

# 基于图卷积记忆网络对珠海臭氧时空预测\*

孙磊<sup>1,2</sup>, 蓝玉峰<sup>1</sup>, 梁秀姬<sup>1</sup>, 孙弦<sup>1</sup>,  
聂会文<sup>2</sup>, 苏烨康<sup>1</sup>, 贺芸萍<sup>1</sup>, 王静<sup>1</sup>, 夏冬<sup>1,2</sup>

1. 珠海市公共气象服务中心, 广东 珠海 519000  
2. 珠澳气象创新与应用研究中心, 广东 珠海 519000

**摘要:** 臭氧(O<sub>3</sub>)已成为影响珠三角(乃至广东)空气质量达标的首要因素。数据驱动的统计模型(较数值模式)虽展现出改进的预报能力,但多数未能解析站点数据(非欧结构)之间的空间依赖性。本文基于珠海市6个环保国控站及其周边气象站监测数据,通过构建时空协同的图卷积记忆网络(GCN-LSTM)开展多站点未来3天逐小时O<sub>3</sub>质量浓度预报。结果表明:GCN\_LSTM在不同预报时效均准确还原了O<sub>3</sub>的年、季节和昼夜变化特征,但对日变化的预报技巧随预报时效增加下降明显。通过与业务数值模式(GRACEs)和长短期记忆网络(LSTM)对比发现:GCN-LSTM表现最优,其72 h预报时效内RMSE和R均值分别为27.13 μg/m<sup>3</sup>和0.64, LSTM表现次之(RMSE=28.44 μg/m<sup>3</sup>; R=0.61),而GRACEs与统计模型存在明显差距(RMSE=40.93 μg/m<sup>3</sup>; R=0.33)。此外,相较于LSTM,GCN-LSTM全局考虑所有站点及其之间的相互联系,不仅将计算速度提高了71%,而且在不同站点的表现也更为优秀和稳定,同时捕捉秋季O<sub>3</sub>污染事件的能力也有所提高。最后,敏感性实验揭示出考虑相关性较高的变量作为预报因子可以提高模型能力。

**关键词:** 臭氧; 时空预报; 机器学习; 图卷积记忆网络

中图分类号: P467 文献标志码: A 文章编号: 2097-0137(2024)03-0048-12

## The spatio-temporal prediction of ozone in Zhuhai based on graph convolutional memory network

SUN Lei<sup>1,2</sup>, LAN Yufeng<sup>1</sup>, LIANG Xiuji<sup>1</sup>, SUN Xian<sup>1</sup>,  
NIE Huiwen<sup>2</sup>, SU Yekang<sup>1</sup>, HE Yunping<sup>1</sup>, WANG Jing<sup>1</sup>, XIA Dong<sup>1,2</sup>

1. Zhuhai Public Meteorological Service Center, Zhuhai 519000, China

2. Zhuhai-Macao Collaborative Research Center for Meteorological Innovation and Application, Zhuhai 519000, China

**Abstract:** Ozone (O<sub>3</sub>) has become the primary factor affecting air quality over the Pearl River Delta and even the entire Guangdong Province. Although data-driven statistical models have shown improved forecast capabilities compared to numerical models, most of them operate grid-by-grid and cannot resolve the spatial dependence between site data of non-Euclidean structures. Based on in-situ measurements from national environmental stations and surrounding weather stations in Zhuhai, this study performs hourly O<sub>3</sub> concentration forecasts for up to three days over multiple sites by constructing a graph convolution memory network (GCN-LSTM). The results show that GCN\_LSTM forecasts at different lead times could accurately reproduce the annual, seasonal, and diurnal variations of O<sub>3</sub>, but the capa-

\* 收稿日期: 2023-12-17

录用日期: 2024-01-09

网络首发日期: 2024-03-20

基金项目: 广东省气象局科技项目(GRMC2022Q16)

作者简介: 孙磊(1992年生), 男; 研究方向: 空气污染预报与来源解析; E-mail: sunleimodel@gmail.com

通信作者: 蓝玉峰(1990年生), 女; 研究方向: 空气污染预报与来源解析; E-mail: lyfnuist@163.com

bility of capturing daily variations decreases significantly with the increase in lead time. Further comparisons with the operational numerical model (GRACEs) and Long Short-Term Memory (LSTM) reveal that GCN-LSTM performs the best, with mean RMSE=27.13  $\mu\text{g}/\text{m}^3$  and  $R=0.64$ , LSTM is the second (RMSE=28.44  $\mu\text{g}/\text{m}^3$ ;  $R=0.61$ ), and GRACEs presents distinct results (RMSE = 40.93  $\mu\text{g}/\text{m}^3$ ;  $R=0.33$ ) in 72h forecasting. Compared with LSTM, GCN-LSTM considers all sites and their interconnections, it not only increases the calculation speed by 71% but also performs better and more stably over different sites. Moreover, it is also optimal for capturing  $\text{O}_3$  pollution events in cold seasons. Additional sensitivity experiments reveal that considering more correlated variables improves forecasting capabilities.

**Key words:** ozone ( $\text{O}_3$ ); spatial-temporal forecast; machine learning; graph convolution memory network

经过多年治理,珠三角大气中的一次污染物已经得到有效控制。但二次污染物臭氧( $\text{O}_3$ )的年均质量浓度却上升了25%,已经成为珠三角甚至广东省首要大气环境污染问题(吴乃庚等,2020)。 $\text{O}_3$ 是强氧化剂,其在近地面主要是由氮氧化物( $\text{NO}_x$ )和挥发性有机物(VOCs)等污染物在太阳光的照射下,经系列光化学反应后生成。高浓度 $\text{O}_3$ 不仅会对人体产生危害(例如刺激眼睛、呼吸道粘膜甚至中枢神经),还会对植物、农作物与生态系统产生一定程度影响(Krzyzanowski et al., 2008)。

$\text{O}_3$ 预报可以分为2类(Zhang et al., 2012):一类是基于物理化学过程的数值模式预报,从排放、平流输送、湍流扩散、化学反应到清除的完备过程,已被广泛使用和业务化(邓雪娇等,2016;王自发等,2009;周广强等,2016;Eder et al., 2006),但对高性能计算要求高,预报能力受到多源不确定性的影响(李勋等,2020;Li et al., 2017;Bassett et al., 2020;Kitayama et al., 2019);另一类则是统计预报,其多数基于气象环境要素对 $\text{O}_3$ 的影响,利用传统的回归统计方法或日益流行的机器学习方法,构建了一系列统计预报模型,有效地提高了 $\text{O}_3$ 的预报能力(夏冬等,2012;Gao et al., 2018;Pak et al., 2018;Zhang et al., 2017;杨雨佳等,2023)。

$\text{O}_3$ 统计预报仍存在局限性。首先,多数模型依赖于数值天气模式提供的未来气象条件预测构建回归关系,因此业务应用的便捷性受到很大限制(Cabaneros et al., 2019)。其次,近年来提出的以长短期记忆网络(LSTM)为代表的深度学习模型,虽有效克服了循环神经网络(RNN)在解析长期依赖性方面的不足,在各类时序预测的应用方面表现出不错的能力,但属于单点模型,未能有效结合不同站点在空间上的相互关联和影响。此外,由于站点数据属于非欧结构,目前广泛应用于空间

特征的卷积神经网络CNN也无法使用(Sun et al., 2023)。最近,一种以图结构为研究对象的图卷积神经网络(GCN)被提出,为一些非欧数据空间特征的提取提供了新的解决思路,并通过与记忆网络(如LSTM)进行关联,成功在交通流量预测、社交网络分析预测等多个领域得到有效应用(Yu et al., 2017)。需要注意的是,该技术虽在空气质量预报方面取得进展,但是目前应用多是针对颗粒物(Qi et al., 2019; Xiao et al., 2022; 祁柏林等, 2021),尚未有研究对其在 $\text{O}_3$ 预报的适用性进行探索,这也是本文的首要出发点。

