



四川大學
SICHUAN UNIVERSITY

机器学习在时空数据建模中的应用

报告人：伍元凯

时间：2022.06.08

目录

CONTENT

- 
- 01 时空数据建模概述
 - 02 图神经网络空间插值方法
 - 03 高分辨率城市气候建模
 - 04 解绑表征城市通勤数据建模



时空数据概述

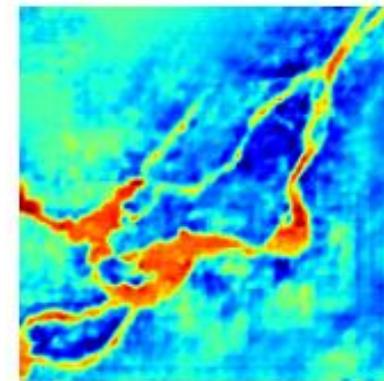
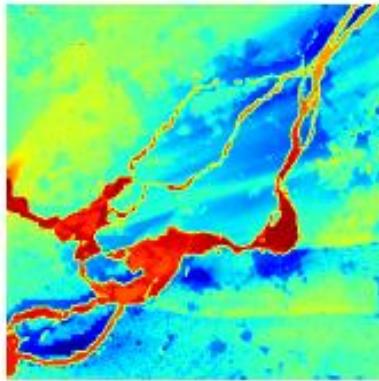


四川大学
SICHUAN UNIVERSITY

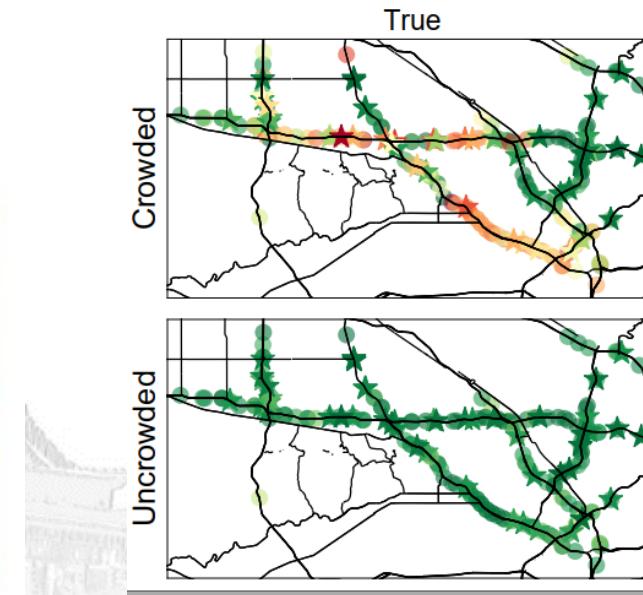
时空数据: 具有空间分布且随着时间变化而变化的数据的集合。如. 路段交通流量、人口密度、大气污染指数、网约车订单、国民经济发展、发生的事故。

来源: 气候监控、神经科学、社会科学、病毒传播学、**交通工程**、犯罪学、地球科学。

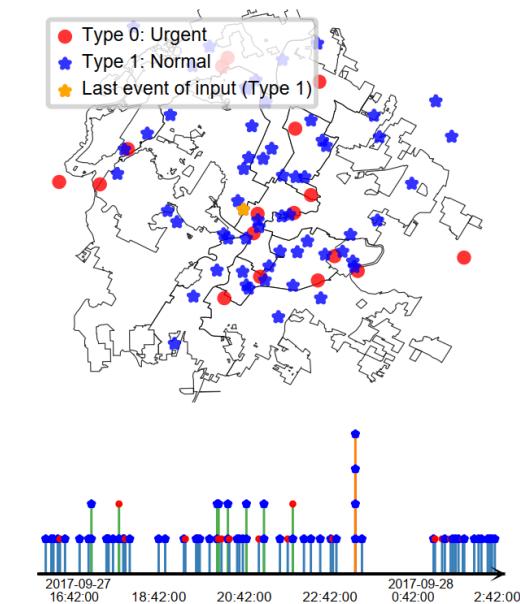
数据形态:



网格型：城市的温度分布



图型：路网传感器车速



事件型：城市的事故的时空分布

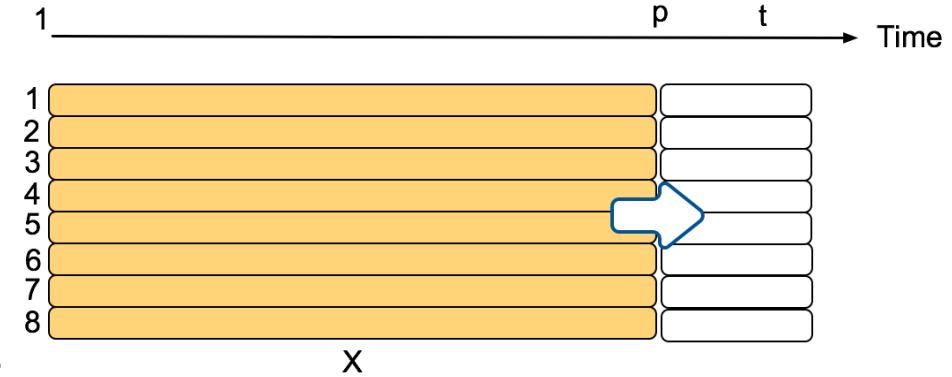
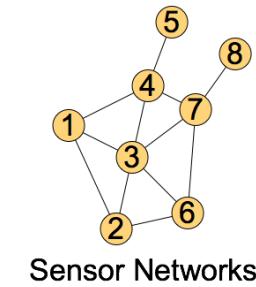


时空数据建模-典型应用

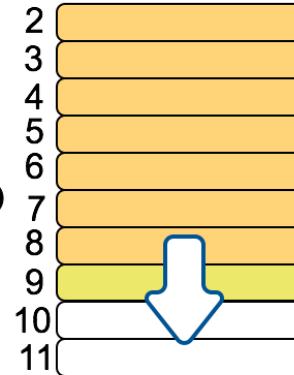
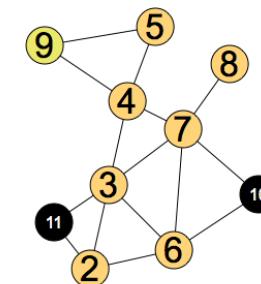


四川大学
SICHUAN UNIVERSITY

预测: 根据历史的信息预测未来的信息



插值 (克里金) : 根据有检测器位置的信号插值无检测器位置的信息

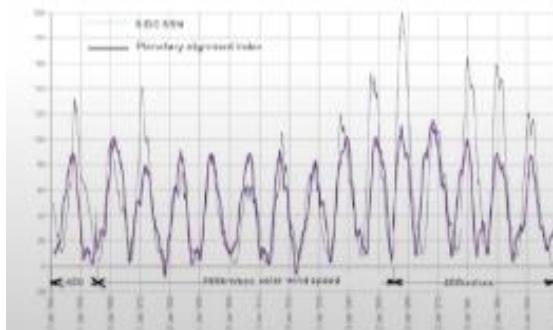




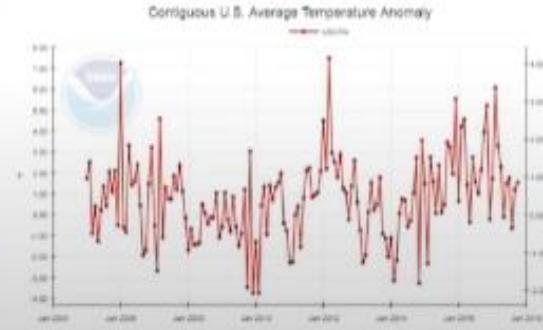
预测，为什么深度学习？



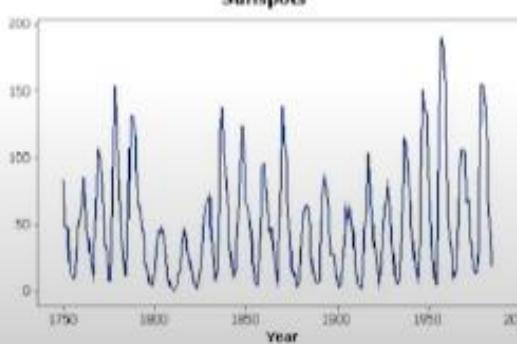
四川大学
SICHUAN UNIVERSITY



风速



温度



太阳盒子

统计特征：过少的信息

几何特征：过于复杂的启发式规则

分解：过于简单

ARIMA系列时间序列方法：只是不同时间之间的自相关性

深度学习：全能逼近能力，非线性，卷积神经网络可以学习到局部特征

[Deep Learning: the final Frontier for Time Series Analysis?](#)
[- YouTube](#)



传统预测方法-自回归模型



四川大学
SICHUAN UNIVERSITY

Rob J Hyndman
George Athanasopoulos

FORECASTING
PRINCIPLES AND PRACTICE

A comprehensive introduction to the latest forecasting methods using R. Learn to improve your forecast accuracy using dozens of real data examples.

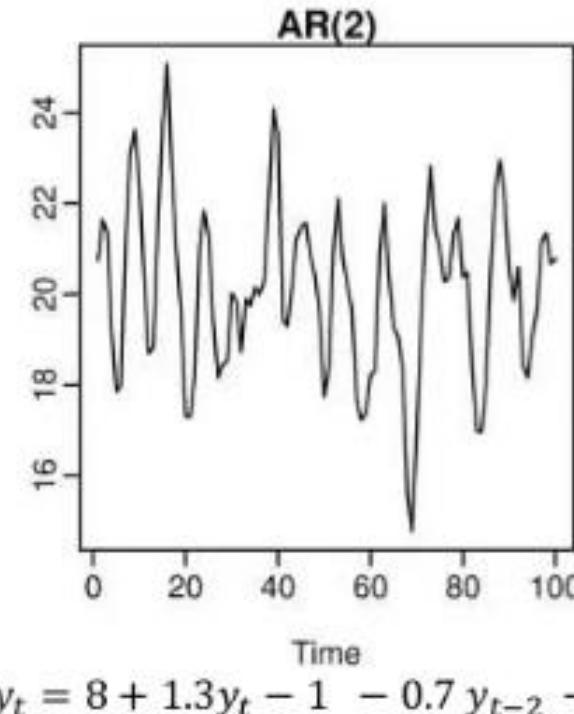
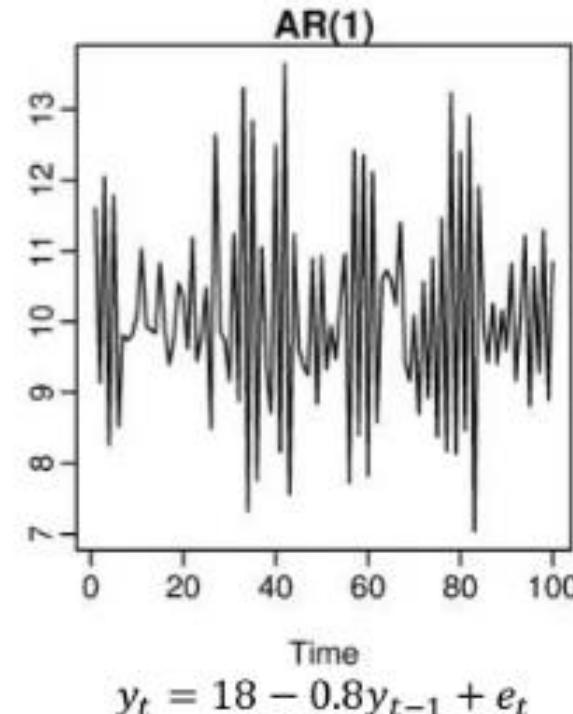


[Hyndsiht | Rob J Hyndman](#)

自回归模型可以拟合多种时间序列变化的模式

在书中RJ Hyndman将自回归模型定义为：在自回归模型中，我们将**未来变量的值表示为过去变量值的线性加权和**。自回归指的是变量使用自己的过去预测自己的未来。

$$\text{AR}(p): \quad y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + e_t$$





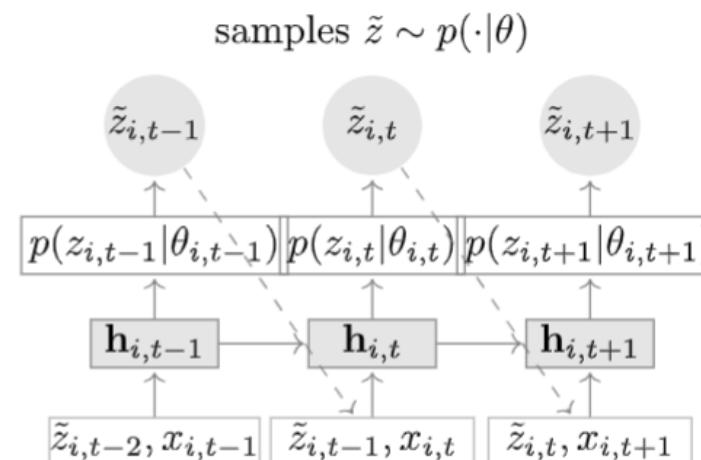
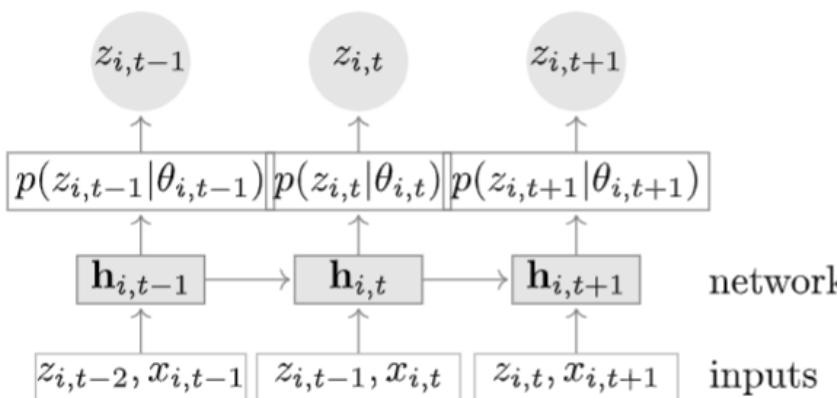
深度自回归模型



大多数自回归方法被设计用来预测单个或非常少量的时间序列模型，无法处理以下的新问题：

1. 大规模海量的时间序列预测问题：如淘宝的商品数量超过6亿种
2. 训练宏观模型的时候，需要处理不同时间序列的尺度问题。
3. AR模型没有考虑环境变量的输入问题
4. 预测新时间序列时有冷启动问题（Cold Start Problem）

