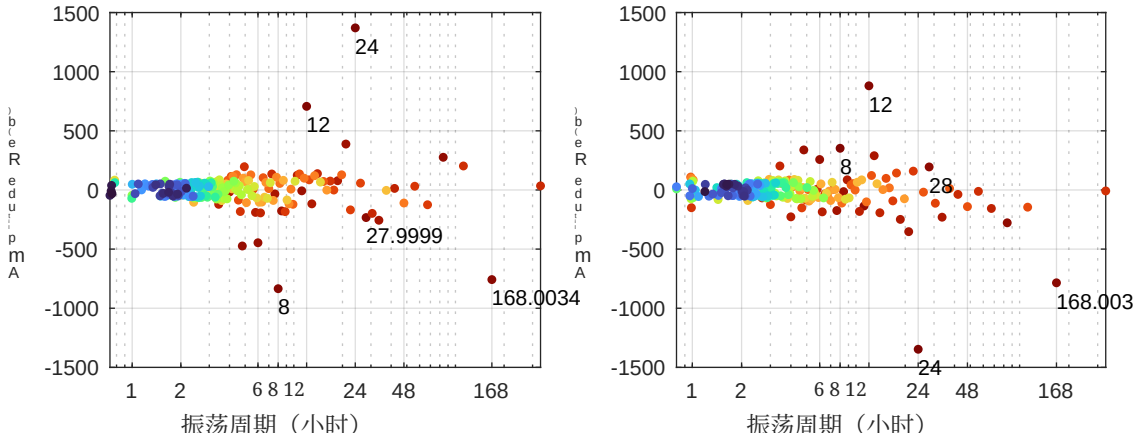

图6：南向（左）和北向（右）的振荡周期及振幅 $\boldsymbol{b}$。每个点表示一个振荡周期（小时）。图中未包含无限周期（$Re(b) = -2.099 \times 10^4$）和负周期。

空间模态 $\boldsymbol{\Phi}$ 反映了传感器之间的空间一致性。接下来，我们考察空间模态如何随其对应的振幅 $\boldsymbol{b}$ 和振荡周期演变。为此，我们定义一个重塑算子——reshape($\boldsymbol{\phi}_i b_i, N, \tau$)—将每个空间模态 $\boldsymbol{\phi}_i \in \mathbb{R}^{N\tau}(i = 1, \cdots, r)$及其振幅整理为一个大小为 $N \times \tau$的矩阵。本研究将延迟嵌入长度设为 $K = 288 \times 3$（3天），因此动态模态自然地表示空间模式在3天内的演变情况。图7展示了由 circDMDsp 识别出的441个模态中最显著的五个模态在前24小时内的演变情况，另外前十个模态见附录中的图14 和图15 。我们可以发现，交通速度数据由一系列具有不同振幅/权重的周期性模式构成。值得注意的是，circDMDsp 检测到三种占主导地位的时空结构：恒定/均值模式（模式 1）、日模式（模式 2）和周模式（模式 5）。与仅分离空间模式和时间模式的其他矩阵分解方法不同，所提出的模型基于局部线性假设，能够识别出具有一致性的时空模式 $\boldsymbol{x}_{t+1} = \boldsymbol{A}\boldsymbol{x}_t$。

为了更直观地展示，我们在图 8 中，以 2 小时为间隔，将西雅图高速公路上周期为 12 小时的模态（南向的模式 3、北向的模式 4）从 6:00 到 16:00 的情况绘制出来，并展示一天中不同时段获得的空间结构。我们可以检测出 12 小时周期内的拥堵点。拥堵点是指速度为负的附近传感器，即所有模态的聚合速度有所下降的区域。例如，在南向交通中，6:00 至 8:00 以及 16:00 时发现了一些拥堵点。这些拥堵点位于交叉口附近，例如 I-5 与 SR-520 之间，或 I-90 与 I-405 之间。到了午后早些时候，情况有所缓解（速度转为正值）。北向交通中的拥堵点位置和时间有所不同，更可能位于 I-5 与 I-90 的交叉口，或 I-405 与 SR-520 的交叉口。此外，北向交通中，8:00 时 I-5 上有更多传感器被视为拥堵点。动态模态的异质分布表明，交通特征受到交通流方向的