珠海作为粤港澳大湾区核心城市,一直以环境优美,生态良好享誉全国。但从珠海市2019—2021年的环境监测数据来看, $\text{O}_3$ 作为首要污染物比例高达85%,2021年比例更是高达94%(赵燕等,2022)。可以说, $\text{O}_3$ 污染是制约珠海AQI(air quality index)达标率提升的最重要因素。

综上所述,本文选取珠海市臭氧质量浓度( $\rho(\text{O}_3)$ ,下同)预报为研究对象,基于历史逐小时6个环保国控站及其周边气象站监测信息,通过构建时空协同的图卷积记忆网络,具体分别利用GCN和LSTM进行空间特征和多尺度时间特征解析,实现各站点未来3天逐小时 $\text{O}_3$ 质量浓度预报。

## 1 数据与方法

### 1.1 数据

本文所用数据来自珠海市6个环保国控站及其周边气象站(图1),时间分辨率为逐小时,跨度为2019年1月至2023年9月。环保国控站点提供的环境质量要素,包括: $\text{O}_3$ 、 $\text{NO}_2$ 、颗粒物( $\text{PM}_{2.5}$ 和 $\text{PM}_{10}$ ),其中 $\text{O}_3$ 同为预报因子和预报对象,其他要素则为预报因子,缺测则使用临近时间或周边站点的记录进行填补。周边气象站提供的气象要素,包括:

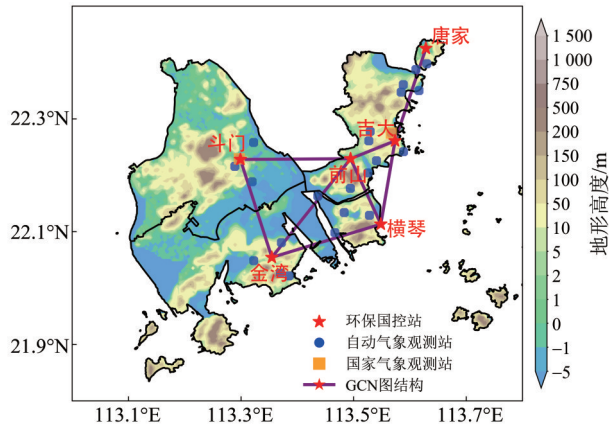


图1 研究区域观测站点的地理位置叠加地形高度分布以及GCN模型的图结构

Fig. 1 The geographical location of the measurements sites overlaid with the topographic heights as well as the graph structure of the GCN model

近地气温 T2、近地平均风速 WS10、日照时数 SH (来自于2个国家气象站, 图1), 这些要素都作为预报因子, 通过最邻近插值方法插值到对应环保国控站(若最邻近站点记录缺失, 则使用次邻近站点, 依此类推)。

本文选取中国气象局广州热带海洋气象研究

所研发的华南区域大气成分数值模式系统(简称 GRACEs; 邓雪娇等,2016)进行对比。该模式采用水平三重嵌套, 气象条件来自于国产数值模式 CMA-GD(陈德辉等,2012), 使用的排放源清单在融合多个源清单的基础上, 利用大气成分卫星遥感资料和站点观测进行了优化(邓雪娇等, 2016)。本文使用的模式数据来自于李婷苑等(2021)和孙弦等(2023), 具体为2019年每日20时起报的4个环保国控站点(即为唐家、前山、吉大和斗门)未来3 d逐小时 $\rho(\text{O}_3)$ 数据。

### 1.2 模型

1.2.1 问题概述 本文构建深度学习预测模型去有效连接GCN与LSTM, 前者用于空间特征提取, 后者用于多尺度时间解析。图2展示了 $\text{O}_3$ 时空预测的主要流程, 可以用如下表达式进行概括:

$$[O_{3T+1}, O_{3T+2}, O_{3T+3}, \dots, O_{3T+72}] = f([X_{T-N}, \dots, X_{T-2}, X_{T-1}, X_T]; G),$$

其中 $f$ 为构建的深度学习预报模型, 预报输入为不同环保站点组成的无向图 $G$ (图1)和据此构建的过去 $N$ h的预报因子向量 $X_t$ , 预报输出为各站点未来72 h $\rho(\text{O}_3)$ 。本文设置 $N=24$ , 即利用过去1 d的信息去开展未来3 d的预报。

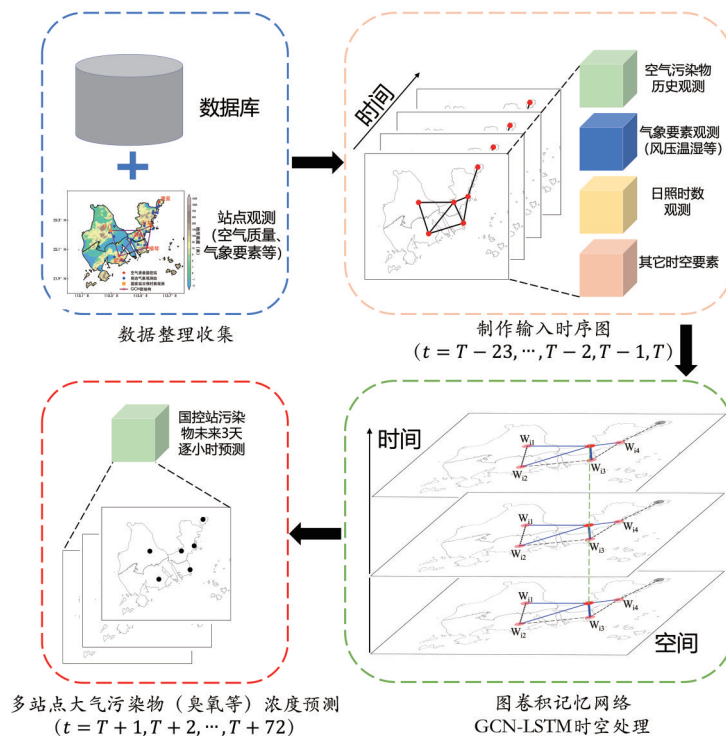


图2 图卷积记忆网络进行 $\text{O}_3$ 时空预测示意图

Fig. 2 Schematic diagram of  $\text{O}_3$  spatio-temporal prediction using graph convolutional memory network

**1.2.2 空间依赖性** GCN由Bruna et al.(2013)提出,并广泛应用到非欧结构数据的空间特征提取。在本文中,根据图结构的定义, $N$ 个节点(即珠海市6个环保国控站点,图1)每个时刻 $t$ 的输入特征可以转化为图信号的特征矩阵 $X_t \in \mathbb{R}^{N \times M}$ ,其中 $M$ 是每个节点的特征数。不同节点之间存在的空间关系可以表示为一张确定的无向图 $G=(V, E, W)$ ,其中 $V$ 代表图的节点, $E$ 代表节点之间连接的边, $W \in \mathbb{R}^{N \times N}$ 为邻接矩阵,代表节点间的定量空间关系( $w_{i,i}=0$ ),若任意2个节点之间的空间距离在设定阈值内(本文设置为25 km),则设置为1,若超出则为0,最终据此得到确定的一张关系图(图1)。1层的图卷积过程表达为

$$H^l = \sigma \left( \tilde{D}^{-\frac{1}{2}} \tilde{W} \tilde{D}^{-\frac{1}{2}} H^{l-1} \Theta^l \right),$$

其中 $H^{l-1}$ 和 $H^l$ 分别为输入和输出, $\tilde{W}$ 和 $\tilde{D}$ 分别为修改的邻接矩阵及其度矩阵, $\sigma$ 为非线性激活函数, $\Theta^l$ 是 $l$ 层的可学习参数。