深度自回归（DeepAR）： $P(\mathbf{z}_{i,t_0:T} | \mathbf{z}_{i,1:t_0-1}, \mathbf{x}_{i,1:T}) \xrightarrow{\text{环境变量}} Q_{\Theta}(\mathbf{z}_{i,t_0:T} | \mathbf{z}_{i,1:t_0-1}, \mathbf{x}_{i,1:T}) = \prod_{t=t_0}^T Q_{\Theta}(z_{i,t} | \mathbf{z}_{i,1:t-1}, \mathbf{x}_{i,1:T})$



$$= \prod_{t=t_0}^T p(z_{i,t} | \theta(\mathbf{h}_{i,t}, \Theta)),$$

$$p_G(z | \mu, \sigma) = (2\pi\sigma^2)^{-\frac{1}{2}} \exp(-(z - \mu)^2 / (2\sigma^2))$$

$$\mu(\mathbf{h}_{i,t}) = \mathbf{w}_\mu^T \mathbf{h}_{i,t} + b_\mu,$$

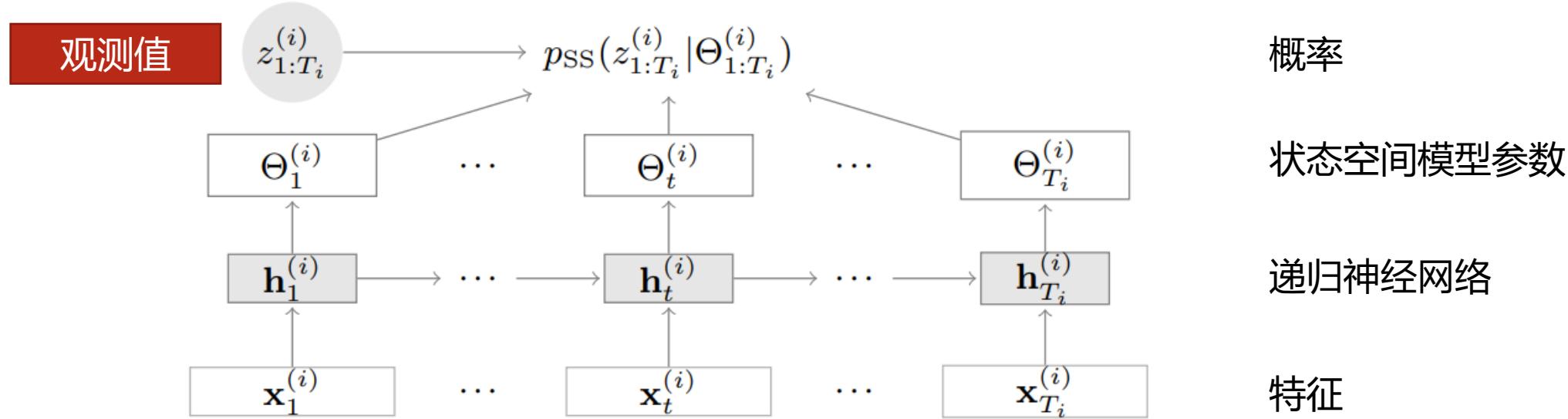
$$\sigma(\mathbf{h}_{i,t}) = \log(1 + \exp(\mathbf{w}_\sigma^T \mathbf{h}_{i,t} + b_\sigma)).$$



更复杂的结构假设



四川大学
SICHUAN UNIVERSITY



状态空间方程

$$\epsilon_t \sim N(0, 1), l_0 \sim N(\mu_0, \Sigma_0), l_t = F_t l_{t-1} + g_t \epsilon_t$$

$$y_t = a_t^T l_{t-1} + b_t, z_t = y_t + \sigma_t \epsilon_t$$

学习参数 $\Phi_t = (\mu_0, \Sigma_0, a_t, F_t, g_t, b_t, \sigma_t)$

训练损失

$$p(z_t^{(i)} | z_{1:t-1}^{(i)}, \Phi_{1:t-1}^{(i)}) = N(z_t^{(i)} | \mu_t^{(i)}, \Sigma_t^{(i)})$$

$$\mu_1^{(i)} = a_1^{(i)T} \mu_1^{(i)}, \quad \Sigma_1^{(i)} = a_1^{(i)T} \Sigma_0^{(i)} a_1^{(i)} + \sigma_1^{(i)2} \quad t = 1$$

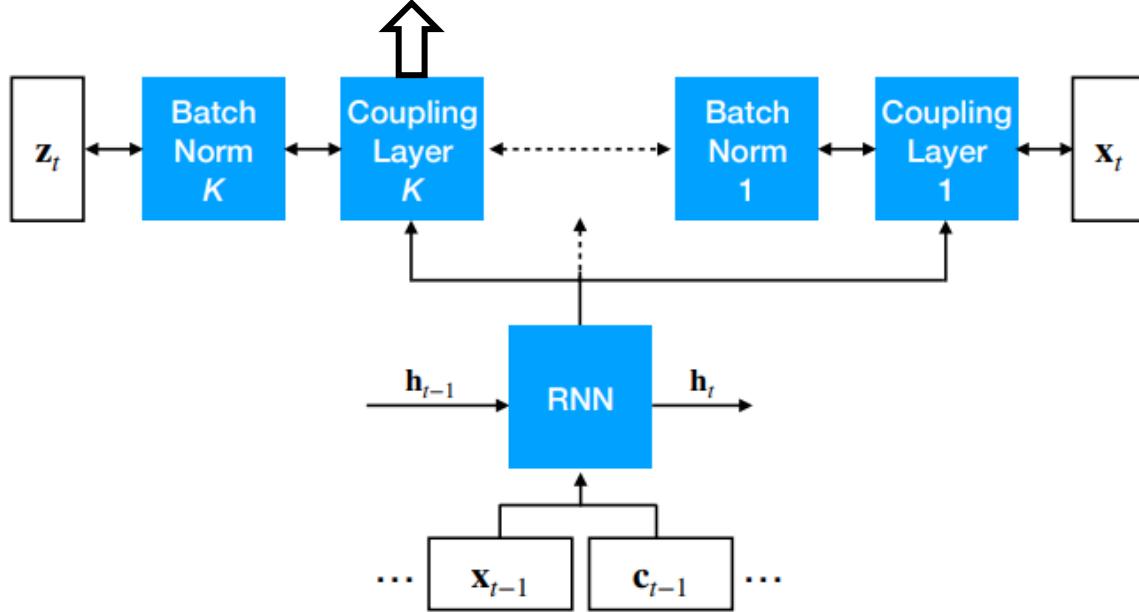
$$\mu_t^{(i)} = a_1^{(i)T} F_t^{(i)} f_{t-1}^{(i)}, \quad \Sigma_1^{(i)} = a_t^{(i)T} (F_t^{(i)} S_{t-1}^{(i)} F_t^{(i)T} + g_t^{(i)} g_t^{(i)T}) a_t^{(i)} + \sigma_t^{(i)2} \quad t > 1$$



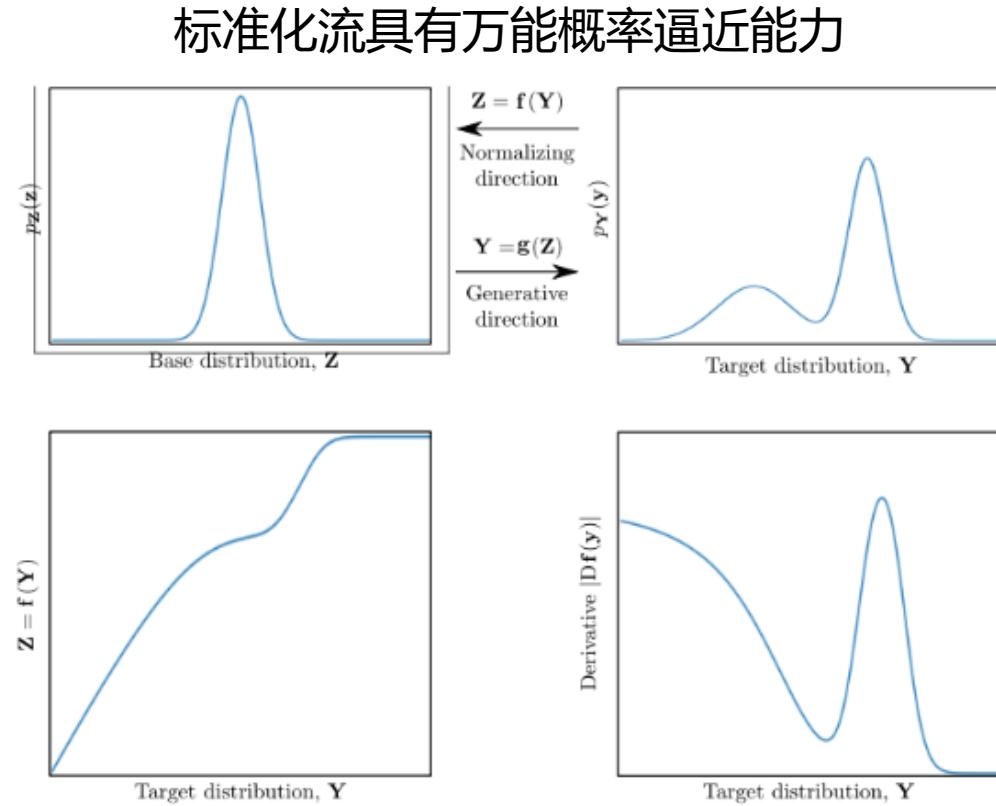
更复杂的结构假设



$$\begin{cases} \mathbf{y}^{1:d} = \mathbf{x}^{1:d} \\ \mathbf{y}^{d+1:D} = \mathbf{x}^{d+1:D} \odot \exp(s(\mathbf{x}^{1:d})) + t(\mathbf{x}^{1:d}), \end{cases}$$



标准化流



| Data set | Vec-LSTM ind-scaling | Vec-LSTM lowrank-Copula | GP scaling | GP Copula | LSTM Real-NVP | LSTM MAF |
|-------------|-------------------------|----------------------------|---------------|--------------|----------------------|----------------------|
| Exchange | 0.008±0.001 | 0.007±0.000 | 0.009±0.000 | 0.007±0.000 | 0.0064 ±0.003 | 0.005 ±0.003 |
| Solar | 0.391±0.017 | 0.319±0.011 | 0.368±0.012 | 0.337±0.024 | 0.331±0.02 | 0.315 ±0.023 |
| Electricity | 0.025±0.001 | 0.064±0.008 | 0.022±0.000 | 0.024±0.002 | 0.024±0.001 | 0.0208 ±0.000 |
| Traffic | 0.087±0.041 | 0.103±0.006 | 0.079±0.000 | 0.078±0.002 | 0.078±0.001 | 0.069 ±0.002 |
| Taxi | 0.506±0.005 | 0.326±0.007 | 0.183±0.395 | 0.208±0.183 | 0.175 ±0.001 | 0.161 ±0.002 |
| Wikipedia | 0.133±0.002 | 0.241±0.033 | 1.483±1.034 | 0.086±0.004 | 0.078±0.001 | 0.067 ±0.001 |

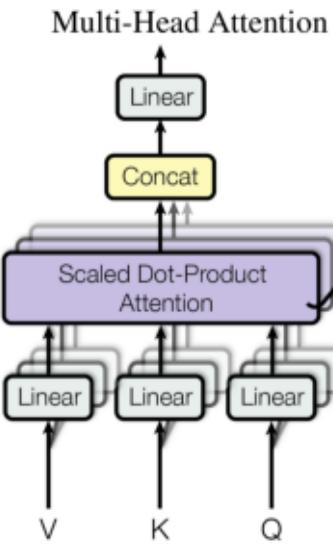
Rasul, Kashif, Abdul-Saboor Sheikh, Ingmar Schuster, Urs M. Bergmann, and Roland Vollgraf. "Multivariate Probabilistic Time Series Forecasting via Conditioned Normalizing Flows." In International Conference on Learning Representations. 2020.



