

基于多尺度序列间相关性的臭氧浓度预测

朱振业¹ 唐超礼²

ZHU Zhenye TANG Chaoli

摘要

为提高地表臭氧浓度逐小时预测的精度,文章提出了一种基于多尺度序列相关性的深度学习模型。该模型通过频域分析、图卷积网络和时间卷积网络,提取并学习时间序列中的多尺度时间信息。在真实数据集上的实验结果表明,该模型在多项指标上表现出更高的预测精度,尤其在极值拟合方面具有显著优势,可为城市空气质量管理 and 政策制定提供有效的数据支持。

关键词

臭氧浓度预测;深度学习;频域分析;图卷积;时间卷积

doi: 10.3969/j.issn.1672-9528.2025.04.031

0 引言

随着社会经济和城市化的快速发展,空气污染逐渐从微粒污染转向光化学污染^[1-2],其中地表臭氧污染日益突出,严重制约了中国空气质量的改善。与其他污染物不同,地表臭氧是二次污染物,由氮氧化物和挥发性有机化合物在阳光下通过复杂的光化学反应生成^[3-5]。这些臭氧前体物质主要来自机动车尾气、发电厂和炼油厂的排放,因此城市中更容易发生臭氧污染,对居民健康、生态环境和农业生产带来显著影响。研究表明,臭氧暴露会增加呼吸系统和心血管疾病的风险,尤其对老年人和儿童影响较大,同时还会降低农作物产量并损害建筑材料。因此,准确预测地表臭氧浓度及其变化趋势,对于保护人类健康和减少污染损害至关重要。

臭氧的预测方法主要分为两类:基于知识驱动的方法和基于数据驱动的方法。前者如 WRF-Chem 模型^[6],基于物理化学过程,适用于简单预测,但由于臭氧影响因素复杂且模型计算时间较长,准确预测具有挑战性。后者使用统计方法和机器学习,从大量历史数据中提取有价值的信息,广泛应用于臭氧预测。例如,王一龙等人^[7]结合 ARMA 和 GARCH 模型预测烟台市臭氧浓度,能较好地描述时间序列中的自相关性和波动性,但缺乏对非线性因素、季节变化和突发事件的处理。杨晓彤等人^[8]提出基于随机森林的预测方法,能有效处理复杂的非线性关系并减少异常值影响,但在捕捉臭氧极值时表现不佳,并且需要大量训练样本。王凯等人^[9]结合

GCN 和 LSTM 模型,分别捕捉空间和时间变化特征,达到了较高的预测精度,但在远距离空间依赖建模和长时间序列处理上有所不足。

针对现有模型在长时间序列处理上的不足,本文提出了一种基于离散小波变换(DWT)、自适应图卷积(AGCN)和时间卷积(TCN)的组合模型,用于逐小时预测地表臭氧。该模型通过 DWT 对序列进行时频域分析,得到不同尺度的信号,并根据每个尺度选择合适的卷积核进行 AGCN 和 TCN 操作,从而同时捕捉序列间和序列内的相关性,提升了臭氧浓度预测的效果。

1 数据采集与模型构建

1.1 数据采集

本文使用的空气质量和气象数据时间分辨率为小时,跨度为2021年1月1日—2023年12月31日,共26 280条数据。空气质量数据来自中国环境监测总站的全国城市空气质量实时发布平台,选取了北京12个站点的臭氧和二氧化氮数据。气象数据来源于欧洲中期天气预报中心(ECMWF)的ERA5大气再分析数据集,包括露点温度、空气温度和地表气压(高度为2 m)。

1.2 模型框架

本文结合 DWT、ADGCN 和 TCN,构建了如图1所示的模型。在模型的每个块中,先利用 DWT 提取序列间的尺度,再对序列进行自适应图卷积,然后在前两个模块的基础上进行时间卷积。最后通过残差连接的方式输出预测结果。

1.3 DWT

DWT 是一种用于分析非平稳信号的有效工具,特别适用于处理多尺度特征^[10],将 DWT 用在这里处理不同时间尺度下的序列间相关性。

1. 安徽理工大学计算机科学与工程学院 安徽淮南 232001

2. 安徽理工大学电气与信息工程学院 安徽淮南 232001

[基金项目] 安徽省研究生创新创业实践项目(2023cxycsj089);

安徽理工大学研究生创新基金(2023cx2092)