**1.2.3 时间依赖性** LSTM是建立在RNN上的一种新型神经网络(Hochreiter et al., 1997),它通过引入单元状态 $c$ (cell state)和门(gate)的结构,改进RNN的记忆能力,有效解决其难以处理的长期依赖问题。原始RNN只利用状态 $h$ 用于进行短时间序列的记忆,LSTM在此基础上引进了新的状态 $c$ 来存储长期的历史信息。此外,LSTM通过输入门、遗忘门和输出门有效地优化和改进了每个网络单元的参数信息。LSTM基本记忆块包含1个或多个记忆单元和3个非线性门,分别为遗忘门 $f_t$ ,输入门 $i_t$ 和输出门 $o_t$ ,前两者控制 $t$ 时刻单元状态 $c_t$ 的内容更新,后者控制 $c_t$ 映射到最终输出值 $h_t$ 的内容,表述为

$$c_t = f_t \odot c_{t-1} + i_t \odot \sigma_c(W_c x_t + U_c h_{t-1} + b_c),$$

其中 $W_c$ 是输入向量 $x_t$ 在时刻 $t$ 的权重矩阵, $U_c$ 为分配给前一时刻输出值 $h_{t-1}$ 的权重矩阵, $b_c$ 则为偏差向量, $\sigma_c$ 为激活函数用于给模式添加非线性, $\odot$ 为哈达玛积。

### 1.3 实验

本文构建的GCN-LSTM模型基于Python-Keras框架,包含图卷积层、LSTM层和全链接层,

输出单元数分别为64、10和72。本文数据跨度为2019年1月到2023年9月,采用5折交叉验证的方式进行模型训练和预报。模型训练次数和每次利用样本数分别设置为5 000和为64,采用自适应算法Adam进行优化,损失函数和学习率分别设置为均方误差MSE和0.000 1。当模式性能在连续100次训练中没有进一步提升,则停止训练。

不同变量和 $O_3$ 的相关性存在明显差异(表1),本文控制实验GCN\_LSTM\_Ctrl的输入预报因子为 $O_3$ 、T2、 $NO_2$ 、SH、WS10。在控制实验基础上,依次剔除剩余预报因子中相关性最低的因子开展实验GCN\_LSTM\_1(预报因子为 $O_3$ 、T2、 $NO_2$ 、SH)、GCN\_LSTM\_2(预报因子为 $O_3$ 、T2、 $NO_2$ )、GCN\_LSTM\_3(预报因子为 $O_3$ 、T2)、和GCN\_LSTM\_4(预报因子为 $O_3$ )。为证明图卷积的附加价值,本文还利用LSTM模型去开展 $O_3$ 预报,该模型的预报方案和超参数设置与GCN-LSTM基本保持一致,但存在根本区别:GCN-LSTM模型的训练和预报是多点全局的(模型个数为1),而LSTM是单点局部的(模型个数为6),需要针对各个站点去单独开展模型构建、训练和预报。LSTM实验的预报因子与GCN-LSTM控制实验保持一致。

### 1.4 评价

本文使用均方根误差RMSE和相关系数 $R$ 对模型预报能力进行检验,计算公式为

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (O_i - P_i)^2},$$

$$R = \frac{\sum_{i=1}^N (P_i - \bar{P})(O_i - \bar{O})}{\sqrt{\sum_{i=1}^N (P_i - \bar{P})^2} \sqrt{\sum_{i=1}^N (O_i - \bar{O})^2}},$$

其中 $O$ 代表观测值, $P$ 代表预报值, $N$ 为样本总数, $\bar{O}$ 为观测值样本平均, $\bar{P}$ 为预报值样本平均。需要说明的是, $O_3$ 的日值定义为24 h内滑动8 h平均的最大值,记为 $O_{3\_8h}$ ,其空气质量分指数iAQI(进位取整)在0~50、51~100、101~150、151~200和>200分别对应 $O_3$ 的优、良、轻度、中度、重度5个等级。进一步利用TS评分对等级预报准

表1 各预报因子与 $\rho(O_3)$ 的相关系数<sup>1)</sup>

Table 1 Correlation coefficient between  $O_3$  and its predictors

预报因子	$NO_2$	$PM_{2.5}$	$PM_{10}$	T2	SH	WS10
相关系数	-0.42*	-0.015	0.016	0.56*	0.19*	0.14*

1) 样本为随机选择的1 000个样本;\*表示通过99.9%显著性检验。

确率进行检验, 公式为

$$TS = \frac{N_A}{N_A + N_B + N_C},$$

其中 $N_A$ 、 $N_B$ 和 $N_C$ 分别是预报正确、空报和漏报次数。

## 2 时空预报能力评估

观测显示,  $\rho(\text{O}_3\text{-8h})$ 及其iAQI在2019年达到近几年峰值后(年均值达到 $94.5 \mu\text{g}/\text{m}^3$ ), 在2020年出现了明显下降, 然后在2021—2022年呈现出上升趋势(图3)。一年之中,  $\rho(\text{O}_3\text{-8h})$ 在春季和秋季呈现出双峰, 峰值出现在4和10月, 月均值分别为 $92.1$ 和 $128.8 \mu\text{g}/\text{m}^3$ ; 而在夏季呈现低谷, 最低值出现在6月。总体而言, GCN\_LSTM准确还原了 $\text{O}_3\text{-8h}$ 的年变化和季节变化特征, 变化趋势与实况吻合程度较佳。但是, 其对 $\text{O}_3\text{-8h}$ 在冬季的峰值存在明显低估, 尤其在10月误差达到 $19.3 \mu\text{g}/\text{m}^3$ 。这点与其他研究结论一致(李婷苑等, 2021), 主要原因是由于统计预报模型以最小化条件均值为优化目标, 因此对极值的捕捉能力存在不足(Sun et al., 2023; Henzi et al., 2021), GCN\_LSTM模型也无法克服这一缺陷。从24、48和72 h不同预报时

效来看, 模型对于年变化和季节变化的预报结果较为接近, 这充分说明了模型具备良好的 $\text{O}_3$ 短期预报的能力。但是需要说明的是, 年和季节变化时间尺度较长, 还需探究模型对 $\text{O}_3$ 小尺度变化(即日变化和昼夜变化)的捕捉能力。

模型24 h预报的日值在春、夏、秋、冬的 $R^2$ 值分别为0.42、0.41、0.55和0.43, 各拟合线的斜率也在1附近, 这表明模型可以真实捕捉 $\text{O}_3\text{-8h}$ 在不同季节的日变化(图4)。相应的iAQI的 $R^2$ 稍有下降, 拟合线斜率也明显小于1, 这与iAQI与 $\rho(\text{O}_3\text{-8h})$ 在不同区间的差异对应关系有关。当预报时效达到48 h, 模型的模拟能力出现明显下滑, 尤其 $R^2$ 在夏季已不足0.16。当预报时效进一步达到72 h, 模式虽仍具备一定技巧(如相关性), 但与实况的拟合程度低。这是由于 $\rho(\text{O}_3\text{-8h})$ 的短期日变化与气象条件的演变密切相关。而统计模型仅依靠过去观测, 未考虑天气形势的短期变化, 其集成的记忆网络虽存在长期记忆, 可以部分考虑到历史情况并将其考虑到未来预报中, 但仍与实际情形存在明显差异。此外,  $\rho(\text{O}_3\text{-8h})$ 和气象条件在不同季节(甚至不同月份)存在明显差异。