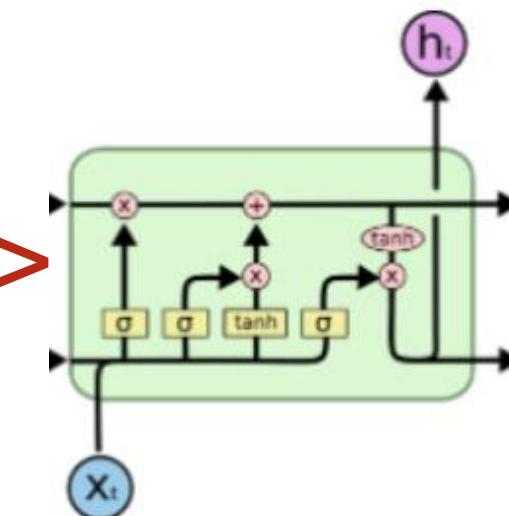
更强大的时间关联性建模能力



四川大学
SICHUAN UNIVERSITY

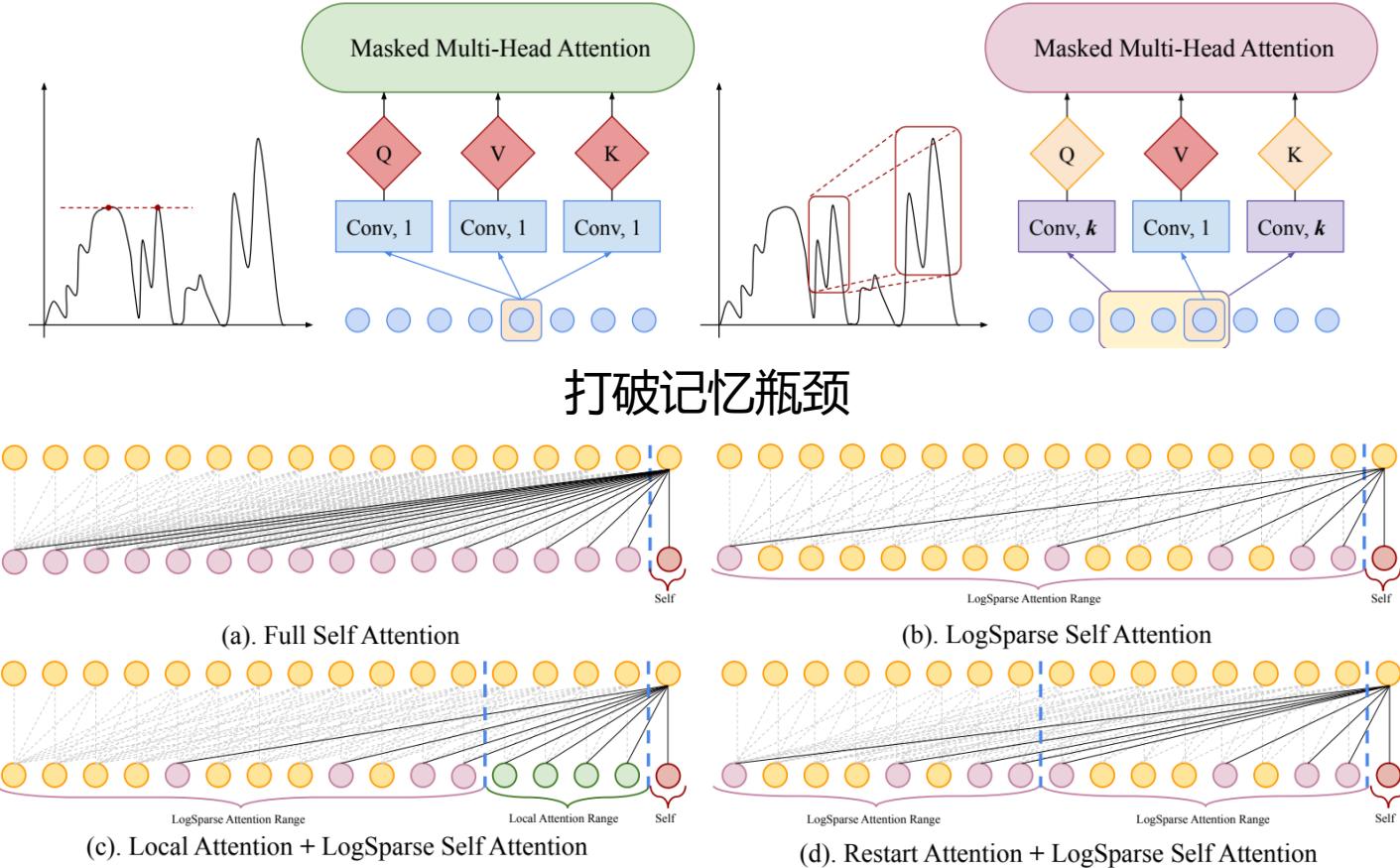


非序列化
自注意力
位置编码



序列化
过去信息-隐藏状态

增强Transformer局部规律的获取能力

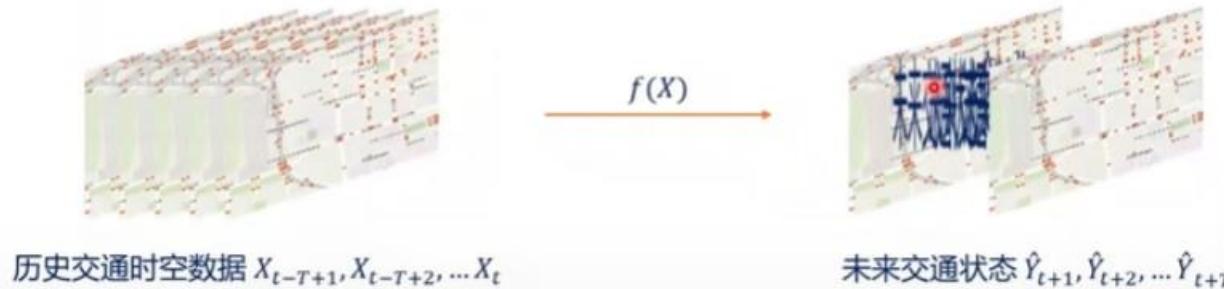




如何考慮多变量关系?



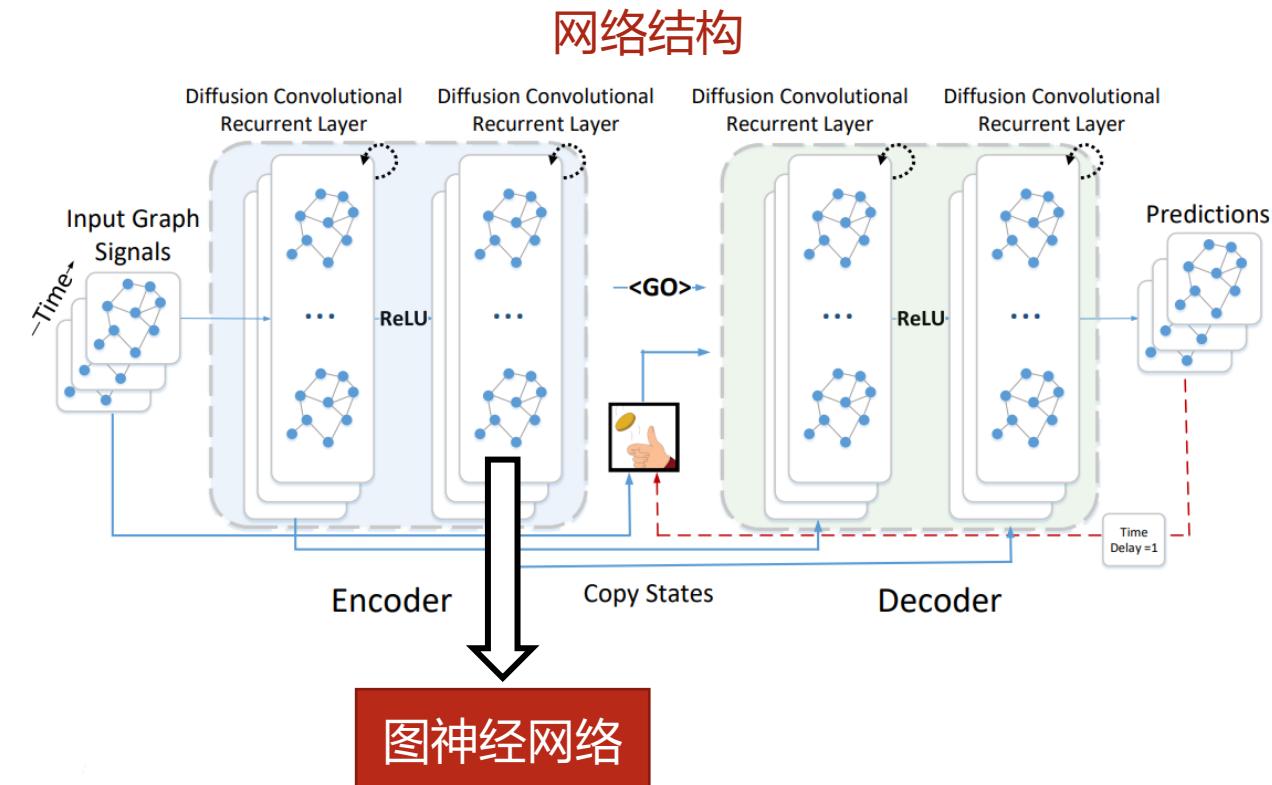
四川大学
SICHUAN UNIVERSITY



传统方法

| | T | Metric | HA | ARIMA _{Kal} | VAR | SVR | FNN | FC-LSTM | DCRNN |
|---------|--------|--------|-------|----------------------|-------|-------|-------|---------|--------------|
| METR-LA | 15 min | MAE | 4.16 | 3.99 | 4.42 | 3.99 | 3.99 | 3.44 | 2.77 |
| | | RMSE | 7.80 | 8.21 | 7.89 | 8.45 | 7.94 | 6.30 | 5.38 |
| | | MAPE | 13.0% | 9.6% | 10.2% | 9.3% | 9.9% | 9.6% | 7.3% |
| | 30 min | MAE | 4.16 | 5.15 | 5.41 | 5.05 | 4.23 | 3.77 | 3.15 |
| | | RMSE | 7.80 | 10.45 | 9.13 | 10.87 | 8.17 | 7.23 | 6.45 |
| | | MAPE | 13.0% | 12.7% | 12.7% | 12.1% | 12.9% | 10.9% | 8.8% |
| | 1 hour | MAE | 4.16 | 6.90 | 6.52 | 6.72 | 4.49 | 4.37 | 3.60 |
| | | RMSE | 7.80 | 13.23 | 10.11 | 13.76 | 8.69 | 8.69 | 7.59 |
| | | MAPE | 13.0% | 17.4% | 15.8% | 16.7% | 14.0% | 13.2% | 10.5% |

| | T | Metric | HA | ARIMA _{Kal} | VAR | SVR | FNN | FC-LSTM | DCRNN |
|----------|--------|--------|------|----------------------|------|------|-------|---------|-------------|
| PEMS-BAY | 15 min | MAE | 2.88 | 1.62 | 1.74 | 1.85 | 2.20 | 2.05 | 1.38 |
| | | RMSE | 5.59 | 3.30 | 3.16 | 3.59 | 4.42 | 4.19 | 2.95 |
| | | MAPE | 6.8% | 3.5% | 3.6% | 3.8% | 5.19% | 4.8% | 2.9% |
| | 30 min | MAE | 2.88 | 2.33 | 2.32 | 2.48 | 2.30 | 2.20 | 1.74 |
| | | RMSE | 5.59 | 4.76 | 4.25 | 5.18 | 4.63 | 4.55 | 3.97 |
| | | MAPE | 6.8% | 5.4% | 5.0% | 5.5% | 5.43% | 5.2% | 3.9% |
| | 1 hour | MAE | 2.88 | 3.38 | 2.93 | 3.28 | 2.46 | 2.37 | 2.07 |
| | | RMSE | 5.59 | 6.50 | 5.44 | 7.08 | 4.98 | 4.96 | 4.74 |
| | | MAPE | 6.8% | 8.3% | 6.5% | 8.0% | 5.89% | 5.7% | 4.9% |



$$\mathbf{X}_{:,p} \star_{\mathcal{G}} f_{\theta} = \sum_{k=0}^{K-1} \left(\theta_{k,1} (\mathbf{D}_O^{-1} \mathbf{W})^k + \theta_{k,2} (\mathbf{D}_I^{-1} \mathbf{W}^\top)^k \right) \mathbf{X}_{:,p} \quad \text{for } p \in \{1, \dots, P\}$$

Li, Y., Yu, R., Shahabi, C., & Liu, Y. (2018, February). Diffusion Convolutional Recurrent Neural Network: Data-Driven Traffic Forecasting. In *International Conference on Learning Representations*.



传统空间插值方法（克里金法）



四川大学
SICHUAN UNIVERSITY

克里金方法-高斯过程回归：



Georges Fran ois
Paul Marie Matheron

$$P(\mathbf{f} | \mathbf{X}) = \mathcal{N}(\mathbf{f} | \boldsymbol{\mu}, \mathbf{K})$$

$$\begin{array}{l} \text{观测点} \\ \text{未观测点} \end{array} \begin{bmatrix} \mathbf{f} \\ \mathbf{f}_* \end{bmatrix} \sim \mathcal{N} \left(\begin{bmatrix} \mathbf{m}(\mathbf{X}) \\ \mathbf{m}(\mathbf{X}_*) \end{bmatrix}, \begin{bmatrix} \mathbf{K} & \mathbf{K}_* \\ \mathbf{K}_*^\top & \mathbf{K}_{**} \end{bmatrix} \right)$$

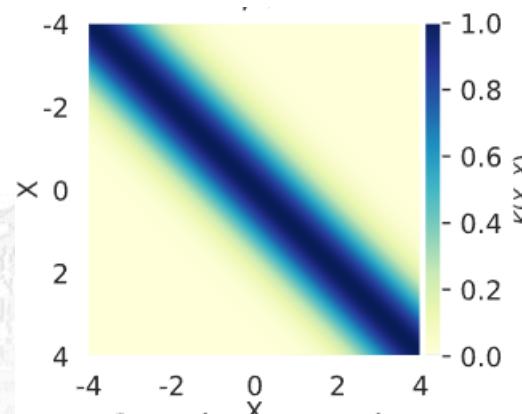
$$\mathbf{f}_* | \mathbf{f}, \mathbf{X}, \mathbf{X}_* \sim \mathcal{N} (\mathbf{K}_*^\top \mathbf{K} \mathbf{f}, \mathbf{K}_{**} - \mathbf{K}_*^\top \mathbf{K}^{-1} \mathbf{K}_*)$$

$$\text{带噪音的观测值 } \mathbf{y} = f(x) + \epsilon \quad \begin{pmatrix} \mathbf{y} \\ \mathbf{f}_* \end{pmatrix} \sim \mathcal{N} \left(\mathbf{0}, \begin{bmatrix} \mathbf{K} + \sigma_n^2 I & \mathbf{K}_* \\ \mathbf{K}_*^\top & \mathbf{K}_{**} \end{bmatrix} \right)$$

$$\begin{aligned} \bar{\mathbf{f}}_* &\triangleq \mathbb{E}[\bar{\mathbf{f}}_* | \mathbf{X}, \mathbf{y}, \mathbf{X}_*] \\ &= \mathbf{K}_*^\top [\mathbf{K} + \sigma_n^2 I]^{-1} \mathbf{y}, \end{aligned}$$

$$\text{cov}(\mathbf{f}_*) = \mathbf{K}_{**} - \mathbf{K}_*^\top [\mathbf{K} + \sigma_n^2 I]^{-1} \mathbf{K}_*$$