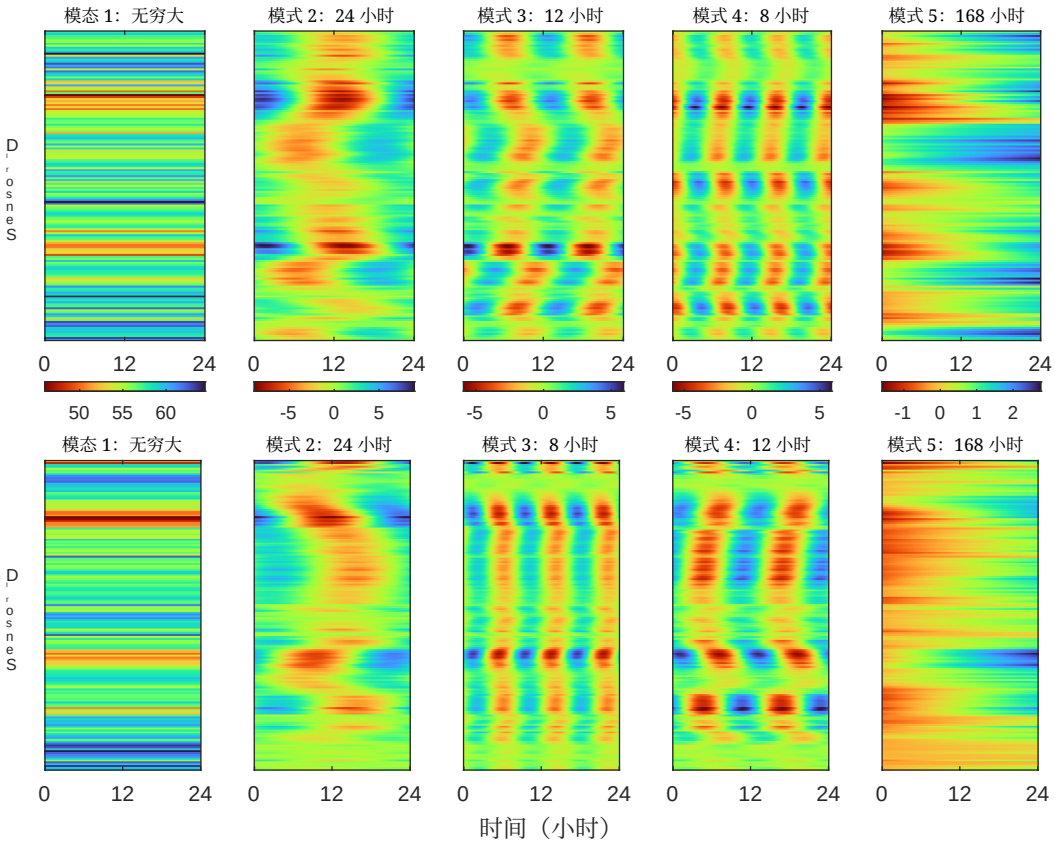

图7：前五个重塑后的动态模态的实部 $\phi_i b_i \in \mathbb{R}^{N\times\tau}$，$i = 1$，.....，5，其对应周期显示在标题中（上排：南向，下排：北向）。

影响。

也可以针对特定高速公路路段分析动态模态。例如，图 9 展示了 I-5高速公路南向交通中 6 个选定动态模态在 24 小时内的演化。I-5 与 SR-520 交叉口处安装的传感器以黑线标出。模型揭示了隐藏的时空波结构。结果表明，在日内模式中，一个表面上看似局部化的簇实际上位于交叉口前方（例如，模式 4、模式 8、模式 9 和模式 15）。这四种模态在高速公路的其他位置各不相同。例如，模式 9 在高速公路末端出现峰值，而模式 8 和模式 15 则出现在高速公路起点。这可能受到 I-5 上其他交叉口的影响，但本研究未纳入相关信息。模式 12 和模式14 属于周内模式，其局部化簇出现在交叉口之后。

4.4. 交通速度的短期可预测性

第 4.3 节表明，交通速度可以分解为一组周期性时空模式。长期交通模式通常与季节性趋势和规律有关，而短期交通模式则与非季节性特征相关（Boukerche 和 Wang，2020）。所提出的 circDMDsp 擅长长期预测（例如每日或每周预测），因为时空模式是周期函数（图 7），能够捕捉全局一致性。相反，在某些短期交通因素对预测至关重要的情境下，例如交通操作造成的不确定波动，circDMDsp 的预测性能可能不尽如人意。接下来，我们使用短期可预测性来评估传感器数据在多大程度上可以由动态模态分解近似。我们使用 circDMDsp 得出的平均绝对百分比误差来评估传感器层级的短期可预测性。