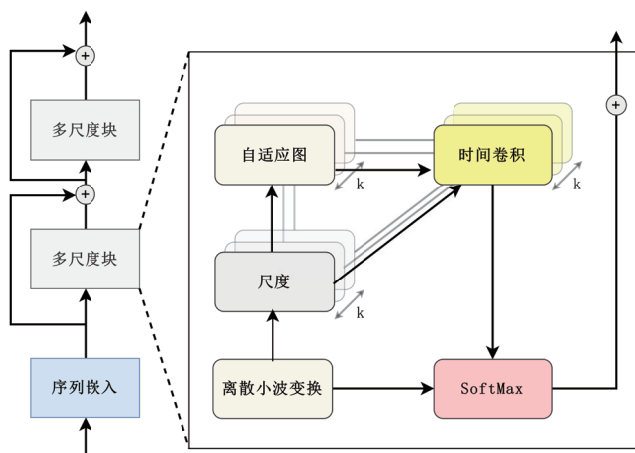


图 1 模型结构

DWT 通过一系列低通和高通滤波器对信号进行分解，获取信号的逼近（低频）和细节（高频）特征。其多尺度分解特性使得 DWT 能够有效捕捉信号中的局部细节和全局趋势。本文聚焦利用小波变化提取的逼近系数来选取前 k 个周期作为主要的周期。

1.4 AGCN

AGCN 是 GCN 的扩展，通过动态调整图结构中的节点和边的权重，使得模型能够根据输入数据的变化自适应地学习图的拓扑结构，增强模型的表达能力^[11]。与传统的图卷积不同，自适应图卷积不仅依赖于预定义的邻接矩阵，还能根据数据的特征，捕捉动态变化的节点间关系。

邻接矩阵 A' 并非固定，而是通过节点特征 X 动态学习得到的：

$$A' = f(X) \quad (1)$$

式中： $f(X)$ 是一个可学习的函数，根据节点特征动态生成邻接矩阵 A' 。在每一次图卷积的操作中，自适应邻接矩阵 A' 结合节点特征 H 进行卷积操作：

$$H_i^{(l+1)} = \sigma(\tilde{A}' H_i^{(l)} W_i^{(l)}) \quad (2)$$

式中： i 表示第 i 个尺度； $\tilde{A}'$ 表示动态生成邻居矩阵 A' 的归一化形式； $H_i^{(l)}$ 表示第 l 层的节点特征； $W_i^{(l)}$ 表示可学习的权重矩阵； σ 表示激活函数。

1.5 TCN

TCN 是一种专门用于处理时间序列数据的卷积网络^[12]。通过因果卷积保证时间步的顺序性，同时通过扩张卷积扩大感受野，从而捕捉长期依赖。在本文模型中，采用 TCN 处理 AGCN 输出的数据。为适应不同尺度，根据 DWT 输出的尺度大小，动态调整卷积核的大小，使得 TCN 能够在多种尺度下进行有效建模。具体公式为：

$$H_i^{\text{TCN}}(t) = \sum_{s=0}^{k-1} W_i(s) \cdot H_i^{\text{AGCN}}(t - r \cdot s) \quad (3)$$

式中： i 是第 i 个尺度； H_i^{AGCN} 是自适应图卷积层的输出特征； $W_i(s)$ 是卷积核的权重； r 是扩张率，用于控制感受野的大小； k 是由 DWT 提取的时间尺度决定的卷积核大小。

2 实验与结果分析

2.1 基准

本文选择 3 种时间序列预测方法进行比较，包括基于 Transformer 架构的 Informer^[13] 模型和 Autoformer^[14] 模型，以及线性模型 Dliner^[15]。

2.2 实验设置

本文实验在 PyCharm 上进行，使用 Windows 11 操作系统、NVIDIA GeForce RTX3060GPU、Python 版本为 3.7.16。历史序列长度为 24，预测长度为 1。这些设置应用于所有模型。初始学习率为 0.000 1，使用批大小为 32，均方误差（MSE）作为损失函数来优化模型。epoch 数为 10，适用时采用提前终止。对于模型中 FFT 的结果，选择其中最显著的 3 个尺度。另外，数据集按照 0.7:0.1:0.2 划分，分别是训练集、验证集和测试集，使用 24 个历史数据预测未来 12 个时刻数据。