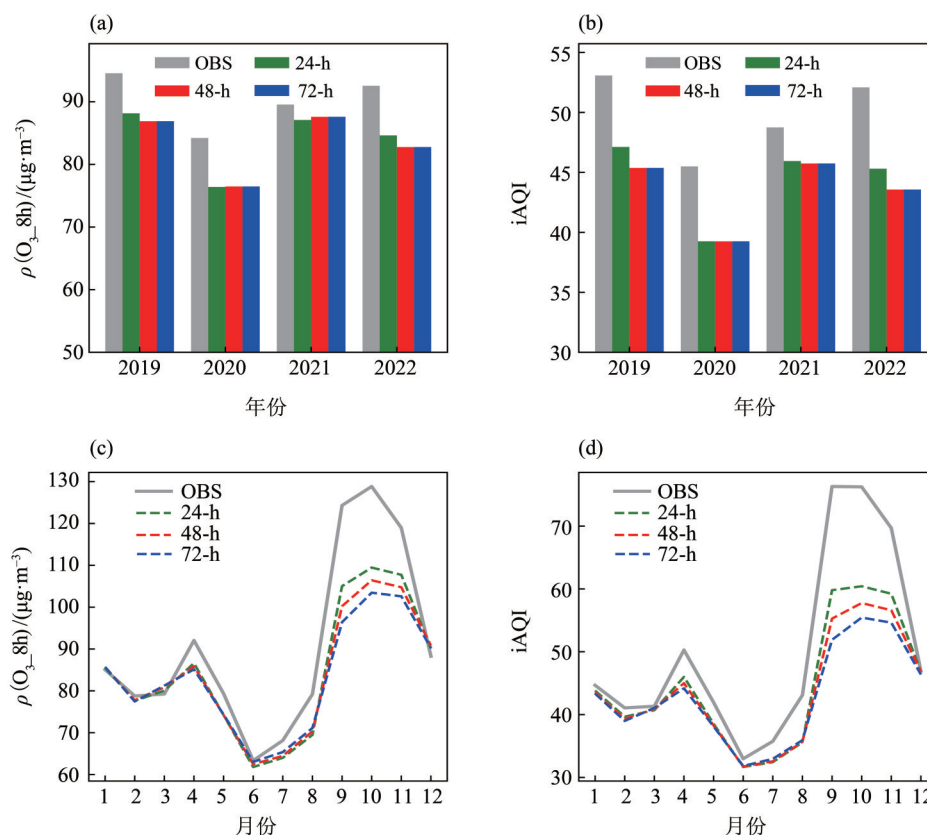
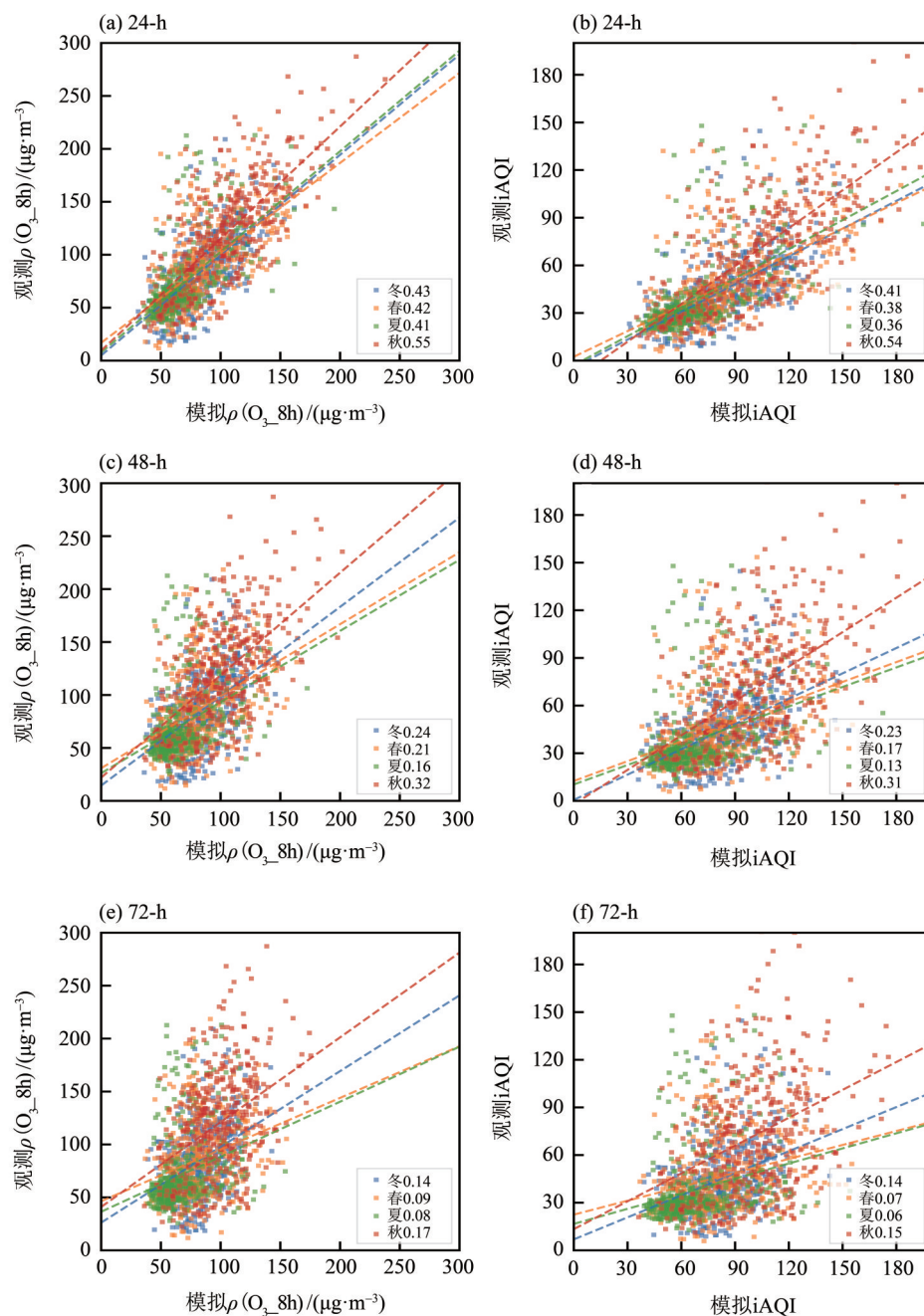


图3 GCN\_LSTM\_Ctrl不同预报时效站点平均的 $\rho(\text{O}_3\text{-8h})$ 及其iAQI的年变化和月变化与观测对比

Fig. 3 Comparison of annual and monthly variations of site-averaged  $\text{O}_3\text{-8h}$  and iAQI between GCN\_LSTM\_Ctrl and observations at different lead times



图中虚线和数字分别为拟合线和 $R^2$ 。

图4 GCN\_LSTM\_Ctrl不同预报时效的站点平均的 $\rho(\text{O}_3, 8\text{h})$ 、iAQI与观测的对比散点图

Fig. 4 Scatter plots of site-averaged daily  $\text{O}_3, 8\text{h}$  and iAQI between GCN\_LSTM\_Ctrl and observations at different lead times, where the dotted line and numbers are the fitting line and  $R^2$ , respectively

本文由于可用数据长度有限,未根据不同季节去分别构建模型,也从一定程度上制约了模型性能。

光照条件不仅是 $\text{O}_3$ 生成的必要条件,还通过进一步影响大气边界层湍流活动对其扩散产生重要影响,因此 $\rho(\text{O}_3)$ 呈现出显著昼夜变化。观测表明各季节 $\rho(\text{O}_3)$ 的昼夜变化的相位差异不大,均在07:00~08:00处于最小值,然后迅速增加,在午后

13:00~15:00达到最大值(图5)。但是,不同季节的昼夜变化振幅存在明显差异,秋季振幅最大,紧接着是冬季,而夏季则最小,这主要与秋冬季节较少的云量和频密的北风有关。GCN\_LSTM对于 $\text{O}_3$ 昼夜变化的模拟非常有效,其预报的昼夜变化的相位和幅度与实况基本吻合(除了在春季夜间和秋季日间存在低估),并且整体表现随着预报时效