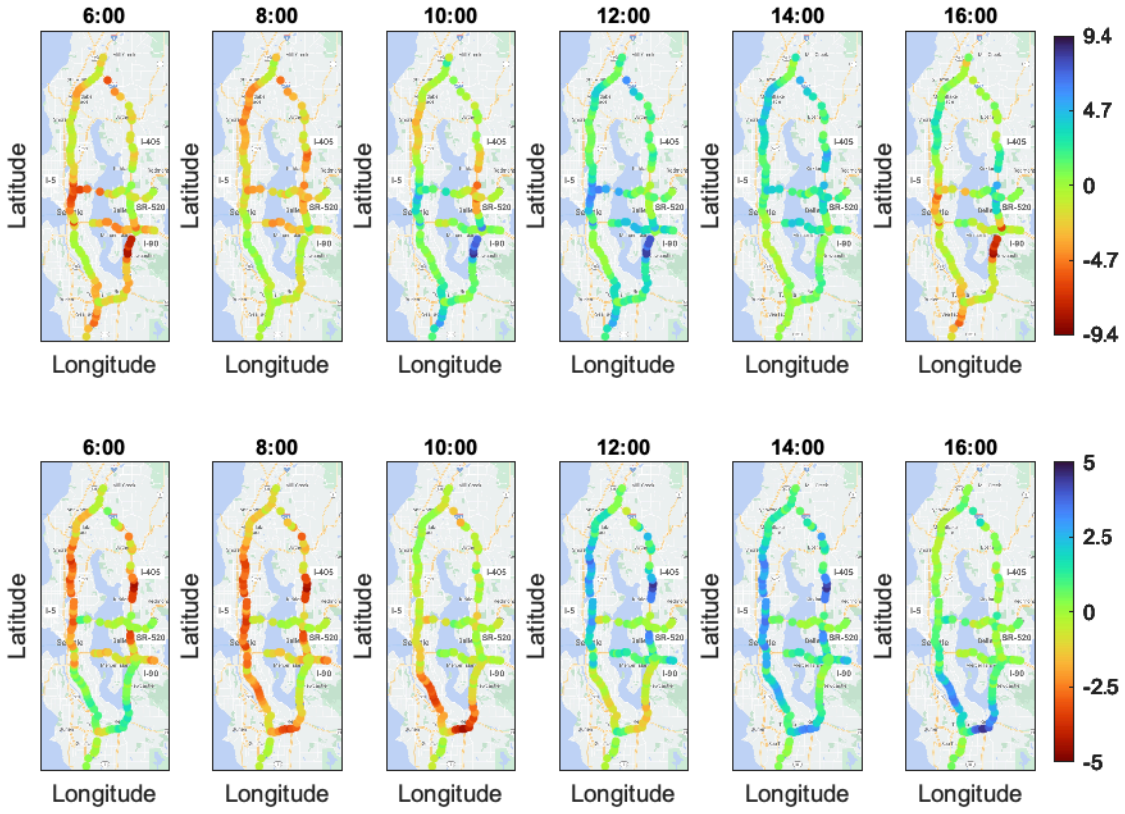

图8：6:00至16:00的12小时周期模态。每个子图的标题为时间戳。每个点代表一个传感器的位置，颜色表示该传感器的幅值（上排：南向，下排：北向）。

图 10 （a）展示了第 4.2 节中西雅图高速公路（I-5、I-90、I-405 和 SR-520）历史重构的平均绝对百分比误差分布。图 10 （b）展示了平均绝对百分比误差直方图，结果表明，大多数传感器的平均绝对百分比误差约为 5%，而两个方向上只有少数传感器的误差超过 10%。我们根据平均绝对百分比误差将所有传感器分为三组。平均绝对百分比误差较小的传感器更容易由模型解释；这也表明，这些传感器基本上由周期性重复模式主导，所提出的模型足以实现较高的预测准确率。另一方面，平均绝对百分比误差较高的传感器位于驶向西雅图的交通路段，或具有显著非重复性模式的高速公路交叉口。这些地点通常交通流量较大，交通操作也更复杂，因此需要更多局部信息来进行预测。对于这些具有显著非重复性模式的传感器，要实现更好的预测性能，其短期预测需要采用更好的模型，例如基于深度学习的时间序列模型。

接下来，我们从每个类别中选取三个传感器，并评估其可预测性。图中显示了 circDMDsp 对这三个选定传感器进行的14天历史重构和7天预测（绿色背景部分）：图11。传感器A在这三个传感器中平均绝对百分比误差最小。可以发现，传感器A采集的交通速度变化规律明显——速度数据几乎保持恒定，且时间变化平滑，没有任何突变。由于发现的时空模式能够很好地捕捉交通特征，因此传感器A的预测结果准确。对于传感器B，整体平均绝对百分比误差超过5%；不过，我们可以看到，最后四天的预测效果很好，而前三天表现不理想，主要是因为预测时段的第三天（星期一）是一个假日（马丁·路德·金纪念日），交通规律与训练时段中前一个星期一的不同。例如，早上可能会出现交通拥堵