$$k(x_a, x_b) = \sigma^2 \exp \left(-\frac{\|x_a - x_b\|^2}{2\ell^2} \right)$$



在克里金方法中，核函数只和距离有关



高斯过程回归的不足

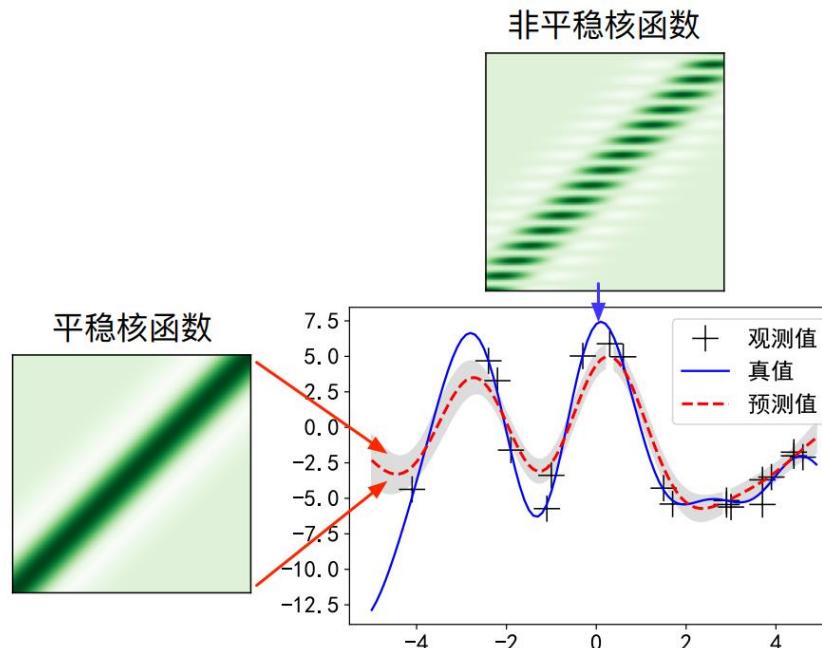


1. 在多变量时失去了有效性—当特征的个数超过几十后

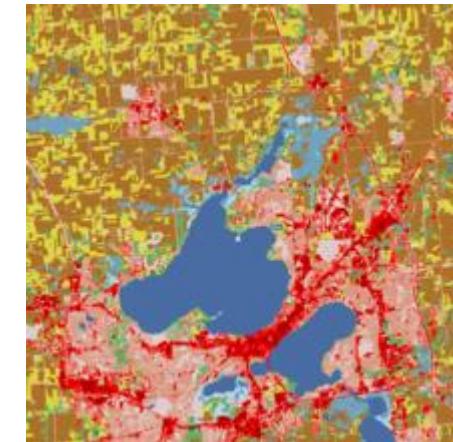
2. 需要完整的样本来进行插值，
当样本较多时会陷入维度灾难

$$\begin{aligned}\bar{\mathbf{f}}_* &\triangleq \mathbb{E}[\bar{\mathbf{f}}_* | \mathbf{X}, \mathbf{y}, \mathbf{X}_*] \\ &= \mathbf{K}_*^\top [\mathbf{K} + \sigma_n^2 \mathbf{I}]^{-1} \mathbf{y}, \\ \text{cov}(\mathbf{f}_*) &= \mathbf{K}_{**} - \mathbf{K}_*^\top [\mathbf{K} + \sigma_n^2 \mathbf{I}]^{-1} \mathbf{K}_*\end{aligned}$$

3. 依赖对核函数的选取



4. 高斯分布在形容空间异质性时会失效



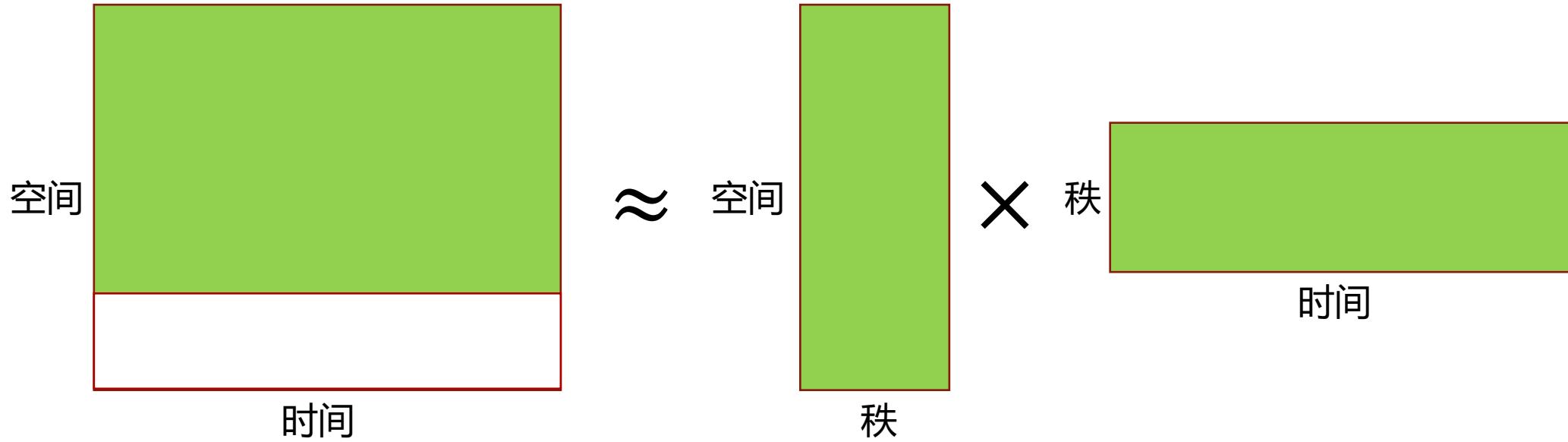
Jiang, B. (2015). Geospatial analysis requires a different way of thinking: The problem of spatial heterogeneity. *GeoJournal*, 80(1), 1-13.



矩阵/张量填充方法的不足



四川大学
SICHUAN UNIVERSITY



- 问题：
1. 直推式方法 (Transductive)，不能推广到训练过程中未出现的时空节点
 2. 某些位置获得新的传感器信息时，需要完全重新优化 (训练) 矩阵填充算法。

目录

CONTENT

01 时空数据概述
03 高分辨率城市气候建模

02 图神经网络空间插值方法
04 解绑表征城市通勤数据建模

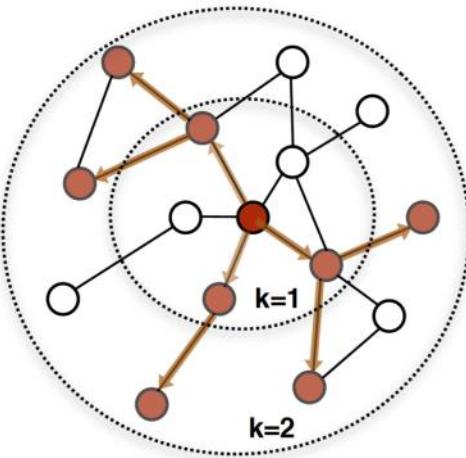




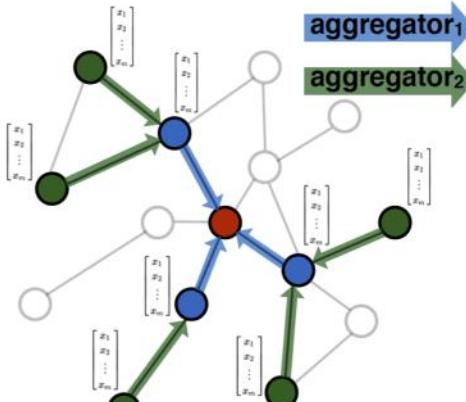
为什么选择图神经网络?



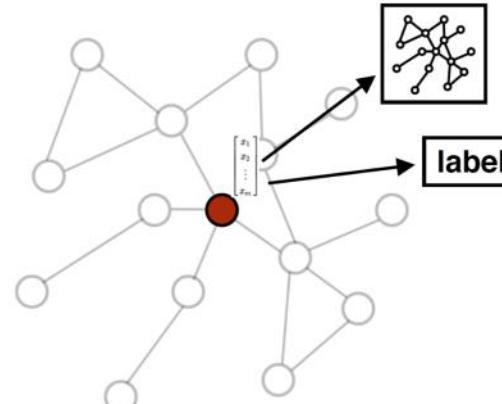
四川大学
SICHUAN UNIVERSITY



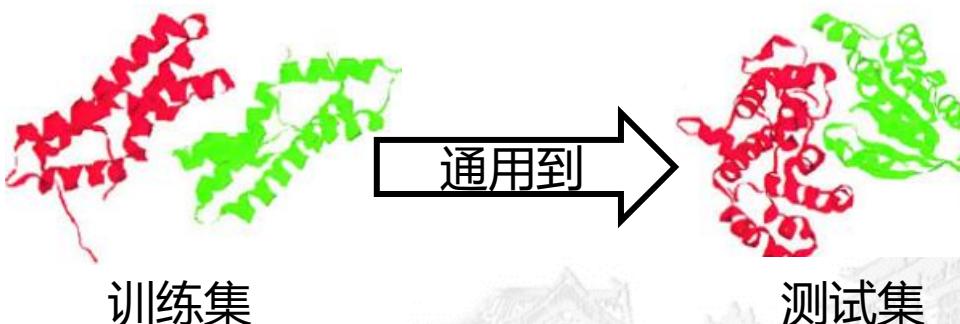
1.随机邻居采样



2.信息聚合



3.预测



| Name | PPI | |
|---------------------|--------------|--------------|
| | Unsup. F1 | Sup. F1 |
| Random | 0.396 | 0.396 |
| Raw features | 0.422 | 0.422 |
| DeepWalk | — | — |
| DeepWalk + features | — | — |
| GraphSAGE-GCN | 0.465 | 0.500 |
| GraphSAGE-mean | 0.486 | 0.598 |
| GraphSAGE-LSTM | 0.482 | 0.612 |
| GraphSAGE-pool | 0.502 | 0.600 |
| % gain over feat. | 19% | 45% |

采用随机采样的模型具备归纳(inductive)能力

Hamilton, W., Ying, Z., & Leskovec, J. (2017). Inductive representation learning on large graphs. Advances in neural information processing systems, 30.



基于归纳图神经网络的克里金方法

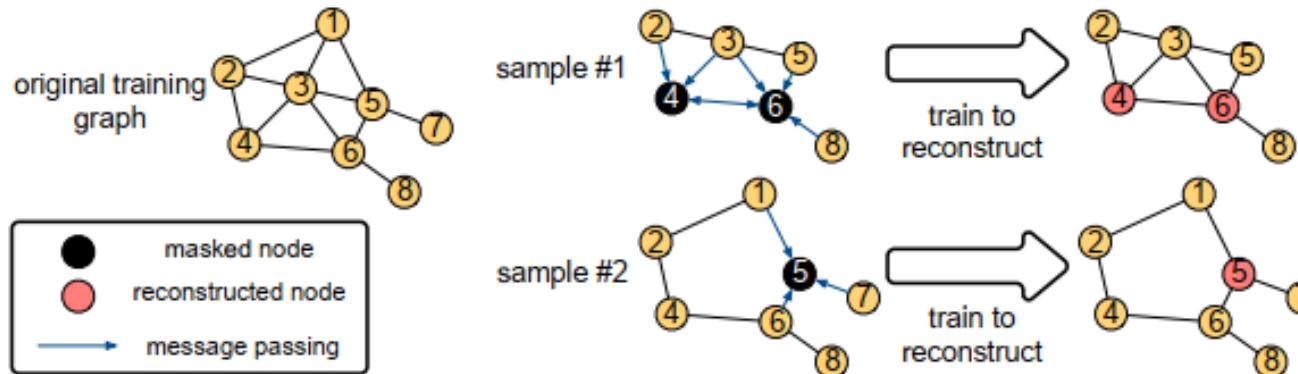


四川大学
SICHUAN UNIVERSITY

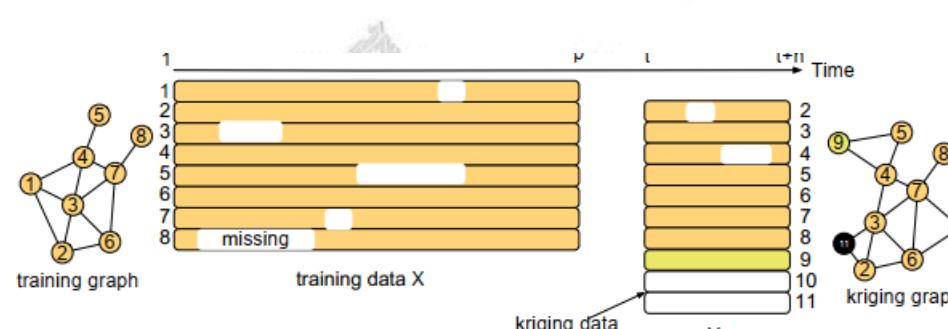
我们提出 Inductive Graph Neural Networks for Kriging (IGNNK)

在随机子图，而不是完整图上重建时空数据

可被推广到未曾见过的节点之上，甚至完全的新图之上



测试过程





子图采样算法

子图采样方法：

1. 从训练数据集中随机采样出一个子集
2. 根据采样的子集构建出一个图及其邻接矩阵
3. 将构建的子图随机遮蔽一部分节点（设为0，仿真位置节点）
4. 用图神经网络重建整个图的信号

Algorithm 1 Subgraph signal and random mask generation

Require: Historical data X from sampled locations over period $[1, p]$ (size $n \times p$).
Parameters: window length h , sample size each iteration S , and maximum iteration I_{\max} .

```
1: for iteration = 1 :  $I_{\max}$  do
2:   for sample = 1 :  $S$  do
3:     Generate random integers  $n_o$  (number of nodes selected as observed) and  $n_m$  (number of nodes selected as missing) with  $n_o + n_m \leq n$ .
4:     Randomly sample  $n_o + n_m$  indices without replacement from  $[1, n]$  to obtain  $I_{\text{sample}} = \{i^1, \dots, i^{n_o}, \dots, i^{n_o+n_m}\}$ .
5:     Randomly choose a time point  $j$  within range  $[1, p - h]$ . Let  $J_{\text{sample}} = [j, j + h)$ .
6:     Obtain submatrix signal  $X_{\text{sample}} = X[I_{\text{sample}}, J_{\text{sample}}]$  with size of  $(n_o + n_m) \times h$ .
7:     Construct adjacency matrix  $W_{\text{sample}} \in R^{(n_o+n_m) \times (n_o+n_m)}$  for nodes in  $I_{\text{sample}}$ .
8:     Generate a mask matrix  $M_{\text{sample}}$  of size  $(n_o + n_m) \times h$ ,  $M_{\text{sample}}[i, :] = \begin{cases} 1, & \text{if } i \in [1, n_o], \\ 0, & \text{otherwise.} \end{cases}$ 
9:   end for
10:  Use sets  $\{X_{1:S}\}, \{M_{1:S}\}, \{W_{1:S}\}$  to train GNNs.
11: end for
```