2.3 预测效果评估

使用回归模型常用的几个评价指标来评估模型的效果，分别是均方误差（MSE）、均方根误差（RMSE）、平均绝对误差（MAE）和决定系数（ R^2 ），其定义公式为：

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (o_i - \hat{o}_i)^2 \quad (4)$$

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (o_i - \hat{o}_i)^2} \quad (5)$$

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |o_i - \hat{o}_i| \quad (6)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (o_i - \hat{o}_i)^2}{\sum_{i=1}^n (o_i - \bar{o})^2} \quad (7)$$

式中： n 是样本数量； O_i 是真实值； $\hat{o}_i$ 是预测值； $\bar{o}$ 是真实值的均值。

2.4 各模型预测结果对比

表 1 展示了本文模型和各基准模型的预测结果的指标，从中可以看出，模型在其中达到了最高的预测精度，决定系数到达了 0.97，其次是 Dliner、Autoformer 和 Informer，决定系数分别是 0.93、0.85 和 0.81。为了更直观地展示模型的优势，随机截选了测试集中的一段结果，如图 2 所示。比较图 2 中 4 个模型的预测值和真实值曲线，可以看出，模型在极值方面拟合的效果最佳，Informer 效果最差。此外，图 3 展示了各模型散点图，相较于基准模型，本文模型的散点密集分布在真实值的周围。图 4 为本文模型的预测值与真实值拟合曲线。

表 1 不同模型预测臭氧浓度的评价指标

Model	MSE	RMSE	MAE	R^2
Informer	575.24	23.98	18.63	0.77
Autoformer	632.42	25.14	19.14	0.75
Dliner	585.40	24.20	18.28	0.77
Our model	374.99	19.36	14.05	0.85