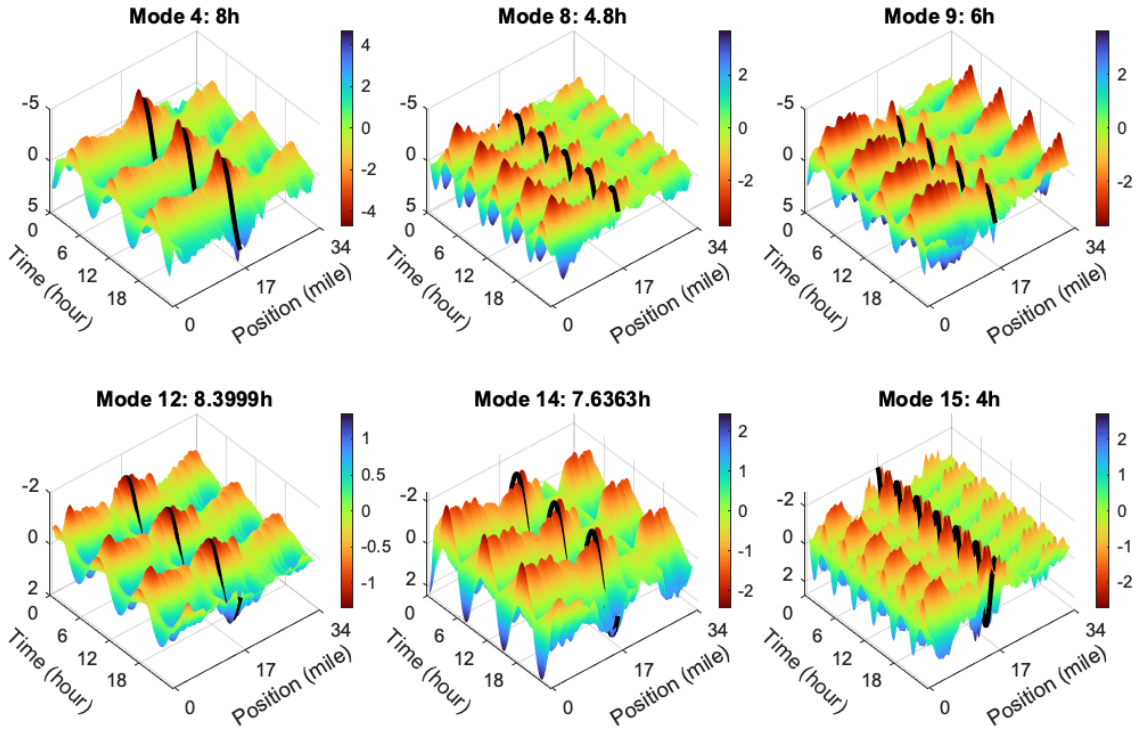

图9：I-5高速公路南向沿线6个动态模态及其对应振幅在24小时内的演化。I-5与SR-520的交叉口以黑线标示。z轴方向反转。

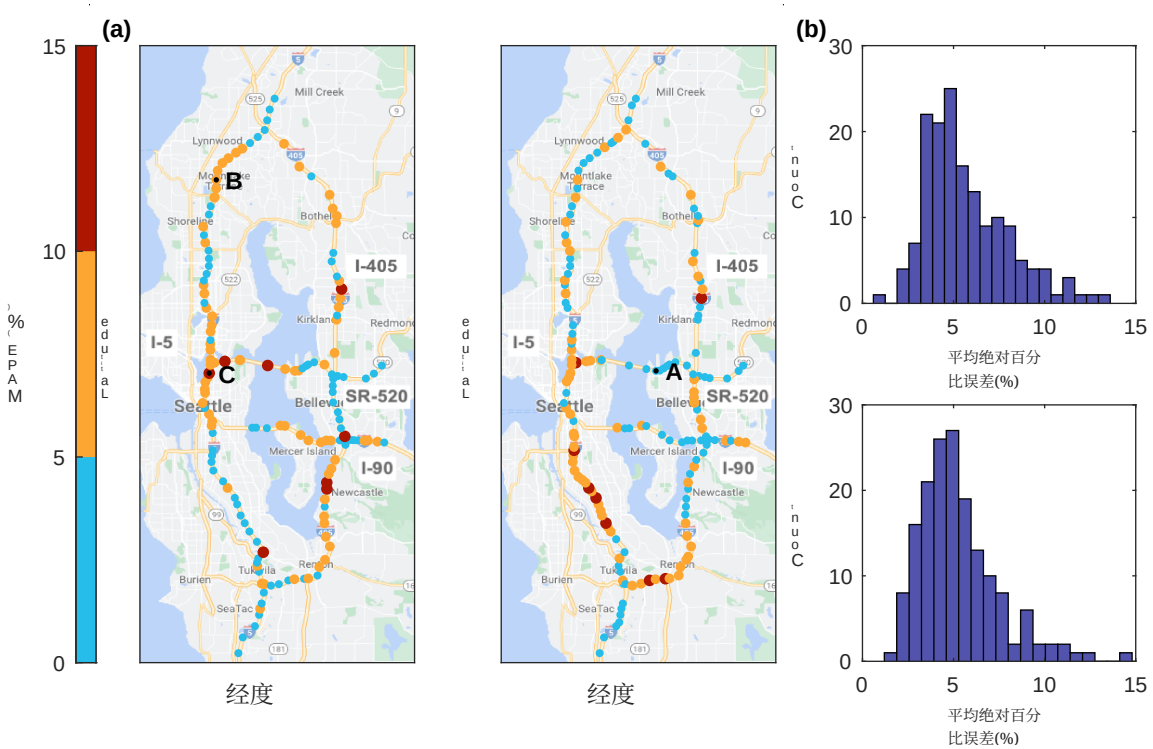

图10：（a）：南向（左）和北向（右）短期可预测性的平均绝对百分比误差分组；（b）：南向（上）和北向（下）的平均绝对百分比误差直方图。地图上标出了传感器A、传感器B和SensorC的位置。（北向上）

星期一达到峰值，但由于假日，在预测阶段不会出现这一现象。SensorC 的平均绝对百分比误差超过 10 %。预测 SensorC 的难点在于，该传感器位于 I-5 和 SR-520 的交汇处，交通流量在自由流与严重交通拥堵之间存在许多非重复的剧烈变化。这类变化主要受短期信息影响，仅依据周期性很难准确匹配峰值时段。对于这类传感器，我们建议采用更先进的短期预测框架，例如基于深度学习的时间序列模型。