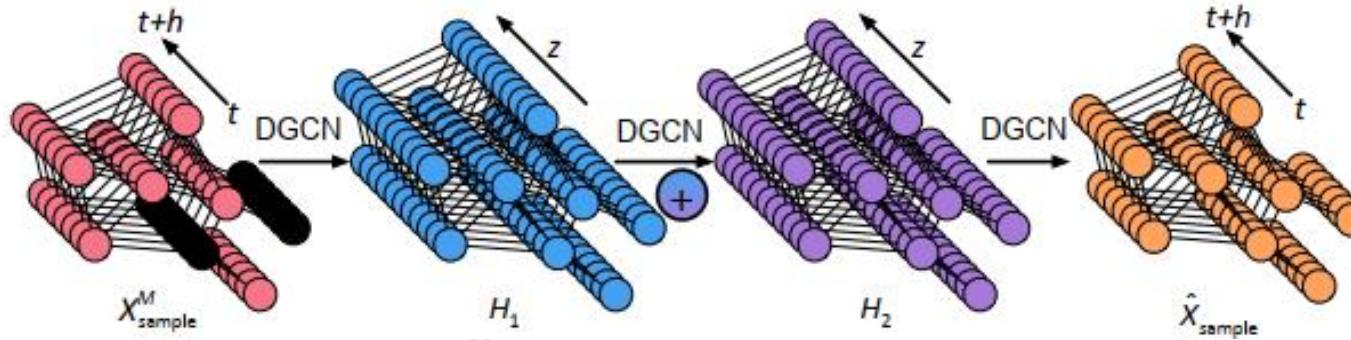
图神经网络结构



四川大学
SICHUAN UNIVERSITY

图神经网络结构：

为了建模数据的非对称空间关系，我们采用扩散图神经网络（有前向和后向图卷积核）



$$H_{l+1} = \sum_{k=1}^K T_k(\bar{W}_f) H_l \Theta_{b,l}^k + T_k(\bar{W}_b) H_l \Theta_{f,l}^k,$$

为了增强模型的通用能力，我们不仅重建遮蔽节点的信号，也重建整个随机子图的信号

$$J = \sum_{\text{sample}} \|\hat{X}_{\text{sample}} - X_{\text{sample}}\|_F^2$$



实验结果

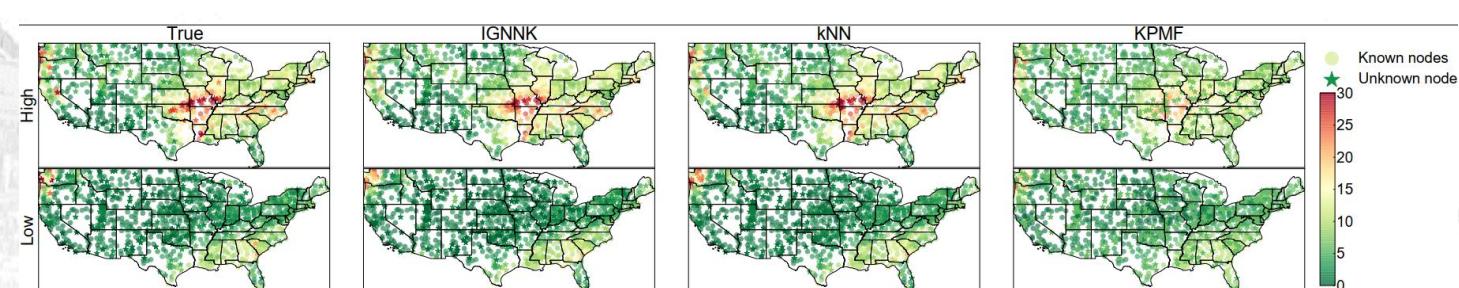
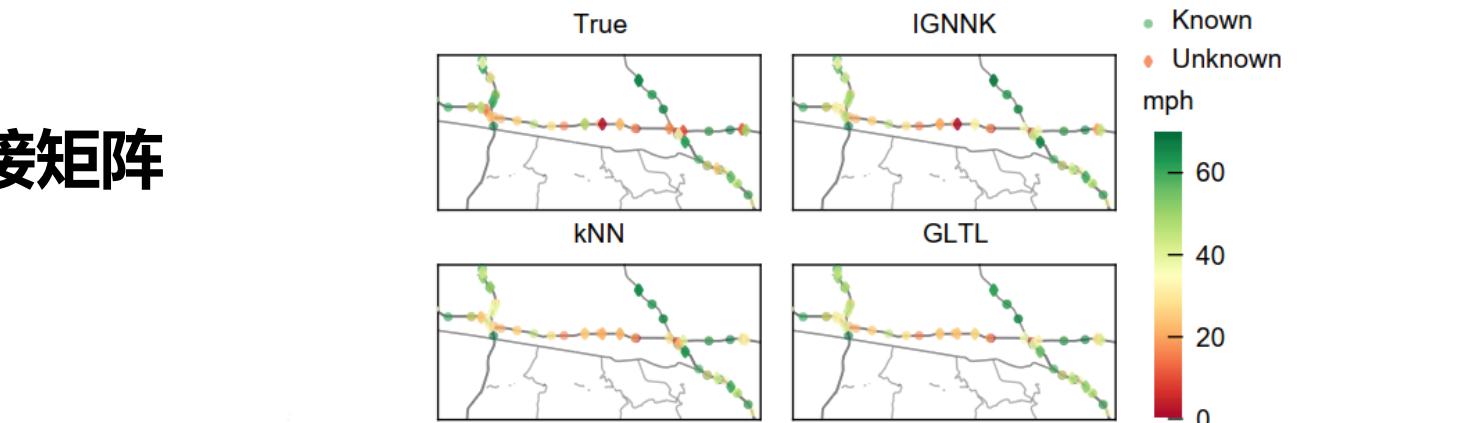
| Model | METR-LA | | | NREL | | | USHCN | | | SeData | | |
|----------|--------------|--------------|----------------|--------------|--------------|----------------|--------------|--------------|----------------|--------------|--------------|----------------|
| | RMSE | MAE | R ² |
| IGNNK | 9.048 | 5.941 | 0.827 | 3.261 | 1.597 | 0.885 | 3.205 | 2.063 | 0.771 | 6.863 | 4.241 | 0.537 |
| kNN | 11.071 | 6.927 | 0.741 | 4.192 | 2.850 | 0.810 | 3.400 | 2.086 | 0.742 | - | - | - |
| KPMF | 12.851 | 7.890 | 0.652 | 8.771 | 7.408 | 0.169 | 6.663 | 4.847 | 0.011 | 13.060 | 8.339 | -0.673 |
| GLTL | 9.668 | 6.559 | 0.803 | 4.840 | 3.372 | 0.747 | 5.047 | 3.396 | 0.432 | 6.989 | 4.285 | 0.520 |
| OKriging | - | - | - | 3.470 | 2.381 | 0.869 | 3.231 | 1.999 | 0.767 | - | - | - |

NREL, METR-LA以及USHCN的邻接矩阵

$$W_{ij} = \exp \left(- \left(\frac{\text{dist}(v_i, v_j)}{\sigma} \right)^2 \right)$$

SeData的邻接矩阵

$$W_{ij} = \begin{cases} 1, & \text{if } i \text{ and } j \text{ are neighbors,} \\ 0, & \text{otherwise.} \end{cases}$$





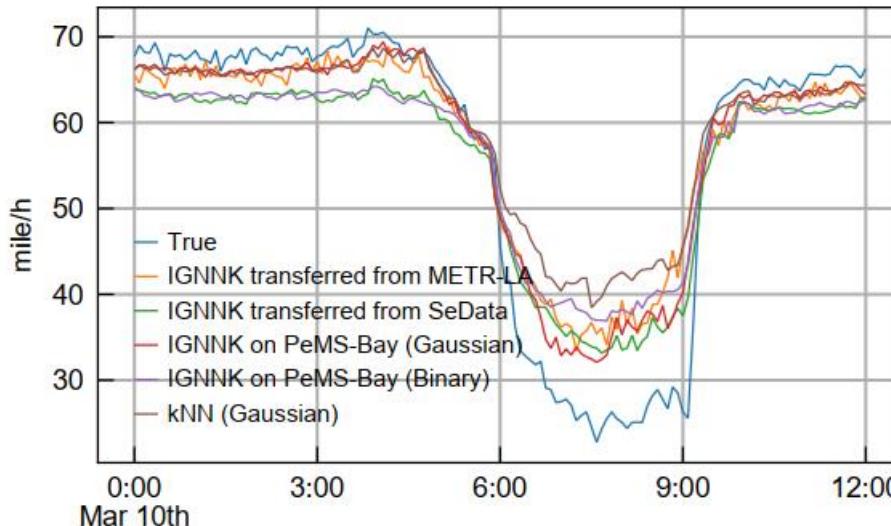
迁移学习结果



四川大学
SICHUAN UNIVERSITY

迁移到PeMS

| Model | Gaussian | | | | Binary | | | |
|----------------|--------------|--------------|--------------|--------------|--------|-------|--------|--------|
| | RMSE | MAE | MAPE | R^2 | RMSE | MAE | MAPE | R^2 |
| IGNNK | 6.093 | 3.663 | 8.16% | 0.574 | 9.245 | 5.394 | 13.26% | 0.161 |
| kNN | 7.431 | 4.245 | 9.13% | 0.458 | - | - | - | - |
| KPMF | 7.332 | 4.293 | 9.21% | 0.472 | 10.065 | 5.985 | 16.03% | 0.005 |
| GLTL | 8.846 | 4.486 | 10.25% | 0.232 | 8.504 | 4.962 | 12.24% | 0.290 |
| IGNNK Transfer | METR-LA | | | | SeData | | | |
| | 6.713 | 4.173 | 9.19% | 0.525 | 11.484 | 6.456 | 15.10% | -0.388 |



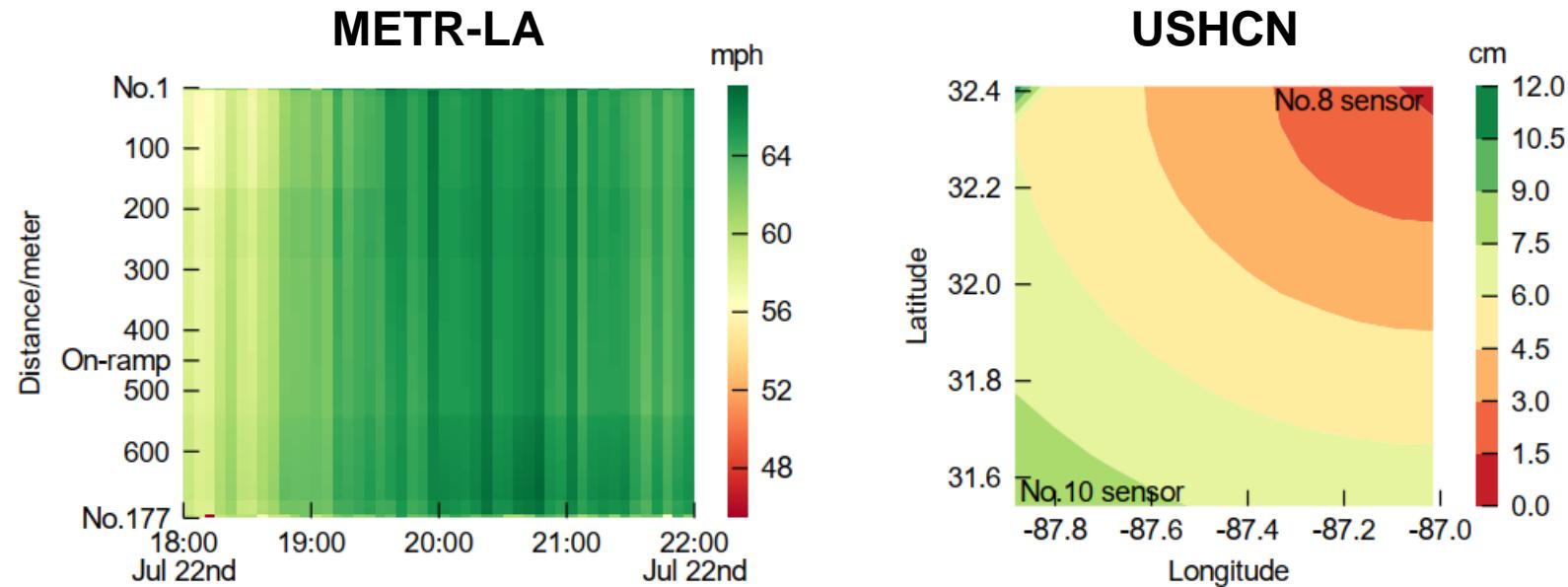
基于距离邻接矩阵的模型的迁移能力比基于二值0-1邻接矩阵的模型更强，说明了距离在空间插值问题中的重要性



虚拟节点仿真



四川大学
SICHUAN UNIVERSITY



IGNNK更多根据距离信息估计虚拟节点的信息

代码：[Kaimaoge/IGNNK \(github.com\)](https://github.com/Kaimaoge/IGNNK)

Wu, Y., Zhuang, D., Labbe, A., & Sun, L. (2021, May). Inductive Graph Neural Networks for Spatiotemporal Kriging. In *Proceedings of the AAAI Conference on Artificial Intelligence* (Vol. 35, No. 5, pp. 4478-4485).