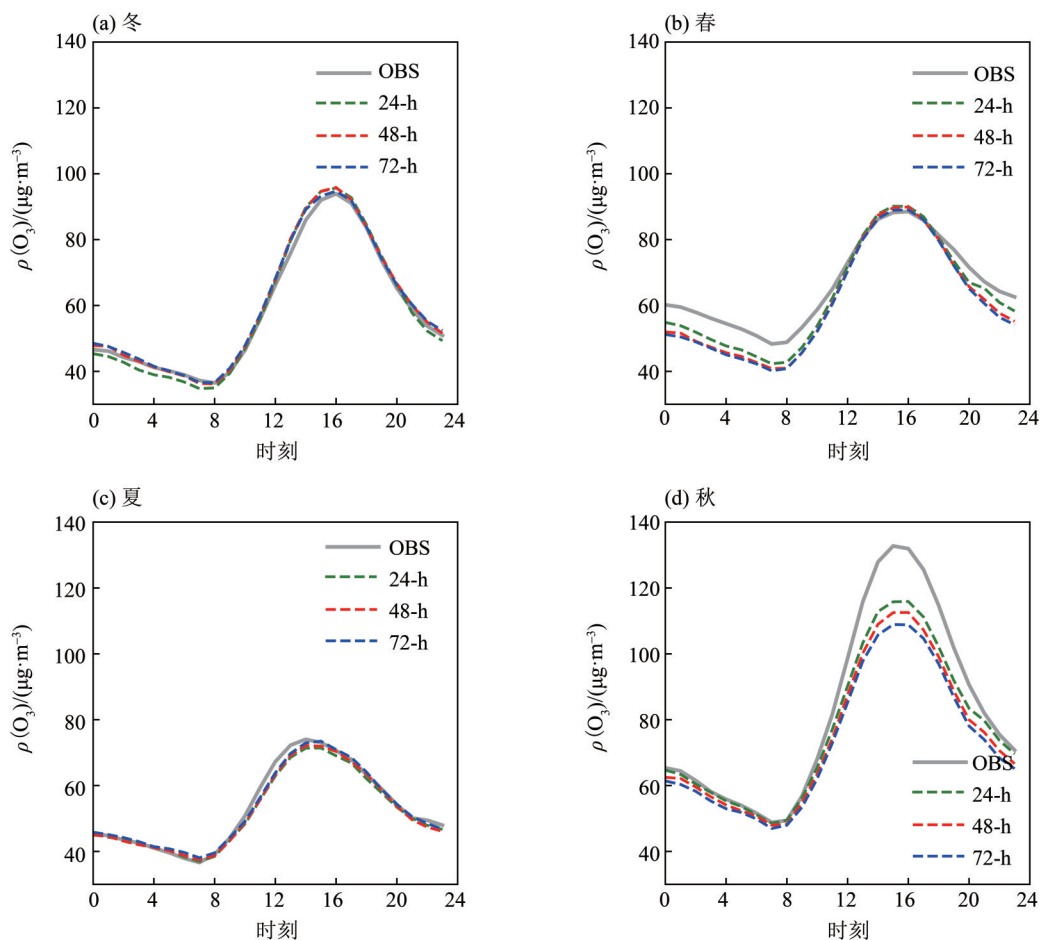


图5 GCN\_LSTM\_Ctrl不同预报时效的站点平均 $\rho(\text{O}_3)$ 在不同季节的昼夜变化与观测对比

Fig. 5 Comparison of diurnal variations of site-averaged  $\text{O}_3$  concentration between GCN\_LSTM\_Ctrl and observations at different lead times

的增加降低十分有限。模型在春季夜间和秋季日间的偏差主要是由于模型未针对不同季节进行训练, 预报结果趋于整体平均最优。臭氧一年之中在春季和秋季呈现双峰(图3c), 而模型整体对极值的捕捉能力存在不足, 因此低估了春季昼夜变化的谷值和秋季昼夜变化的峰值。

其较优的表现归功于GCN\_LSTM自身可以有效解析输出变量的时间变化信息, 还与在构建训练数据时将滑动窗口(长度为 $24+72=96$ )步长设置有关。本文将步长设置为24, 对应为昼夜变化的时间尺度, 而测试实验中将步长调整为1, 虽可以增加训练数据的数量, 但明显降低了模型对昼夜变化的捕捉能力, 从而引起模型整体性能的下落。

$\text{O}_3$ 不仅存在不同时间尺度的变化, 还具有明显的空间变化(Lu et al., 2019)。不同站点之间(表2), 唐家站的 $\rho(\text{O}_3_{8\text{h}})$ 均值最大, 达到 $95.9 \mu\text{g}/\text{m}^3$ , 其次是横琴站( $94.5 \mu\text{g}/\text{m}^3$ )和吉大( $92.7 \mu\text{g}/\text{m}^3$ ), 前山

最低( $86.0 \mu\text{g}/\text{m}^3$ )。GCN\_LSTM基于站点构建的时序图开展全局训练, 可以较为合理地还原 $\rho(\text{O}_3_{8\text{h}})$ 的空间差异, 各站点的 $R$ 都达到0.65以上, 但存在明显低估(各站点的RMSE均高于 $30 \mu\text{g}/\text{m}^3$ )。尤其是唐家站, RMSE达到了 $38.0 \mu\text{g}/\text{m}^3$ , 这主要与上述提到的统计模型自身限制有关。需要说明的是, 模式在不同站点的表现存在一些差异, 这主要与各站点气象因子对 $\text{O}_3$ 预报不同的重要性有关。以唐家站为例, 该站点周围山体环绕且存在大片树林, 地形对臭氧的扩散产生影响, 同时植物VOCs的排放对于臭氧本地生成起到了重要作用。本文构建的预报模型仅考虑了气象因素的影响, 因此模型在该站的整体预报水平最低。此外, 唐家站位于珠海东北角(图1), 本文关注珠海, 因此在构建图结构时未考虑上游站点对其的影响。后续应用若将研究区域扩展到珠三角, 应可以部分提高该站的预报能力。

表2 日 $\rho(\text{O}_3_{8\text{h}})$ 在各站点的实况均值以及GCN\_LSTM\_Ctrl预报均值、RMSE和 $R$ Table 2 The observed means of daily  $\text{O}_3_{8\text{h}}$  from different sites and predicted means, RMSE, and  $R$  from GCN\_LSTM\_Ctrl experiment

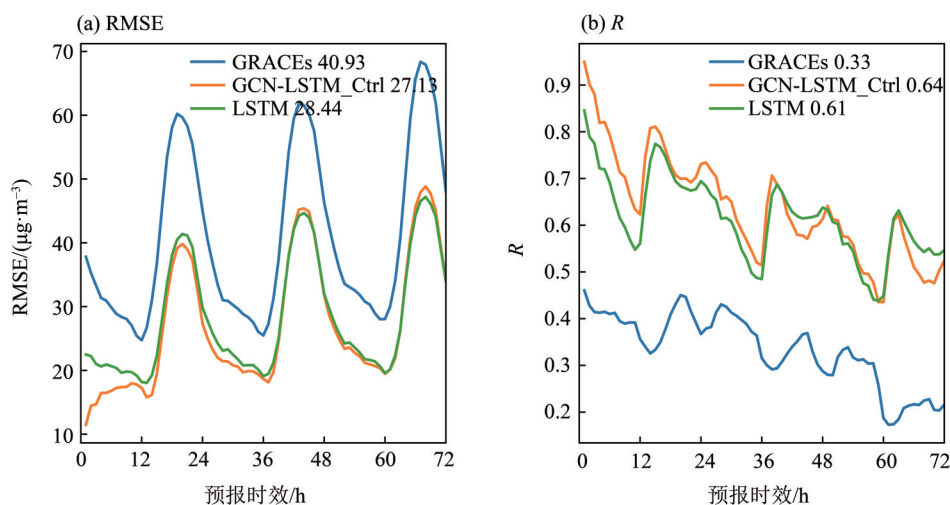
统计量	唐家	前山	吉大	斗门	横琴	金湾
实况均值/ $(\mu\text{g}\cdot\text{m}^{-3})$	95.9	86.0	92.7	87.7	94.5	89.0
预报均值/ $(\mu\text{g}\cdot\text{m}^{-3})$	85.2	80.2	86.2	81.9	87.3	84.0
RMSE/ $(\mu\text{g}\cdot\text{m}^{-3})$	38.0	30.2	33.8	32.7	31.9	30.8
$R$	0.65	0.71	0.69	0.70	0.69	0.71