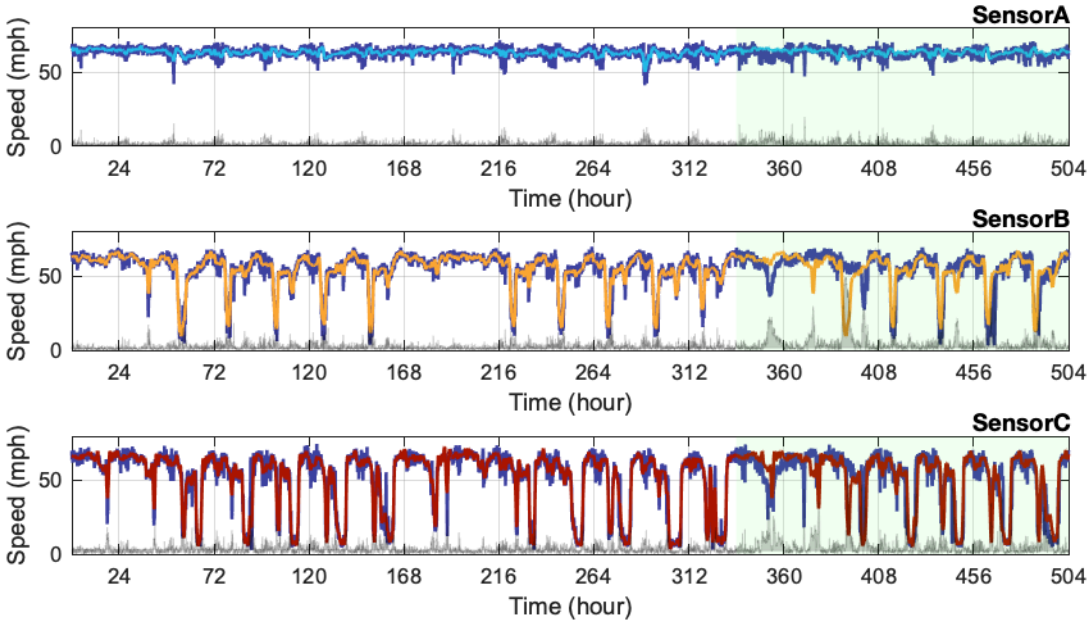

图11：circDMDsp模型对三个选定传感器进行14天估计和7天预测（绿色背景）的结果。紫色线：各传感器的真实值；蓝色、橙色和红色线：各传感器的估计值；灰色线：各传感器的绝对残差。传感器A、传感器B和SensorC的位置标注在图10 （a）中。

为进一步阐明传感器的短期可预测性，我们分析了历史重构的残差过程（图 11中的灰线）的自相关函数（ACF）。ACF揭示了不同时间滞后下的自相似性。每个时间滞后的系数表示相关性。图 12分别展示了传感器A、B和C的ACF，以及置信界限中的三倍标准误差（蓝线）。传感器A表现出非常弱的自相关性，这表明circDMDsp本身已经实现了非常好的重构，且从残差过程中可学习的信息并不多。对于传感器B和传感器C，当时间滞后分别约为30分钟和60分钟时，自相关性达到0。这一结果表明，传感器B/传感器C的重构/预测确实可以受益于更好的时间序列模型，该模型能够有效刻画残差过程（并适当选择滞后阶数）。

我们可以进一步分析残差过程的时空相关结构。为此，我们计算不同时间滞后下预测残差的相关矩阵，以分析网络层级的短期可预测性。设 $\boldsymbol{\eta}_t \in \mathbb{R}^N$ 为时间戳 t处的残差，则 l 时间滞后下的相关矩阵定义为 $\mathrm{corr}(\boldsymbol{\eta}_{t-l}, \boldsymbol{\eta}_t)$。为排除假日对分析的影响，我们分别对“Pred-F3”（3 天）和“Pred-L4”（4 天）进行相关性分析。图13 展示了 Pred-F3 和 Pred-L4 在时间滞后 $l = 1$、2、6、12（分别对应 5、10、30、60、120 分钟）下的残差相关矩阵；传感器按照其在高速公路上的位置顺序排列。图13 (a) 展示了前三天的相关结构，可以观察到所有传感器对之间都存在较强的正相关。然而，这只是由假日期间“系统性但不可预测”的交通流量造成的简单假象。我们更关注图13 (b) 所示的最后四天的分析结果。可以看出，当时间滞后较小时，确实存在明显的对角线结构，