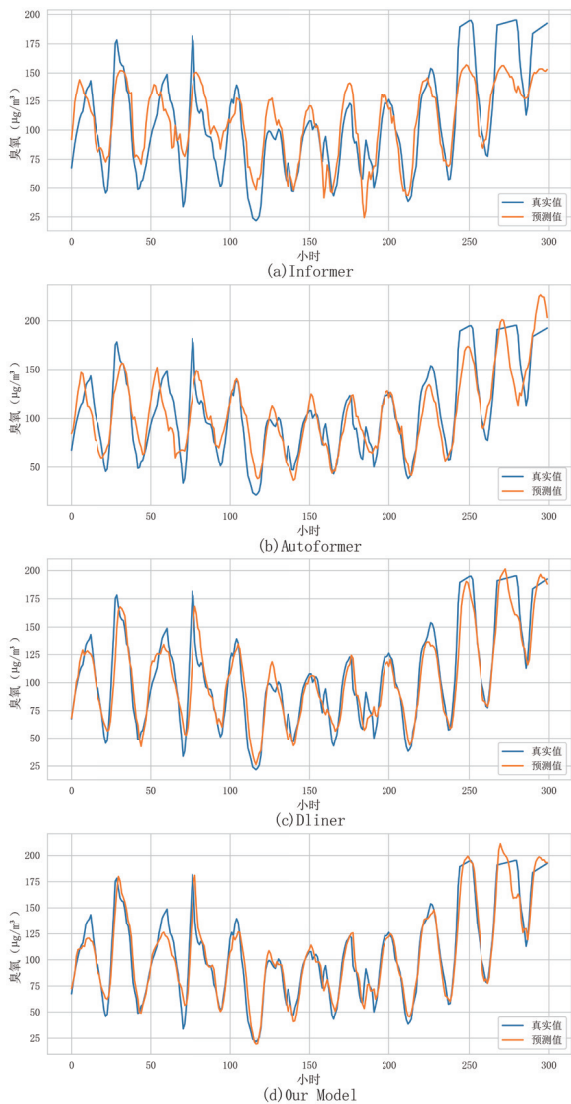


图 2 各模型预测结果的随机截取片段

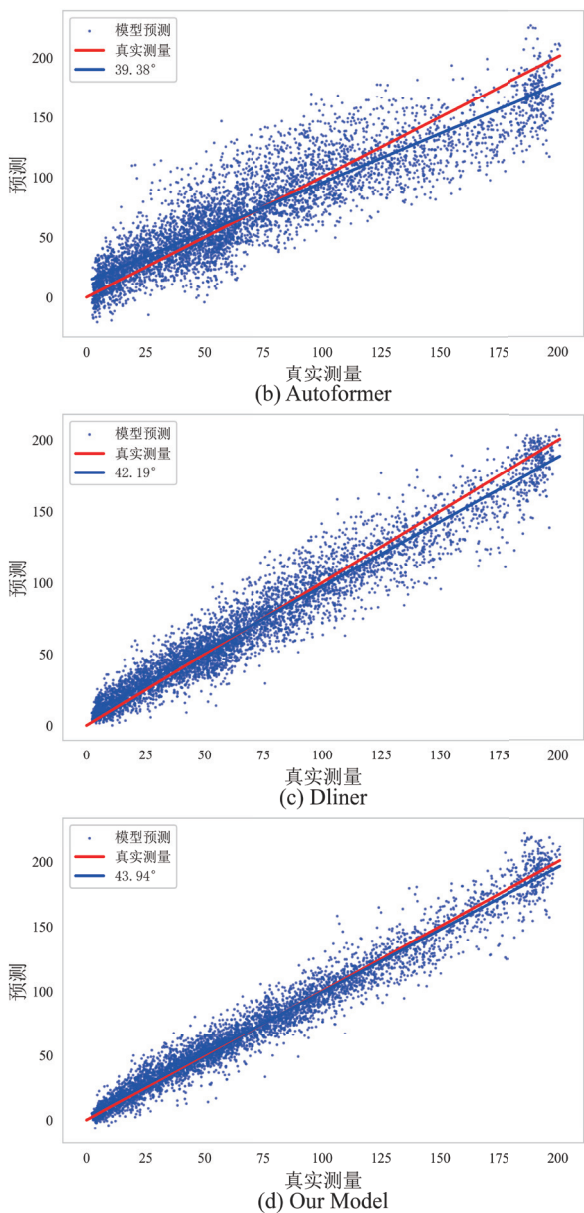
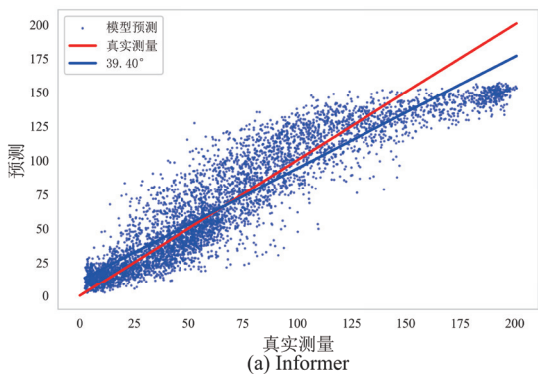


图 3 各模型预测结果的散点对比图

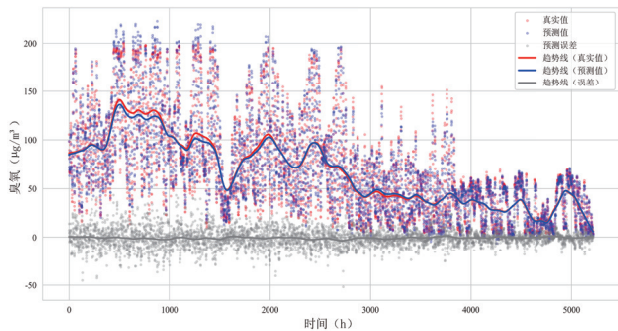


图 4 模型预测值与真实值的比较

3 总结

本文提出的基于多尺度序列间相关性的深度学习模型在地表臭氧浓度预测中展示了较高的准确性和鲁棒性。该模型

结合了 DWT、AGCN 和 TCN，通过多层次的特征提取和序列建模，能够有效捕捉臭氧浓度数据中的多尺度特征，尤其在极值预测和时空相关性方面表现出显著的优势。具体而言，DWT 能够将臭氧浓度数据进行多尺度分解，提取不同频率成分的特征，从而更好地识别数据中的细节变化与周期性趋势；AGCN 则能够捕捉数据之间的空间相关性，特别是在复杂的城市环境中，各地的污染物排放及气象条件可能呈现出强烈的空间依赖关系，AGCN 通过图卷积的方式有效融合不同站点的关联信息；TCN 则专注于时间序列的建模，能够在处理臭氧浓度的时间序列数据时捕捉长时间跨度上的复杂动态模式。