目录

CONTENT

01 时空数据概述
03 高分辨率城市气候建模

02 图神经网络空间插值方法
04 解绑表征城市通勤数据建模

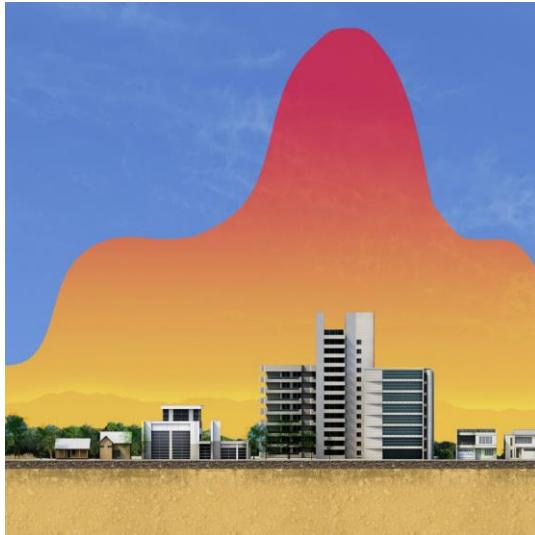




城市热岛效应-高分辨率气候建模



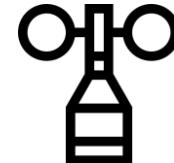
四川大学
SICHUAN UNIVERSITY



城市热岛效应是指城市的气温比郊外的气温高。这种现象通常是由于城市建设发展迅猛，生态环境失调，出现了裸地多，水面少；水泥、沥青地多，草地少；高层楼房多，树林少的现象，再加上车辆拥挤，人群稠密，大量二氧化碳排入空中，加强了温室效应。

缓解城市热岛效应需要掌握高分辨率 (<250m) 的城市局部气候变化信息。

观测手段：



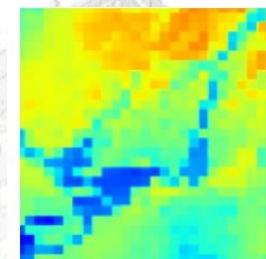
气象站点：分布过于稀疏



遥感观测：存在反演精度问题

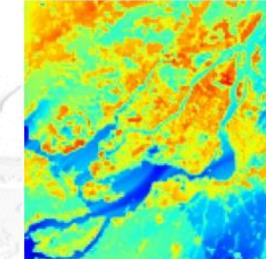
仿真：

2.5km



计算资源
X100

250m



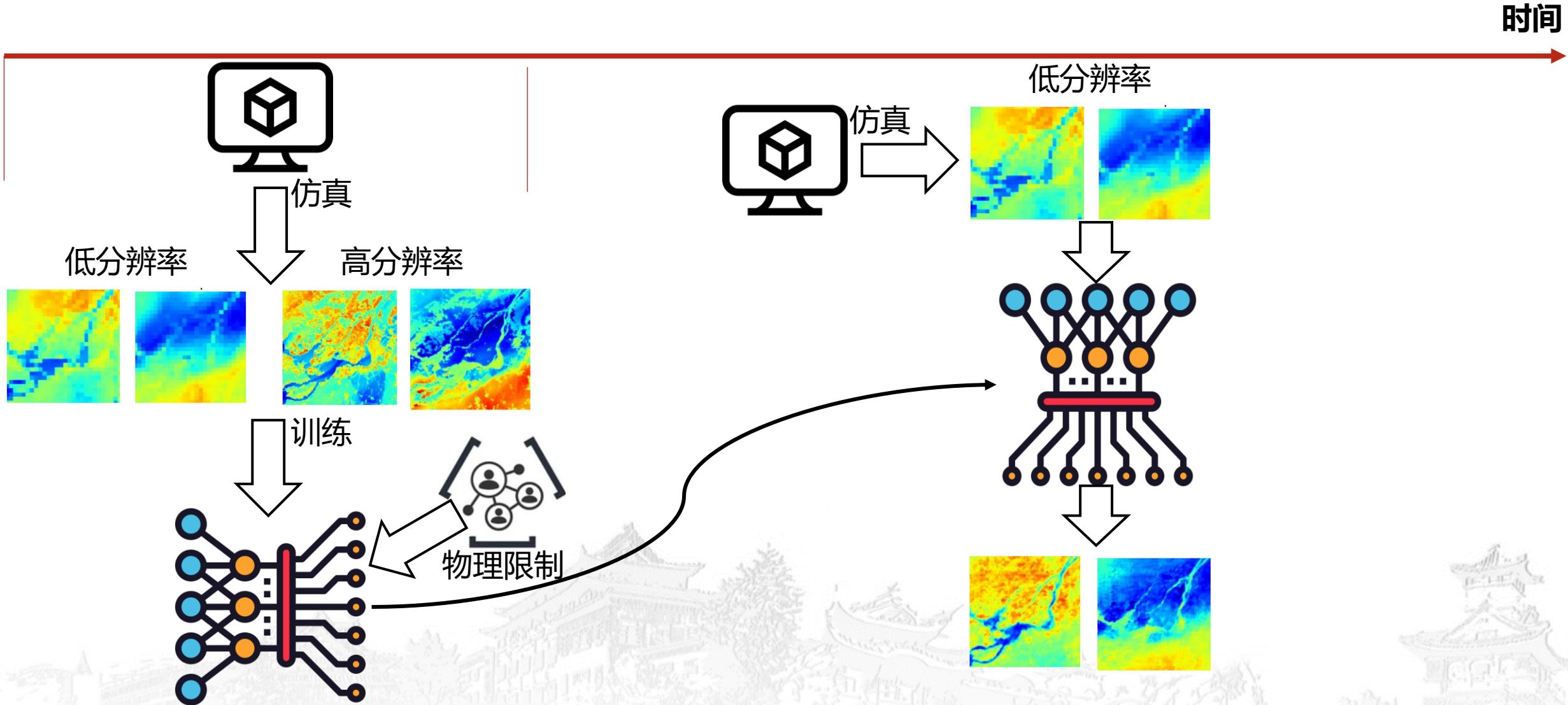


物理驱动深度学习的高分辨率气候建模



四川大学
SICHUAN UNIVERSITY

时间

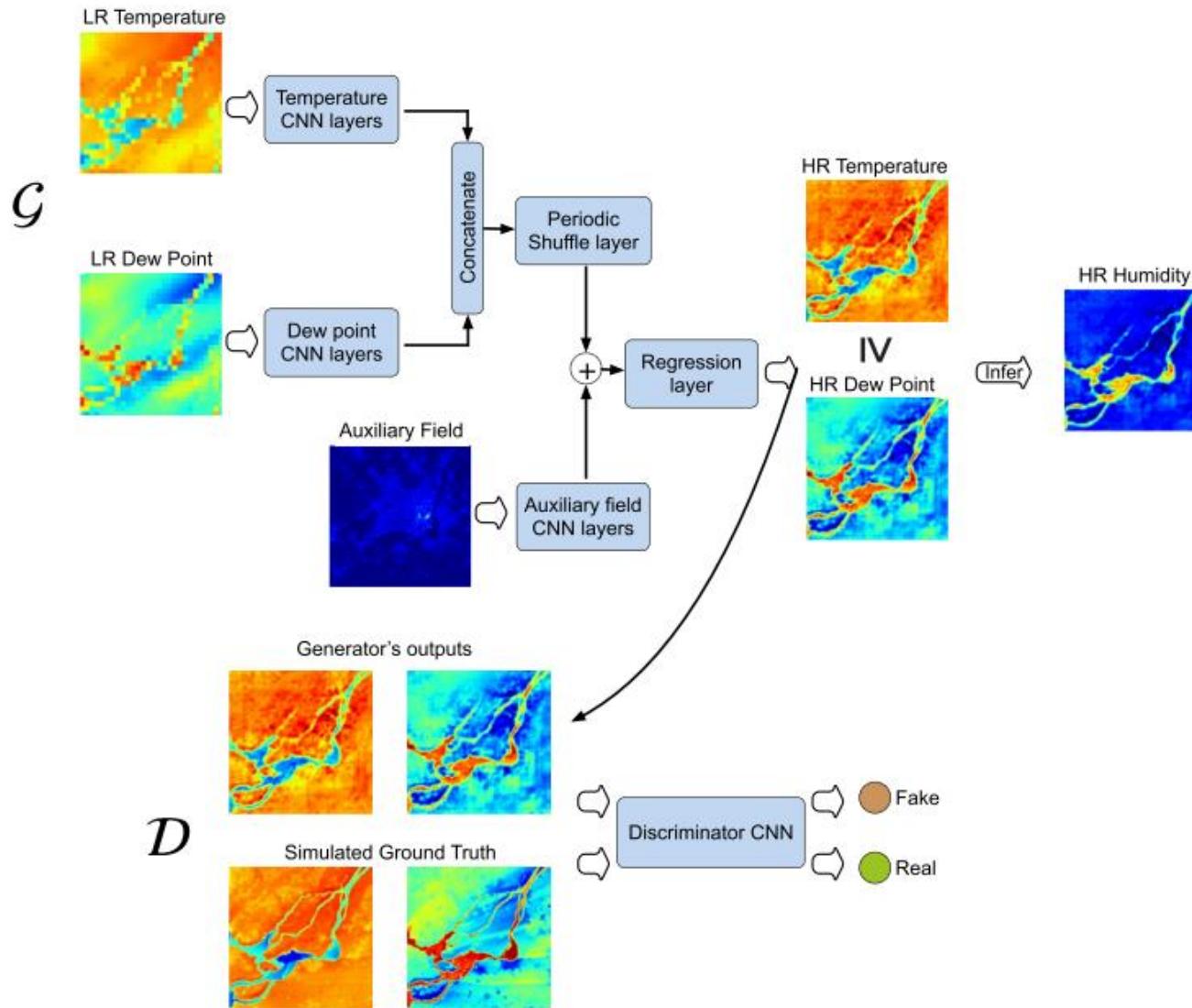




网络结构



四川大学
SICHUAN UNIVERSITY



训练损失函数

$$L_{G\theta} = L_{MSE} + \gamma L_{TV} - \lambda L_W,$$

逐像素损
失函数

全变化损
失函数

WGAN损
失函数

无法确保消除棋盘
气候场的 伪影
物理一致性

确保物
理一致
性



实验结果



四川大学
SICHUAN UNIVERSITY

为了评估使用辅助场、对抗训练和物理约束的好处，这篇文章评估了四种深度学习变体：

a) SPCNN（亚像素卷积神经网络），(b) FSPCNN（融合亚像素卷积神经网络）、(c) FSPCGAN-S (S 指的是单独的模型) 和 (d) FSPCGAN。

| | 是否有鉴别器 | 是否考虑辅助变量 | 是否联合训练温度和露点场 |
|-----------|--------|----------|--------------|
| SPCNN | 无 | 无 | 有 |
| FSPCNN | 无 | 有 | 有 |
| FSPCGAN-S | 有 | 有 | 无 |
| FSPCGAN | 有 | 有 | 有 |

Table 1

Average Errors and Their Standard Deviations for Temperature and Dew Point for the Last 50 Learning Steps; and for Inferred Relative Humidity for the Last Step

| Models | Temperature (°C) | | Dew point (°C) | | Relative humidity (%) | |
|-----------|---------------------------------------|---------------------------------------|---------------------------------------|---------------------------------------|-----------------------|---------------|
| | MAE | RMSE | MAE | RMSE | MAE | RMSE |
| Cokriging | 1.2818 | 1.8468 | 1.1203 | 1.6405 | 7.0178 | 9.6141 |
| SPCNN | 0.7900 ± 0.0206 | 1.0959 ± 0.0216 | 0.9703 ± 0.0236 | 1.4362 ± 0.0265 | 5.3736 | 7.3140 |
| FSPCNN | 0.7892 ± 0.0350 | 1.0851 ± 0.0369 | 0.9674 ± 0.0187 | 1.4326 ± 0.0294 | 5.1908 | 7.1213 |
| FSPCGAN-S | 0.7920 ± 0.0340 | 1.0901 ± 0.0364 | 0.9698 ± 0.0228 | 1.4158 ± 0.0274 | 5.6680 | 7.5672 |
| FSPCGAN | 0.8828 ± 0.0601 | 1.1996 ± 0.0625 | 0.9940 ± 0.0310 | 1.4308 ± 0.0345 | 5.1496 | 7.0700 |

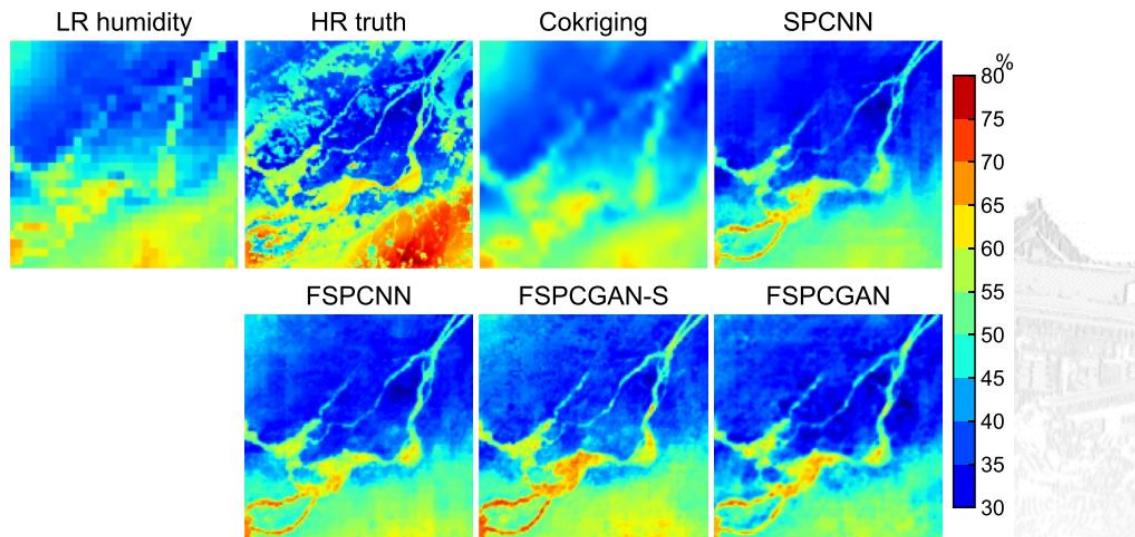
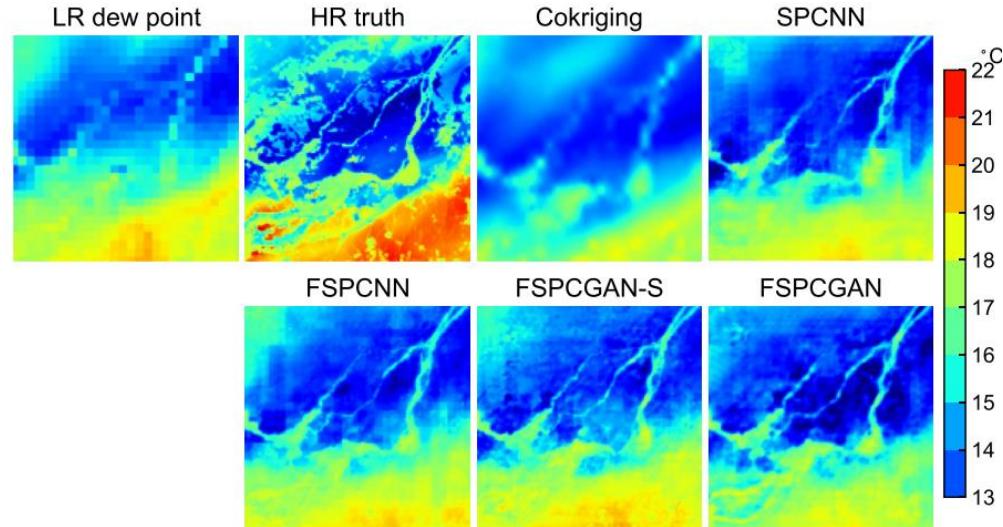
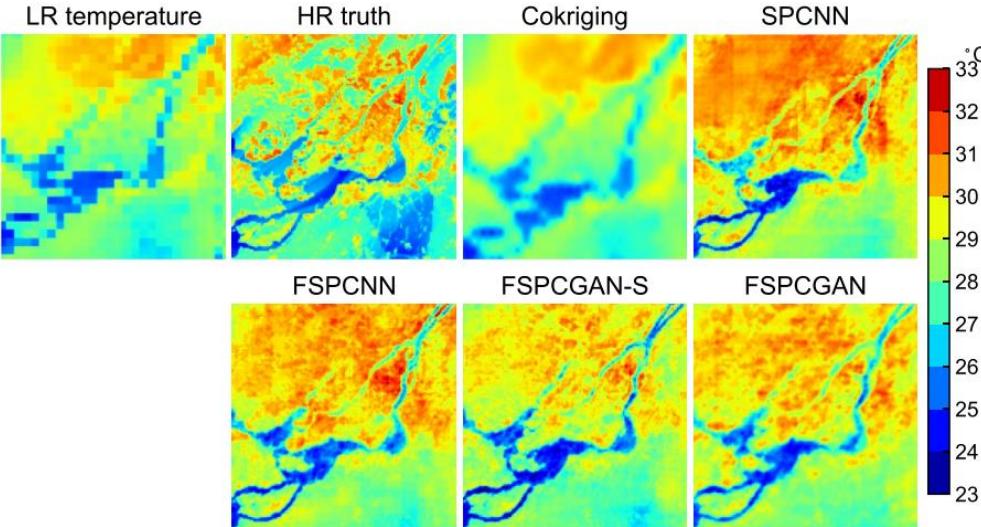
Note. The bold values used to indicate the best performance. MAE, mean absolute error; RMSE, root mean squared error.