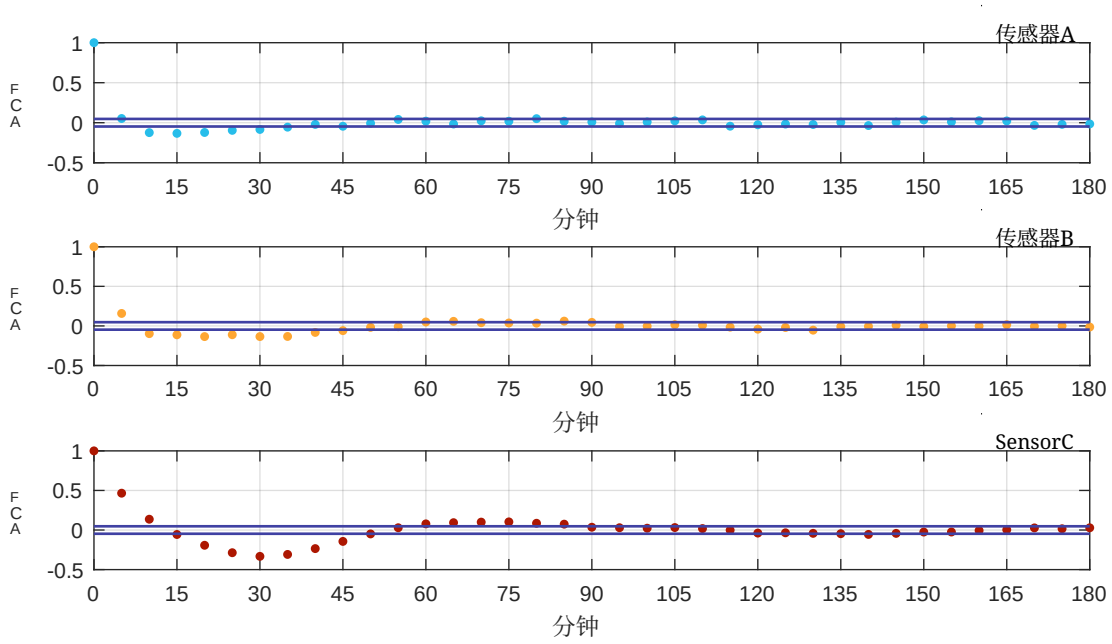

图12：传感器A、传感器B和SensorC的历史重构残差在5分钟时间滞后下的自相关函数（ACF）。蓝线表示置信界限中的3倍标准误。

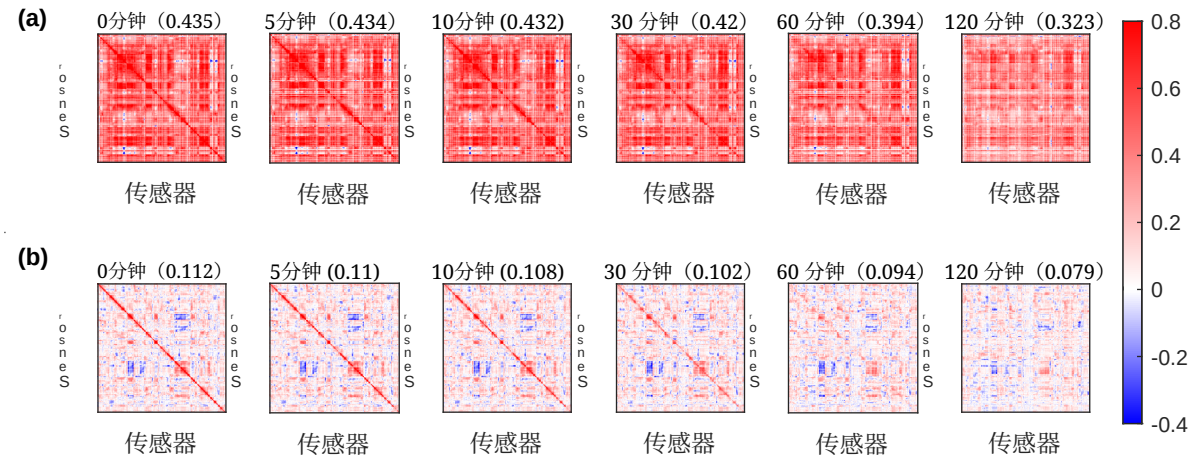

图13：Pred-F3（a）和Pred-L4（b）在不同时间滞后 l 下的残差相关矩阵。括号中显示的是相关系数绝对值的平均值。

这表明残差过程在空间上存在相关性。总体而言，随着时间滞后的增加，相关性逐渐减弱；当 $l = 12$（即 120 分钟）时，相关性几乎降为 0。这进一步证实，对于所分析的西雅图数据，只有当预测窗口小于 2 小时时，短期交通预测才有意义；否则，应依靠利用数据周期性、季节性和重复模式的长期预测框架。我们认为，这一发现适用于任何交通时间序列预测任务。

短期可预测性分析还有助于智能交通系统在实际运行中识别并排序传感器重要性。例如，短期可预测性较好的传感器/路段主要受所提出的 circDMDsp模型所学习到的周期性模式支配，因此对这些传感器/路段进行预测和插值较为简单。相反，短期可预测性不理想的

情况往往表现出非重复模式和由特定事件/事故造成的短期波动；在这些情况下，纳入局部/辅助信息对于实时预测和管理变得更加重要。

5. 结论与未来工作

本文提出了一种新颖的动态模态分解（DMD）框架——稀疏促进反循环动态模态分解（circ DMDsp），从动力系统的视角分析高速公路交通速度数据。所提出的circDMDsp框架将交通速度数据分解为动态模态、时间动态及相应的振幅，使我们能够对输入数据进行重构/去噪，并对未来数据进行长期预测/外推。总体而言，该模型的优势在于不需要大量训练数据，且用于训练模型的超参数很少。在假设circDMDsp能够充分学习并捕捉数据中的重复模式的基础上，我们还考察了高速公路交通速度数据的可预测性。我们在西雅图高速公路数据上的结果表明，当预测窗口小于一小时时，短期交通预测任务具有意义；我们认为，这对应于非重复事件通常的时间影响范围。超过这一范围后，便没有可供学习的短期信息，而鉴于交通数据具有强烈的周期性、重复性和节律性，利用简单的估计方法（如历史平均值和所提出的circDMD）即可实现高精度的长期预测。