### 3 与其他模型对比

以上结果已表明, GCN\_LSTM具备良好的 $\text{O}_3$ 时空模拟能力。为充分阐释其有效性, 本文进一步将其与业务数值模式GRACEs和广泛应用的LSTM模型进行对比。总体而言, 各方法的预报能力随着预报时效的增加而逐渐下降, 并受到 $\text{O}_3$ 自身明显昼夜变化的影响(图6)。以往研究揭示出GRACEs相较于其他模式, 对于华南地区的空气污染物的模拟具有优势(李婷苑等, 2021)。尽管如此, 其预报72 h内的RMSE=40.93  $\mu\text{g}/\text{m}^3$ ,  $R$ 均值为0.33。这与两个统计模型的表现存在较大差距, 这与数值模式在排放清单、气象条件等方面的较大不确定性有关。需要说明的是, 数值模式结果虽未经过后处理订正, 但研究表明后订正技术对模式的提升效果有限, 无法弥补其与统计模型之间的明显差距(孙弦等, 2023)。两个统计模型之间的表现差距不大, 均将RMSE降低至27~28  $\mu\text{g}/\text{m}^3$ , 将 $R$ 提高至0.62附近。其中, GCN\_LSTM表现更优, 优势集中在初始时刻, 因此其RMSE( $R$ )均值

更低(高)。这是由于GCN\_LSTM利用图卷积过程, 有效地解析了不同站点空间上的相互影响, 因此提高了 $\text{O}_3$ 整体的预报能力。除了性能上的有效提升, GCN\_LSTM通过1个模型全局考虑所有站点, 完成5折交叉验证的时间为822.1 s, 计算速度相较于LSTM提高了71%。

图7对比了实况和3种方法在不同站点预报的 $\text{O}_3_{8\text{h}}$ 日值。与上述结论相一致, GRACEs虽可以较为合理地还原 $\text{O}_3$ 的日变化特征, 但性能较统计模型存在差距。尤其针对2019年9~11月期间频发的 $\text{O}_3$ 污染事件, GRACEs几乎模拟失败, 这主要与内部物理化学过程和外部排放清单的不确定性有关(邓雪娇等, 2016)。另一方面, GCN\_LSTM和LSTM显著提升了各站点 $\text{O}_3$ 预报技术, 且对于上述提到的污染事件, 预报能力得到了明显提升, 但仍存在一些低估, 这是由于统计预报模型通常以最小化预报量的条件均值为训练优化目标, 制约了对极值的捕捉能力, 所以表现出对极值的低估(Sun et al., 2023)。其中GCN\_LSTM对于极值的



图中各模型后数字表示均值。

图6 GRACEs、GCN\_LSTM\_Ctrl和LSTM预报的站点平均(4个站)的 $\rho(\text{O}_3)$ 的RMSE和 $R$ 随预报时效的变化Fig. 6 The RMSE and  $R$  of the site-averaged (4 sites)  $\text{O}_3$  concentration predicted by GRACEs, GCN\_LSTM\_Ctrl and LSTM as a function of lead time, with the numbers denoting the means

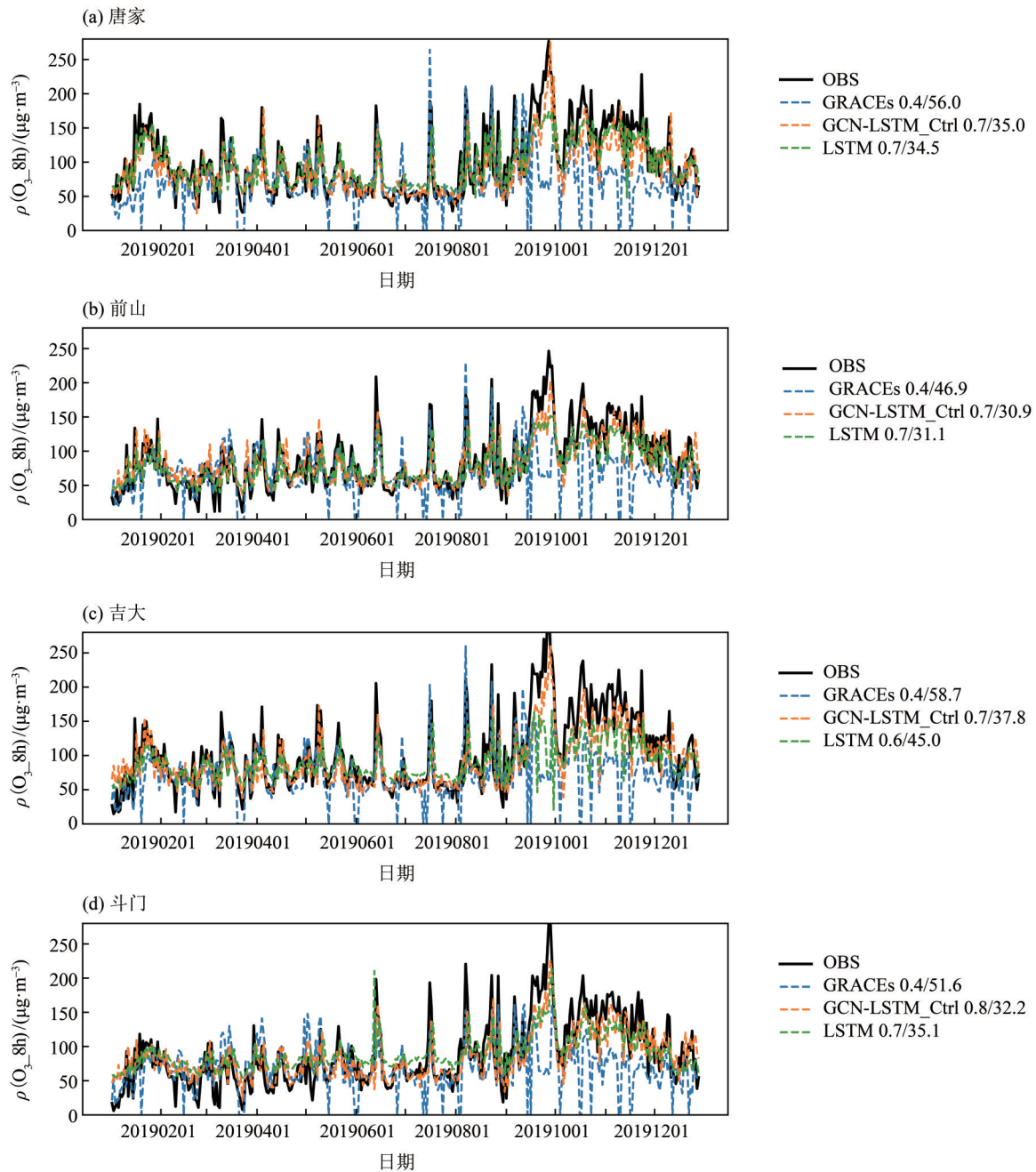


图7 GRACEs、GCN\_LSTM\_Ctrl和LSTM预报的不同站点 $\rho(O_3_{8h})$ 日值的时间序列与观测的对比

Fig. 7 Time series of  $O_3_{8h}$  predicted by GRACEs, GCN\_LSTM\_Ctrl and LSTM against observations at different sites, with the numbers denoting the  $R$  and  $RMSE$

捕捉能力更加优秀。以唐家站为例, 2019年9月下旬(9月21~30日)臭氧污染事件在全年最为严重, 该时段 $O_3_{8h}$ 日均值高达 $219.5 \mu\text{g}/\text{m}^3$ 。GRACEs未能捕捉这一污染事件, 预报均值仅为 $79.2 \mu\text{g}/\text{m}^3$ 。LSTM模型取得明显改进(均值为 $158.9 \mu\text{g}/\text{m}^3$ ), 但仍存在明显低估。GCN\_LSTM表现最佳, 预报结果(均值为 $187.4 \mu\text{g}/\text{m}^3$ )与实况最为接近。此外,