实验结果



四川大学
SICHUAN UNIVERSITY



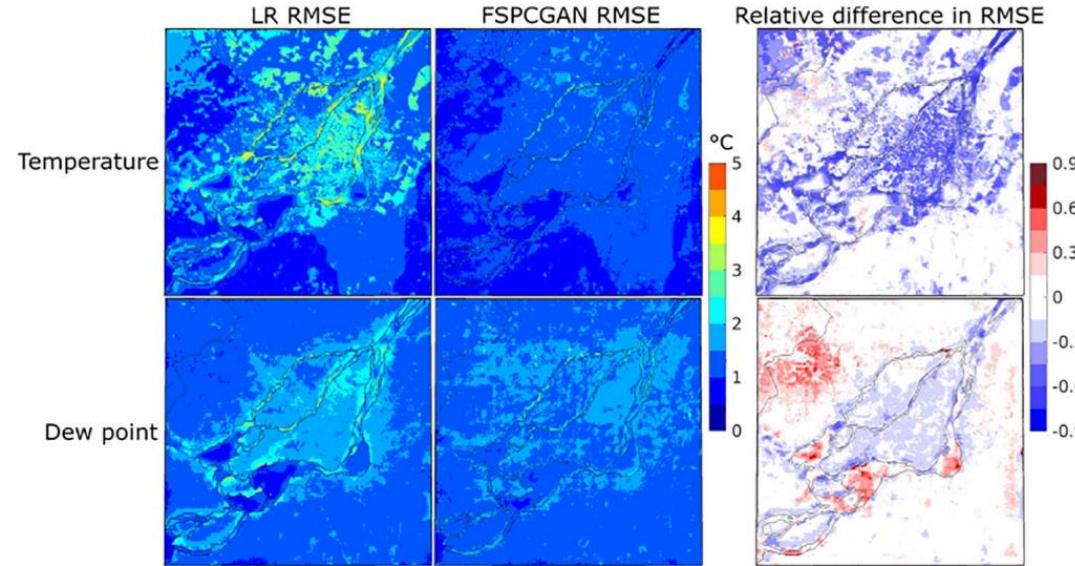
真实HR包含许多 LR 模拟无法重现的小规模特征。正如预期的那样，简单的 Cokriging 方法完全消除了这些特征。与 SPCNN 相比，具有辅助场信息的 FSPCNN 可以额外为输出引入更多的小规模特征。最后，FSPCGAN-S 和 FSPCGAN 的 SR 结果与更精细的小尺度特征相关联。FSPCGAN 的结果在定性上似乎与 HR 模拟更具可比性，尽管它有时会比其他深度学习模型给出更高的像素级错误。



实验结果



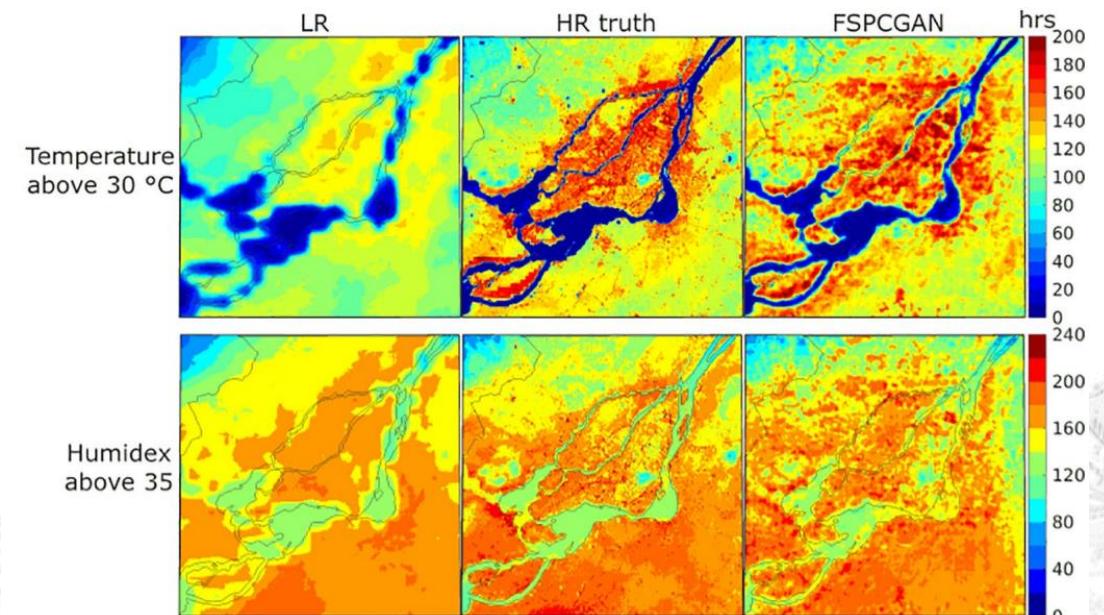
四川大学
SICHUAN UNIVERSITY



LR 数据在土地覆盖相对同质的地区（例如东南地区）表现相当好 ($\text{RMSE} < 2^{\circ}\text{C}$)，而在高度异质的地区，诸如 FSPGAN 之类的深度学习方法能够产生更接近 HR 真相的超分辨率数据，通常将温度的 RMSE 降低 50% 以上，露点的 RMSE 降低约 20%。

在常发炎热区域预测上，FSPGAN与真值基本保持一致

Wu, Y., Teufel, B., Sushama, L., Belair, S., & Sun, L. (2021). Deep Learning-Based Super-Resolution Climate Simulator-Emulator Framework for Urban Heat Studies. *Geophysical Research Letters*, 48(19), e2021GL094737.



目录

CONTENT

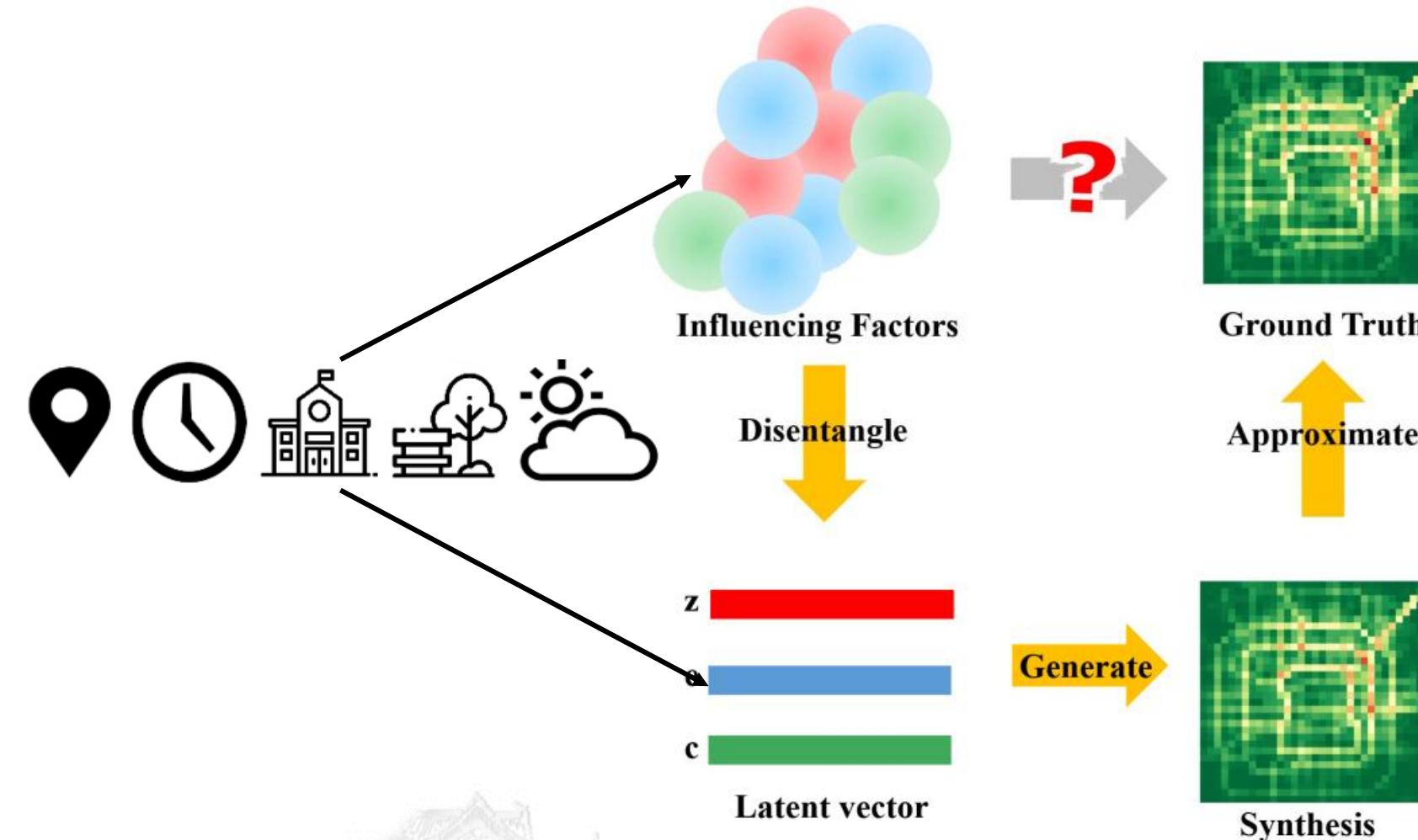
01 时空数据概述
03 高分辨率城市气候建模

02 图神经网络空间插值方法
04 解绑表征城市通勤数据建模





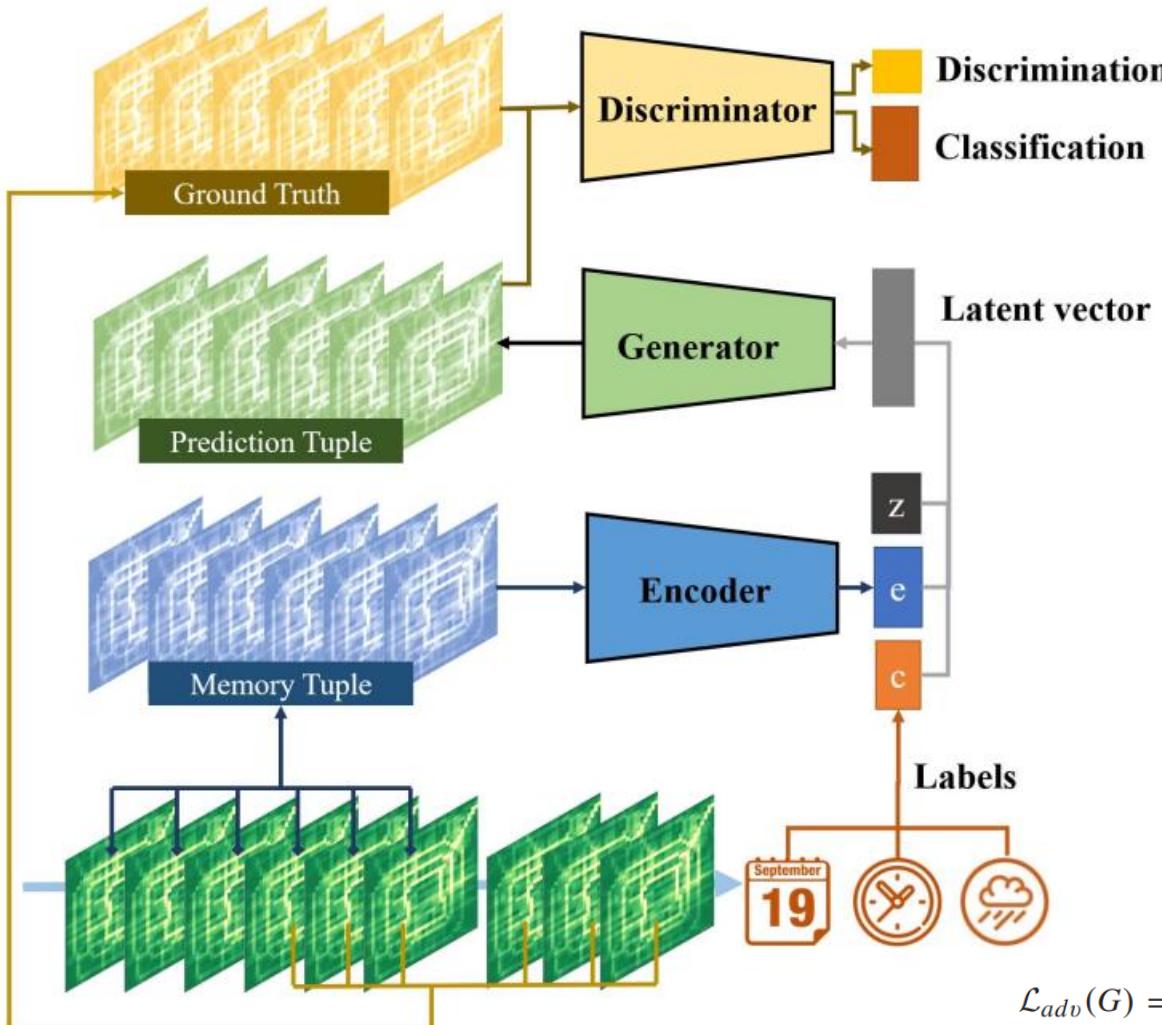
城市通勤数据的成因



城市的人流数据是由多种因素共同作用的结果，这些多种作用**纠缠在一起**，难以对单个因素的影响进行分析。因此我们采用**解绑表征**将这些因素的**纠缠解开**，以便进行单个因素定量定性分析。