未来研究有几个方向。首先，所提出框架得到的时空动态模式可用于许多下游应用/任务。例如，传感器的动态模态 $\mathbf{\Phi}$ 可用于研究动态相似性（Wang 和 Sun，2022），而时间演化 $\mathbf{\Psi}$ 可作为时间嵌入，为基于深度学习的预测模型提供支持。其次，第4.4节中的结果和分析可用于指导短期交通预测任务。例如，将circDMDsp 用作代理模型以去除长程季节性 / 周期性后，我们可以设计更好的短期预测模型，重点提升对非重复现象的预测能力。对于平均绝对百分比误差较小的传感器，circDMDsp 发现的规律模式能够很好地预测其数据；而对于其他传感器，则应纳入更多信息或采用其他模型（例如深度学习模型），以提升预测性能。第三，所提出的框架还可进一步扩展为在线版本，即随着新数据到达而更新模型（例如参见Cheng 等，2022），从而提高短期和长期预测的准确性。最后，一个很有前景的研究方向是开发能够直接处理含有缺失值的受损交通数据的 DMD 方法，利用观测值（例如参见Chen 和 Sun，2021；Lei 等，2022）及其动态特性（例如参见Chen 等，2022）的低秩特性。

致谢

本研究得到了加拿大自然科学与工程研究委员会和加拿大创新基金会的资助。X. Wang 感谢 FRQNT 提供 B2X 博士奖学金。

参考文献

Arbabi, H., Mezic, I., 2017. 遍历理论、动态模态分解与 Koopman 算子谱性质的计算。SIAM 应用动力系统杂志 16, 2096–2126。

Askham, T., Kutz, J.N., 2018. 优化动态模态分解的变量投影方法。SIAM 应用动力系统杂志 17, 380–416。

Avila, A., Mezi´c, I., 2020. 高速公路交通动态的数据驱动分析与预测。《自然·通讯》11, 1–16.

Boukerche, A., Wang, J., 2020. 面向智能交通系统的基于机器学习的交通预测模型。计算机网络 181, 107530. Brunton, B.W., Johnson, L.A., Ojemann, J.G., Kutz, J.N., 2016a. 使用动态模态分解提取大规模神经记录中的时空相干模式。神经科学方法杂志 258, 1–15. Brunton, S.L., Brunton, B.W., Proctor, J. L., Kutz, J.N., 2016b. 用于控制的非线性动力系统的Koopman不变子空间与有限线性表示。PLOS ONE 11, e0150171. Chen, X., Sun, L., 2021. 用于多维时间序列预测的贝叶斯时间因子分解。IEEE模式分析与机器智能汇刊 44, 4659–4673. Chen, X., Zhang, C., Chen, X., Saunier, N., Sun, L., 2022. 通过时变低秩自回归从时空数据中发现动态模式。arXiv预印本 arXiv:2211.15482 . Cheng, Z., Trepanier, M., Sun, L., 2022. 利用高阶加权动态模态分解实时预测地铁起讫点矩阵。交通科学 . Cui, Z., Henrickson, K., Ke, R., Wang, Y., 2019. 交通图卷积循环神经网络：一种用于网络级交通学习与预测的深度学习框架。IEEE智能交通系统汇刊 21, 4883–4894. Dawson, S.T., Hemati, M.S., Williams, M.O., Rowley, C.W., 2016. 表征并校正传感器噪声对动态模态分解的影响。流体实验 57, 42. Erichson, N.B., Brunton, S.L., Kutz, J.N., 2019. 用于背景建模的压缩动态模态分解。实时图像处理杂志 16, 1479–1492. Gavish, M., Donoho, D.L., 2014. 奇异值的最优硬阈值为 $4/\sqrt{3}$。IEEE信息论汇刊 60, 5040–5053. Hemati, M.S., Rowley, C.W., Deem, E.A., Cattafesta, L.N., 2017. 对动态模态分解去偏，以便对含噪数据集进行应用型Koopman谱分析。理论与计算流体力学 31, 349–368. Jovanovi´c, M.R., Schmid, P.J., Nichols, J.W., 2014. 稀疏促进动态模态分解。流体物理学 26, 024103. Koopman, B.O., Neumann, J.v., 1932. 连续谱动力系统。美国国家科学院院刊 18, 255–263. Korda, M., Mezi´c, I., 2018. 关于扩展动态模态分解收敛至Koopman算子。《非线性科学杂志》28, 687–710. Kutz, J.N., Brunton, S.L., Brunton, B.W., Proctor, J.L., 2016a. 动态模态分解：复杂系统的数据驱动建模。工业与应用数学学会。 Kutz, J.N., Fu, X., Brunton, S.L., 2016b. 多分辨率动态模态分解。SIAM 应用动力系统杂志 15, 713–735. Lehmberg, D., Dietrich, F., K¨oster, G., 2021. 墨尔本人建模——利用Koopman算子深入了解人群动力学。交通研究C辑：新兴技术 133, 103437. Lei, M., Labbe, A., Wu, Y., Sun, L., 2022. 用于时空交通数据插补与克里金法的贝叶斯核化矩阵分解。IEEE智能交通系统汇刊 23, 18962–18974.