LSTM在各站点的表现存在差异, 其虽可以在唐家站表现出与GCN\_LSTM相近的模拟能力, 但在其他站点则明显不如。尤其对于吉大站, LSTM( $RMSE=45.0 \mu\text{g}/\text{m}^3$ ;  $R=0.6$ )与GCN\_LSTM( $RMSE=37.8 \mu\text{g}/\text{m}^3$ ;  $R=0.7$ )存在明显差距。

此外, 业务预报和政府决策通常更关注污染物浓度对应的空气质量等级。图8对比3种方法在

不同时效对不同空气质量等级的TS评分。对于24 h预报, 优等级因为样本数最多, 因此各方法均表现出不错的预报能力。随着等级的提升, 各模式预报能力快速下降, LSTM对于轻度污染的预报能力甚至已低于GRACEs。GCN-LSTM在各等级

表现均最佳, 尤其对于中度污染的捕捉能力存在明显优势, 但是在重度等级也同样没有表现出技巧, 这与统计模型对于捕捉极值的内在缺陷和重度污染事件样本太少有关。随着预报时效的增加, GCN-LSTM的表现也出现明显下降, 三者72 h的等级预报技术趋于一致。

### 4 预报因子的影响

以上的结果充分说明了GCN-LSTM控制实验对O<sub>3</sub>时空模拟的不错能力。在控制实验中, 所有存在相关性的变量(表1)都被使用作为预报因子。不少研究已表明预报因子的选择和组合会对模型的预报能力产生影响(夏冬等, 2012; Gao et al., 2018; 张灿等, 2017)。为阐释其对O<sub>3</sub>预报的影响, 图9给出了不同GCN-LSTM实验模拟的 $\rho(O_3)$ 的RMSE和R对预报时效的变化。相较于控制实验, GCN-LSTM1的预报因子中移除了WS10, 其技巧竟略微有所提升, RMSE降低至27.57  $\mu\text{g}/\text{m}^3$ , R提升至0.56。但是, 进一步从预报因子中先后移除了SH(即GCN-LSTM2)和NO<sub>2</sub>(即GCN-LSTM3), 模式的性能出现连续下降。相较于控制实验, RMSE均值分别增加了0.29和0.9  $\mu\text{g}/\text{m}^3$ , R均值分别降低了0.1和0.3。最后, GCN-LSTM4实验进一步移除了T2(即仅使用O<sub>3</sub>自身作为预报因子), 模型的预报能力发生了显著下降, RMSE增加至30.98  $\mu\text{g}/\text{m}^3$ , R也降低至0.47。以上结果表明, 在O<sub>3</sub>预报中考虑相关性较高的变量(尤其是T2)可以明显改进模型的预报能力, 但是低相关性因子的

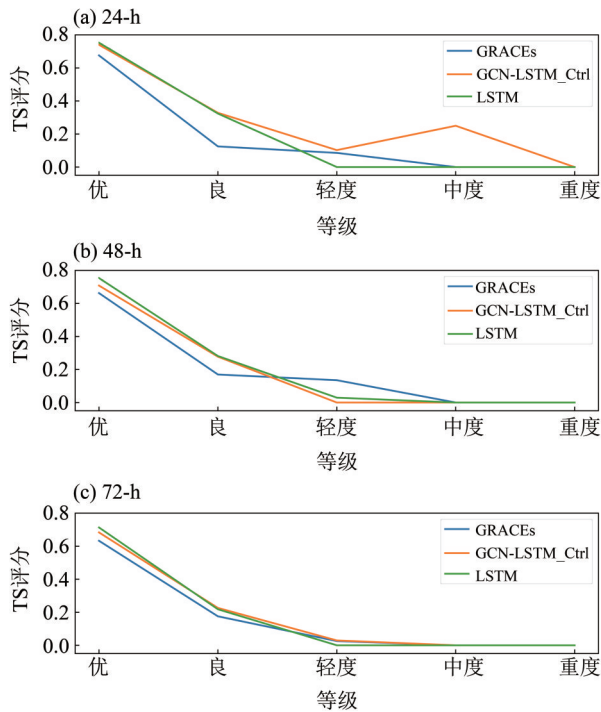


图8 GRACEs、GCN\_LSTM\_Ctrl和LSTM不同预报时效的O<sub>3</sub> 8h不同空气质量等级的TS评分  
 Fig. 8 TS scores of different air quality levels of O<sub>3</sub> 8h predicted by GRACEs, GCN\_LSTM\_Ctrl and LSTM at different lead times

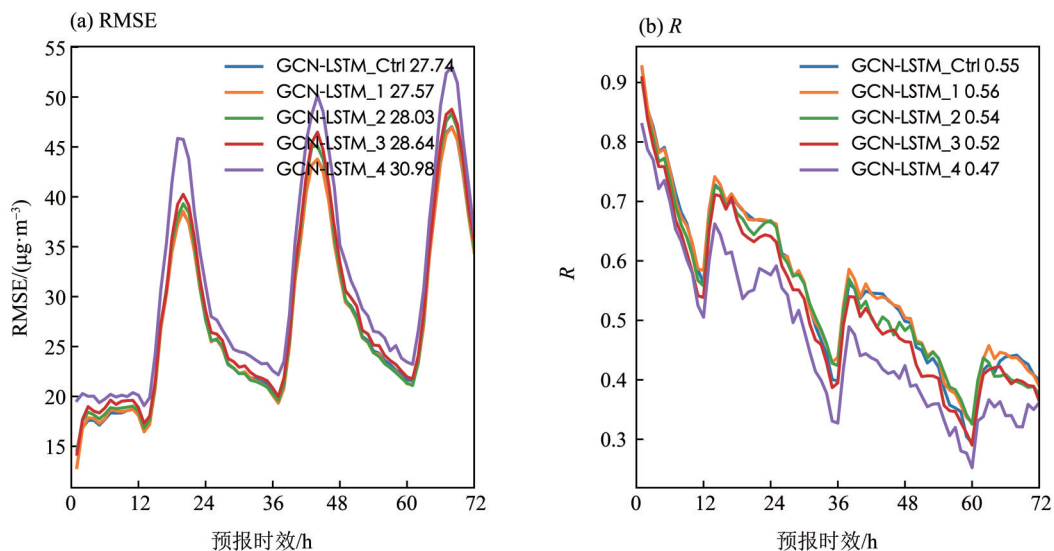


图9 各GCN-LSTM实验预报的站点平均(6个站) $\rho(O_3)$ 的RMSE和R随预报时效的变化  
 Fig. 9 The RMSE and R of the site-averaged O<sub>3</sub> concentration predicted by GCN-LSTM of varying configurations as a function of lead time, with the numbers denoting the means

引入反而会产生不利影响。

## 5 结 论

本研究基于 2019 年 1 月至 2023 年 9 月珠海市环境和气象站点监测数据,通过构建时空协同的图卷积记忆网络 GCN-LSTM 开展珠海市 6 个环保国控站未来 3 d 逐小时 O<sub>3</sub> 预报,得到如下主要结论:

1) GCN\_LSTM 在不同预报时效的预报结果均可以准确地还原 O<sub>3</sub> 的年变化趋势和季节变化特征,但对冬季的峰值存在明显低估。对于日变化,模型 24 h 的预报结果在不同季节的表现都较佳,但随着预报时效的增加,模型性能出现明显下降。对于昼夜变化,各预报时效的结果与实况十分吻合。此外,模型可以较为真实地反映 O<sub>3</sub> 空间分布上的差异。