网络结构



$$\hat{x} = G_{\theta}(c_t, z_t, e_t)$$

- c:** 一些定性的因素，日期、时间以及天气等
- e:** 从历史数据中提取的因素
- Z:** 随机因素

分类器损失函数

$$\begin{aligned}\mathcal{L}_{adv}(D) = & -\mathbb{E}_{\hat{x} \sim P_{data}} [\log D(x)] \\ & - \mathbb{E}_{\substack{z \sim P_{noise}, \\ x \sim P_{data}, \\ c \sim P_{label}}} [\log(1 - D(G(z, c, E(x)))]\end{aligned}$$

$$\begin{aligned}\mathcal{L}_{clc}(D) = & \mathbb{E}_{\hat{x} \sim P_{data}} [\log P_D[\text{label}_x | x]] \\ & + \mathbb{E}_{\substack{z \sim P_{noise}, \\ x \sim P_{data}, \\ c \sim P_{label}}} [\log(P_D[c | G(z, c, E(x))])]\end{aligned}$$

$$\mathcal{L}(D) = \lambda_{adv} \mathcal{L}_{adv}(D) + \lambda_{clc} \mathcal{L}_{clc}(D)$$

生成器损失函数

$$\mathcal{L}_{prd}(G) = \mathbb{E}_{\substack{z \sim P_{noise}, \\ x \sim P_{data}, \\ c \sim P_{label}}} \sqrt{\frac{1}{N} \sum_{i=1:N} (\hat{x}_i - G(z, c, E(x_i)))^2}$$

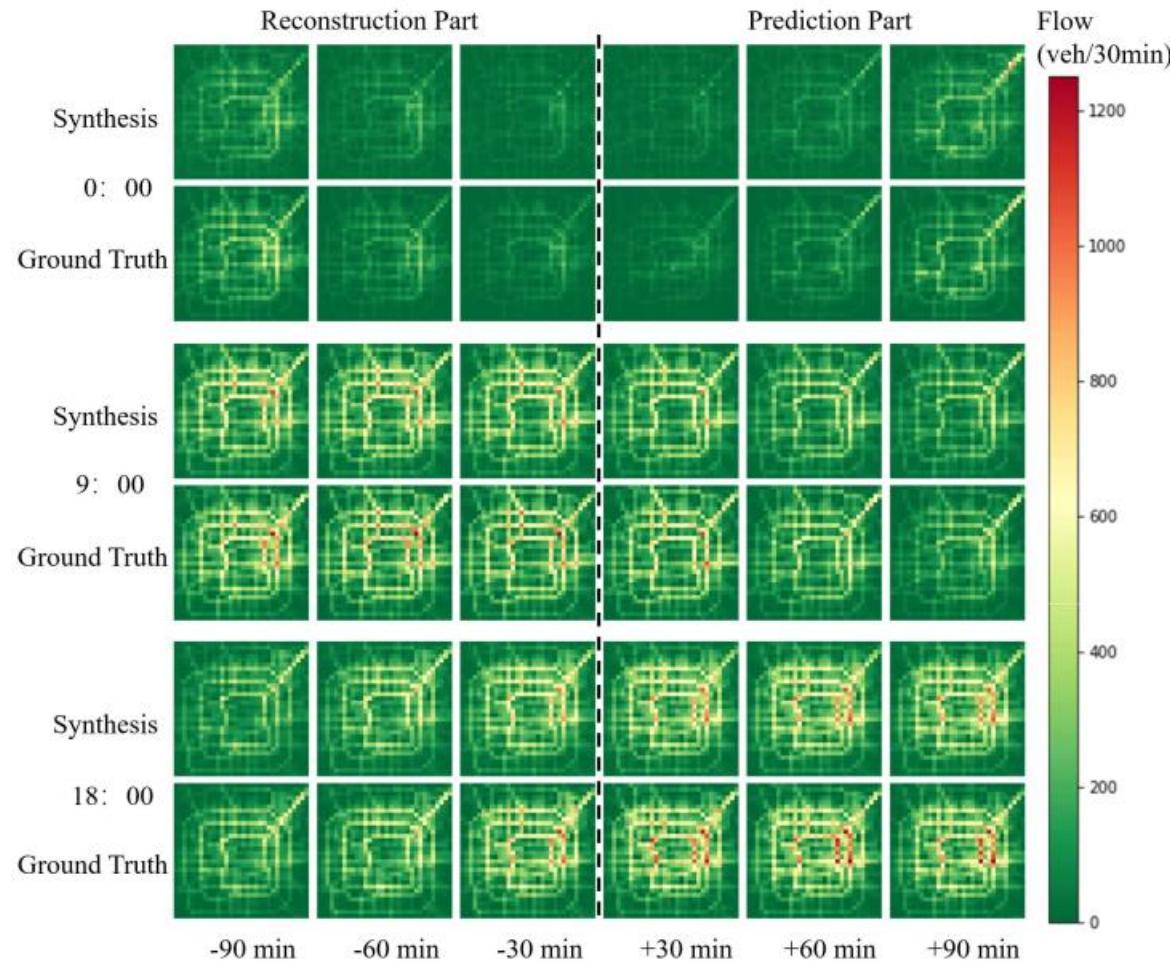
$$\mathcal{L}(G) = \lambda_{adv} \mathcal{L}_{adv}(G) + \lambda_{clc} \mathcal{L}_{clc}(G) + \lambda_{prd} \mathcal{L}_{prd}(G)$$



网络结构



四川大学
SICHUAN UNIVERSITY



| Methods | HA | ARIMA | LSTM | ST-ResNet | GAN | Auto-Encoder | ST-GAN |
|----------|-------|-------|-------|-----------|-------|--------------|--------|
| MAE | 43.21 | 18.54 | 20.13 | 14.28 | 17.51 | 16.17 | 13.6 |
| MAPE | 32.4% | 13.9% | 15.1% | 10.7% | 13.1% | 12.5% | 10.2% |
| RMSE | 51.23 | 20.8 | 24.23 | 17.32 | 19.87 | 18.62 | 16.82 |
| Time (s) | 2.47 | 1.72 | 5.94 | 4.87 | 3.52 | 3.52 | 3.52 |

ST-GAN方法表现优于ST-ResNet等现有方法，且计算时间更少



仿真结果



四川大學
SICHUAN UNIVERSITY

| label data | 1:00 | 3:00 | 5:00 | 7:00 | 9:00 | 11:00 | 13:00 | 15:00 | 17:00 | 19:00 | 21:00 | 23:00 |
|---------------|------|------|------|------|------|-------|-------|-------|-------|-------|-------|-------|
| | | | | | | | | | | | | |
| | | | | | | | | | | | | |
| | | | | | | | | | | | | |
| | | | | | | | | | | | | |
| | | | | | | | | | | | | |
| | | | | | | | | | | | | |



仿真结果



四川大学
SICHUAN UNIVERSITY

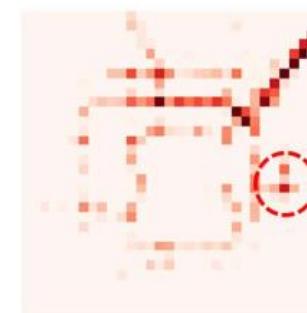


在S12机场快速路增加出租车流量

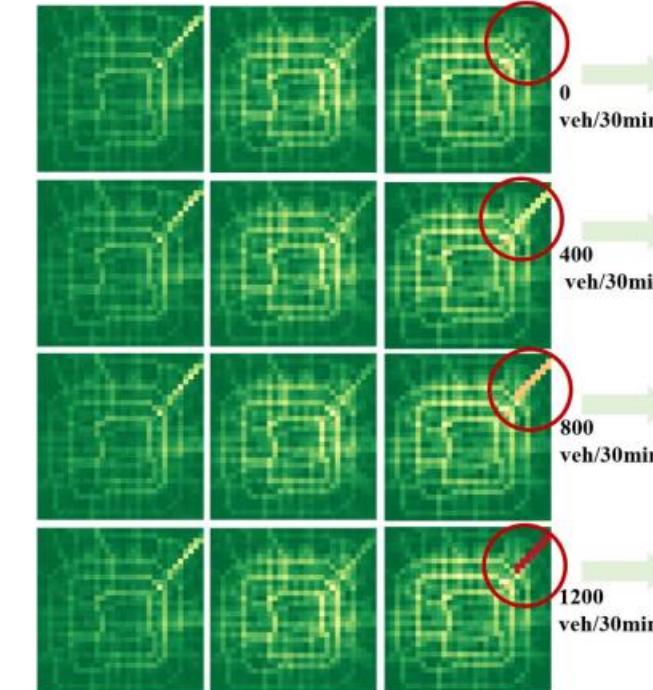
400 veh/30min

800 veh/30min

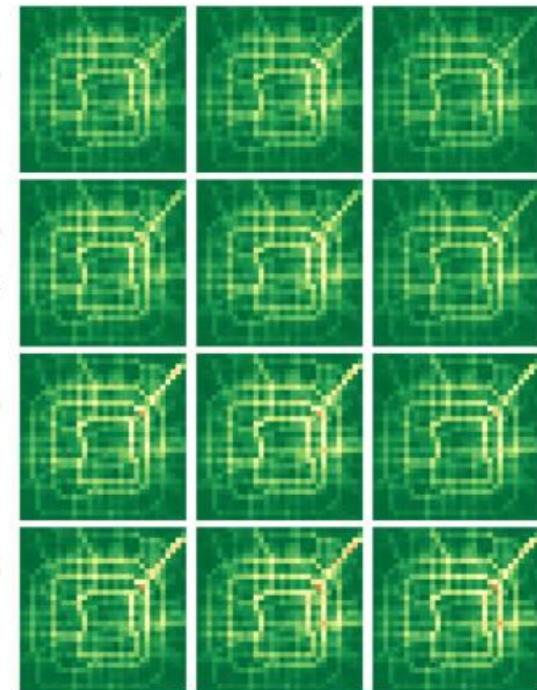
1200 veh/30min



Historical Records



Future Prediction



The accidental increasing flow
in the S12 express road

512机场增加的流量会造成城市整体流量增加，且主要增加在CBD区域，说明解绑表征方法学习到了**城市通勤的语义学特征**。



代表性成果



四川大学
SICHUAN UNIVERSITY

时空大数据建模

- [1] Wu, Y., Zhuang, D., Labbe, A. and Sun, L., 2021, May. Inductive Graph Neural Networks for Spatiotemporal Kriging. In Proceedings of the AAAI Conference on Artificial Intelligence (Vol. 35, No. 5, pp. 4478-4485). ([CCF A类会议](#))
- [2] Zhang, H., Wu, Y.*, Tan, H.*., Dong, H., Ding, F. and Ran, B., 2020. Understanding and modeling urban mobility dynamics via disentangled representation learning. *IEEE Transactions on Intelligent Transportation Systems*. ([中科院1区Top](#))
- [3] Wu, Y., Teufel, B., Sushama, L., Belair, S. and Sun, L., 2021. Deep Learning-Based Super-Resolution Climate Simulator-Emulator Framework for Urban Heat Studies. *Geophysical Research Letters*, 48(19), p.e2021GL094737. ([中科院2区Top](#))
- [4] Wu, Y., Tan, H., Qin, L., Ran, B. and Jiang, Z., 2018. A hybrid deep learning based traffic flow prediction method and its understanding. *Transportation Research Part C: Emerging Technologies*, 90, pp.166-180. ([中科院1区Top, ESI高被引论文](#))
- [5] Wu, Y., Tan, H., Li, Y., Zhang, J. and Chen, X., 2018. A fused CP factorization method for incomplete tensors. *IEEE Transactions on Neural Networks and Learning Systems*, 30(3), pp.751-764. ([中科院1区Top](#))
- [6] Wu, Y., Tan, H., Li, Y., Li, F. and He, H., 2017. Robust tensor decomposition based on Cauchy distribution and its applications. *Neurocomputing*, 223, pp.107-117. ([中科院2区Top](#))
- [7] Tan, H., Wu, Y., Shen, B., Jin, P.J. and Ran, B., 2016. Short-term traffic prediction based on dynamic tensor completion. *IEEE Transactions on Intelligent Transportation Systems*, 17(8), pp.2123-2133. ([中科院1区Top](#))

深度强化学习应用

- [1] Wang, Y., Tan, H., Wu, Y.* and Peng, J., 2020. Hybrid electric vehicle energy management with computer vision and deep reinforcement learning. *IEEE Transactions on Industrial Informatics*, 17(6), pp.3857-3868. ([中科院1区Top](#))
- [2] Wu, Y., Tan, H., Qin, L. and Ran, B., 2020. Differential variable speed limits control for freeway recurrent bottlenecks via deep actor-critic algorithm. *Transportation research part C: emerging technologies*, 117, p.102649. ([中科院1区Top](#))
- [3] Lian, R., Tan, H., Peng, J., Li, Q. and Wu, Y.*, 2020. Cross-type transfer for deep reinforcement learning based hybrid electric vehicle energy management. *IEEE Transactions on Vehicular Technology*, 69(8), pp.8367-8380. ([中科院2区Top](#))
- [4] Lian, R., Peng, J.*., Wu, Y.*., Tan, H. and Zhang, H., 2020. Rule-interposing deep reinforcement learning based energy management strategy for power-split hybrid electric vehicle. *Energy*, 197, p.117297. ([中科院1区Top](#))
- [5] Wu, Y., Tan, H., Peng, J., Zhang, H. and He, H., 2019. Deep reinforcement learning of energy management with continuous control strategy and traffic information for a series-parallel plug-in hybrid electric bus. *Applied energy*, 247, pp.454-466. ([中科院1区Top, ESI高被引论文](#))



四川大學
SICHUAN UNIVERSITY

感谢 Q&A

Email: wuyk0@scu.edu.cn

个人网站: Kaimaoge.github.io