这种多尺度分解与卷积操作的组合方法在处理非线性和复杂的环境数据方面表现出色。传统的臭氧预测模型往往无法有效处理数据中的复杂时空变化模式，而该模型通过层层特征提取和深度建模，不仅提升了对复杂数据模式的适应性，还在面对极端天气事件和特殊污染源时，能够保持较高的预测精度。此外，该模型在不同的预测场景中表现出了较强的鲁棒性，即使在数据中存在噪声或缺失值的情况下，仍能够保证较为准确的预测结果。

在实际应用中，该深度学习模型可以为政策制定者提供实时、精准的臭氧浓度预警信息，有助于在臭氧浓度升高时采取适当的控制措施，减少污染对公众健康和环境的影响。通过进一步优化该模型的计算效率和实时性，未来可以将其应用于更大规模的城市空气质量监测与预测系统中，实现智能化的污染防治和管理。

参考文献:

- [1] MAO W J, JIAO L M, WANG W L. Long time series ozone prediction in China: a novel dynamic spatiotemporal deep learning approach [J]. Building and environment, 2022, 218:109087.
- [2]ZHANG J J, WEI Y J, FANG Z F. Ozone pollution: a major health hazard worldwide[J].Frontiers in immunology, 2019, 10: 2518.
- [3]HU Q H, JI X G, HONG Q Q, et al.Vertical evolution of ozone formation sensitivity based on synchronous vertical observations of ozone and proxies for its precursors: implications for ozone pollution prevention strategies[J].Environmental science & technology,2024, 58 (9): 4291-4301.
- [4]WU W L, FU T M, ARNOLD S R, et al.Temperature-dependent evaporative anthropogenic VOC emissions significantly exacerbate regional ozone pollution[J].Environmental science & technology, 2024, 58 (12): 5430-5441.

- [5] SEINFELD J H, PANDIS S N, NOONE K. Atmospheric chemistry and physics: from air pollution to climate change [J]. Physics today,1998,51(10):88-90.
- [6] GRELL G A, PECKHAM S E, SCHMITZ R, et al. Fully coupled "online" chemistry within the WRF model [J]. Atmos environ, 2005, 39(37): 6957-6975.
- [7] 王一龙,董韶妮,孙丽萍,等. 时间序列模型预测大气臭氧浓度 [J]. 济南大学学报 (自然科学版), 2023, 37(2): 178-183.
- [8] 杨晓彤,康平,王安怡,等. 基于随机森林模型的四川盆地臭氧污染预测 [J]. 环境科学, 2024, 45(5): 2507-2515.
- [9] 王凯,胡冬梅,闫雨龙,等. 基于深度学习的城市臭氧小时浓度预测模型 [J]. 环境化学, 2023, 42(8): 2609-2618.
- [10] RHIF M, ABBES A B, FARAH I R, et al. Wavelet transform application for/in non-stationary time-series analysis: a review [J]. Applied sciences, 2019, 9(7): 1345.
- [11] LI R Y, WANG S, ZHU F Y, et al. Adaptive graph convolutional neural networks[DB/OL].(2018-01-18)[2024-10-02]. <https://doi.org/10.48550/arXiv.1801.03226>.
- [12] BAI S J, KOLTER J Z, KOLTUN V. An empirical evaluation of generic convolutional and recurrent networks for sequence modeling [DB/OL].(2018-04-19)[2024-06-13].<https://doi.org/10.48550/arXiv.1803.01271>.
- [13] ZHOU H Y, ZHANG S H, PENG J Q, et al. Informer: beyond efficient transformer for long sequence time-series forecasting; proceedings of the proceedings of the AAAI conference on artificial intelligence[J]. Proceedings of the AAAI conference on artificial intelligence, 2021, 35(12): 11106-11115.
- [14] WU H X, XU J H, WANG J M, et al. Autoformer: decomposition transformers with auto-correlation for long-term series forecasting [DB/OL]. (2022-01-07)[2024-05-22]. <https://doi.org/10.48550/arXiv.2106.13008>.
- [15] ZENG A L, CHEN M X, ZHANG L, et al. Are transformers effective for time series forecasting[J]. Proceedings of the AAAI conference on artificial intelligence, 2023, 37(9): 11121-11128.

【作者简介】

朱振业 (1999—), 男, 江苏淮安人, 硕士研究生, 研究方向: 时间序列预测, email: zzy171099@163.com。

唐超礼 (1980—), 男, 安徽阜阳人, 博士, 教授, 研究方向: 信息处理技术, email: chltang@mail.ustc.edu.cn。

(收稿日期: 2024-12-04)