2) 72 h 预报时效内,GCN-LSTM 和 LSTM 模型以及数值模式 GRACEs 的性能在超过 24 h 之后均下降明显。其中,GCN-LSTM 表现最优,其 RMSE 和 *R* 均值分别为 27.13 μg/m<sup>3</sup> 和 0.64, LSTM 表现稍次之(RMSE=28.44 μg/m<sup>3</sup>; *R*=0.61),而 GRACEs 的表现与统计模型存在明显差距(RMSE=

40.93 μg/m<sup>3</sup>; *R*=0.33)。

3) GCN-LSTM 通过一个模型全局考虑所有站点及其之间的相互联系,相较于对各站点单独构建 LSTM 模型,不仅将计算速度提高了 71%,而且在不同站点的表现也更为优秀和稳定,对于秋季 O<sub>3</sub> 污染事件的捕捉能力也有所提高。此外,敏感性实验揭示出考虑相关性较高的变量(如近地气温)作为预报因子可以有效改进模型的预报能力。

4) GCN-LSTM 模型仍存在 2 个明显局限性。第一是模型的预报能力下降过快,可能与模型未考虑任何未来预报信息密切相关,后续可以考虑在模型中引入未来天气预测信息,从而克服这一问题,但同时也会增加模型运行对数据的依赖性。同时,O<sub>3</sub> 自身以及影响它的气象条件存在明显季节变化。随着监测数据的持续积累,可以针对不同季节甚至不同月份单独构建模型,从而部分缓解这一问题。第二是模型对极值的捕捉能力存在缺陷,这点可以通过上述方法得到一些改善,但受限于统计模型自身的性质,有效解决该问题还需采用其他方法(例如基于历史的相似预报)。

## 参考文献:

- 陈德辉,薛纪善,沈学顺,等,2012.我国自主研发的全球/区域一体化数值天气预报系统 GRAPES 的应用与展望[J].中国工程科学,14(9):46-54.
- 邓雪娇,邓涛,麦博儒,等,2016.华南区域大气成分业务数值预报 GRACEs 模式系统[J].热带气象学报,32(6):900-907.
- 赖安琪,陈晓阳,刘一鸣,等,2018.珠江三角洲高质量浓度 PM<sub>2.5</sub> 和 O<sub>3</sub> 复合污染特征[J].中山大学学报(自然科学版),57(4):30-36.
- 李婷苑,吴乃庚,邓雪娇,等,2021.华南区域大气成分数值模式 GRACEs 预报性能评估[J].热带气象学报,37(2):207-217.
- 李勋,秦墨梅,李琳,等,2020.基于空气质量模拟的江苏省大气污染物排放清单比较研究[J].南京信息工程大学学报(自然科学版),12(6):705-713.
- 祁柏林,郭昆鹏,杨彬,等,2021.基于 GCN-LSTM 的空气质量预测[J].计算机系统应用,30(3):208-213.
- 孙弦,孙磊,聂会文,等,2023.两种空气质量数值模式的应用评估与集合改进研究[J].热带气象学报,39(3):361-373.
- 王自发,吴其重,GBAGUIDIA,等,2009.北京空气质量多模式集成预报系统的建立及初步应用[J].南京信息工程大学学报(自然科学版),1(1):19-26.
- 吴乃庚,邓玉娇,向昆仑,等,2020.2019年广东省生态气象监测公报[Z].广州:广东省气象局.
- 夏冬,谭浩波,陈玲,等,2012.用逐步回归法预报地面臭氧浓度[J].广东气象,34(1):47-50.
- 杨雨佳,肖庆来,陈健,等,2023.融合空间和统计特征的 CNN-GRU 臭氧浓度预测模型研究[J].南京大学学报(自然科学版),59(2):322-332.
- 张灿,蒋昌潭,罗财红,等,2017.气象因子对臭氧的影响及其在空气质量预报中的应用[J].中国环境监测,33(4):221-228.
- 张远航,郑君瑜,陈长虹,等,2020.中国大气臭氧污染防治蓝皮书(2020年)[C].北京:中国环境科学学会臭氧污染控制专业委员会.
- 赵燕,李大伟,翟宇虹,等,2022.2014—2021年珠海市环境空气质量变化趋势及污染特征研究[J].环境科学与管理,47(12):144-149.
- 周广强,谢英,吴剑斌,等,2016.基于 WRF-Chem 模式的华东区域 PM<sub>2.5</sub> 预报及偏差原因[J].中国环境科学,8(9):2251-2259.

- BASSETT R, YOUNG P J, BLAIR G S, et al, 2020. A large ensemble approach to quantifying internal model variability within the WRF numerical model[J]. *J Geophys Res Atmos*, 125(7): e2019JD031286.
- BRUNA J, ZAREMBA W, SZLAM A, et al, 2013. Spectral networks and locally connected networks on graphs [EB/OL]. arXiv: 1312.6203.
- CABANEROS S M, CALAUTIT J K, HUGHES B R, 2019. A review of artificial neural network models for ambient air pollution prediction [J]. *Environ Model Softw*, 119(C): 285–304.
- EDER B, KANG D, MATHUR R, et al, 2006. An operational evaluation of the Eta-CMAQ air quality forecast model[J]. *Atmos Environ*, 40(26): 4894–4905.
- GAO M, YIN L, NING J, 2018. Artificial neural network model for ozone concentration estimation and Monte Carlo analysis[J]. *Atmos Environ*, 184: 129–139.
- HENZI A, ZIEGEL J F, GNEITING T, 2021. Isotonic distributional regression[J]. *J R Stat Soc Ser B Stat Methodol*, 83(5): 963–993.
- HOCHREITER S, SCHMIDHUBER J, 1997. Long short-term memory[J]. *Neural Comput*, 9(8): 1735–1780.
- KITAYAMA K, MORINO Y, YAMAJI K, et al, 2019. Uncertainties in O<sub>3</sub> concentrations simulated by CMAQ over Japan using four chemical mechanisms [J]. *Atmos Environ*, 198: 448–462.
- KRZYZANOWSKI M, COHEN A, 2008. Update of WHO air quality guidelines[J]. *Air Qual Atmos Health*, 1(1): 7–13.
- LI M, LIU H, GENG G, et al, 2017. Anthropogenic emission inventories in China: A review [J]. *Natl Sci Rev*, 4(6): 834–866.
- LU H, LYU X, CHENG H, et al, 2019. Overview on the spatial-temporal characteristics of the ozone formation regime in China [J]. *Environ Sci Processes Impacts*, 21(6): 916–929.
- PAK U, KIM C, RYU U, et al, 2018. A hybrid model based on convolutional neural networks and long short-term memory for ozone concentration prediction [J]. *Air Qual Atmos Health*, 11(8): 883–895.
- QI Y, LI Q, KARIMIAN H, et al, 2019. A hybrid model for spatiotemporal forecasting of PM<sub>2.5</sub> based on graph convolutional neural network and long short-term memory [J]. *Sci Total Environ*, 664: 1–10.
- SUN L, LAN Y, JIANG R, 2023. Using CNN framework to improve multi-GCM ensemble predictions of monthly precipitation at local areas: An application over China and comparison with other methods [J]. *J Hydrol*, 623: 129866.
- XIAO X, JIN Z, WANG S, et al, 2022. A dual-path dynamic directed graph convolutional network for air quality prediction [J]. *Sci Total Environ*, 827: 154298.
- YU B, YIN H, ZHU Z, 2017. Spatio-temporal graph convolutional networks: A deep learning framework for traffic forecasting [EB/OL]. arXiv: 1709.04875.
- ZHANG J, DING W, 2017. Prediction of air pollutants concentration based on an extreme learning machine: The case of Hong Kong [J]. *Int J Environ Res Public Health*, 14(2): 114.
- ZHANG Y, BOCQUET M, MALLET V, et al, 2012. Real-time air quality forecasting, part I: History, techniques, and current status [J]. *Atmos Environ*, 60: 632–655.

(责任编辑 秦社彩)