Li, L., Su, X., Zhang, Y., Lin, Y., Li, Z., 2015. 交通时间序列分析的趋势建模：一项综合研究。《IEEE智能交通系统汇刊》16, 3430–3439. Liu, C., Huang, B., Zhao, M., Sarkar, S., Vaidya, U., Sharma, A., 2016. 利用高分辨率探测器数据探索交通网络系统动力学，载于：2016年IEEE第55届决策与控制会议（CDC），IEEE。第7629–7634页。Schmid, P.J., 2010. 数值与实验数据的动态模态分解。《流体力学杂志》656, 5–28. Schmid, P.J., 2022. 动态模态分解及其变体。《流体力学年度评论》54, 225–254. Tu, J.H., 2013. 动态模态分解：理论与应用。博士论文。普林斯顿大学。Wang, X., Miranda-Moreno, L., Sun, L., 2021. 用于多变量交通时间序列异常检测的汉克尔结构张量鲁棒主成分分析。arXiv预印本 arXiv:2110.04352。Wang, X., Sun, L., 2022. 通过汉克尔动态模态分解提取动态出行模式，载于：第11届交通分析三年研讨会，毛里求斯岛。Williams, M.O., Kevrekidis, I.G., Rowley, C.W., 2015. Koopman算子的数据驱动近似：拓展动态模态分解。《非线性科学杂志》25, 1307–1346. Yamamoto, R., Hontani, H., Imakura, A., Yokota, T., 2022. 延迟嵌入空间中低秩张量补全的快速算法，载于：IEEE/CVF计算机视觉与模式识别会议论文集，第2058–2066页。

附录

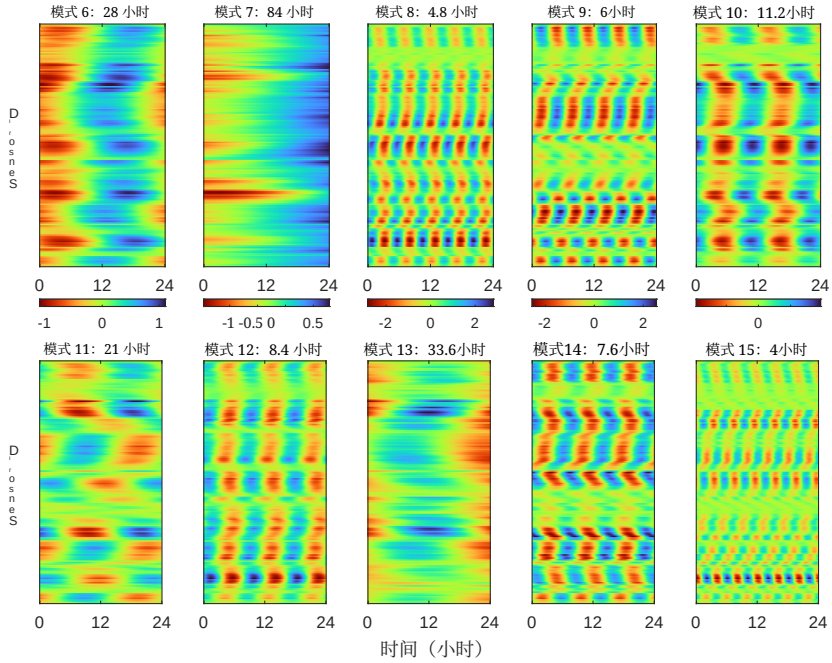

图14：南向交通流量中重构后的第6至第15个动态模态的实部 $\boldsymbol{\phi}_i b_i \in \mathbb{R}^{N\times\tau}$，$i=6$，……，15 及其对应周期。

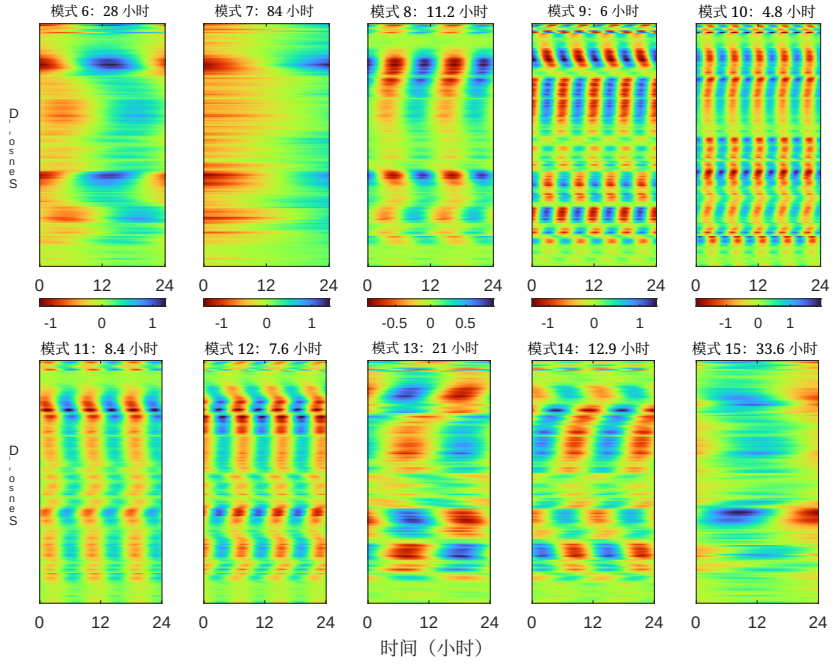

图15：北向交通流量中重构后的第6至第15个动态模态的实部 $\boldsymbol{\phi}_i b_i \in \mathbb{R}^{N\times\tau}$，$i=6$，……，15 及其对应周期。