

基于深度学习和元启发式算法的 臭氧浓度集合预报

曹君乾, 李欢, 莫欣岳

(海南大学 网络空间安全学院(密码学院), 海口 570228)

摘要: 针对普遍而严重的空气污染问题, 特别是日益增加的臭氧污染对环境和公众健康带来的威胁, 提出了一种融合长短期记忆网络、时间卷积网络、Transformer 三种深度学习模型和元启发式算法开普勒优化算法的集合预报模型; 通过分析臭氧与其他空气污染物及气象要素的相关性, 选定了预报因子; 分别搭建了长短期记忆网络、时间卷积网络、Transformer 预报模型并开展独立的预报; 采用开普勒优化算法融合 3 种深度学习模型的预报结果, 最终生成集成预报结果; 实验结果显示在北京地区臭氧小时浓度的多步预报中, 提出的集合预报模型的均方根误差、平均绝对误差和决定系数均优于单一的深度学习模型长短期记忆网络、时间卷积网络和 Transformer 以及统计模型多元线性回归和传统机器学习模型随机森林; 研究结果表明融合深度学习和元启发式算法的集合预报策略能有效提升预报模型的准确性和稳定性, 验证了深度学习集合预报方法对臭氧污染预报的可行性。

关键词: 臭氧浓度预报; 长短期记忆网络; 时间卷积网络; Transformer; 开普勒优化算法; 集合预报

Ensemble Forecast for Ozone Concentration Based on Deep Learning and Metaheuristic Algorithm

CAO Junqian, LI Huan, MO Xinyue

(School of Cyberspace Security (School of Cryptology), Hainan University, Haikou 570228, China)

Abstract: In response to widespread and severe air pollution, especially the increasing threat of ozone pollution to the environment and public health, an ensemble forecasting model is proposed, which combines three deep learning models: long short-term network (LSTM), time convolutional network (TCN), and Transformer, as well as metaheuristic algorithm and Kepler optimization algorithm. Select predictive factors by analyzing the correlation between ozone and other air pollutants and meteorological elements. Build the LSTM, TCN and Transformer respectively and conduct independent forecasting. Use Kepler optimization algorithm to integrate the predictive outputs of three deep learning models, thus obtaining final predictive results. Experimental results indicate that, in multi-step forecast for hourly concentration of ozone in Beijing, the proposed ensemble forecasting model is superior in the RMSE, MAE and R^2 to deep learning models of the LSTM, TCN and Transformer as well as classical multiple linear regression and random forest models. The results show that the ensemble forecasting strategy integrating deep learning with metaheuristic algorithm can effectively improve the accuracy and stability of the forecasting model, and verify the feasibility of deep learning ensemble forecasting method for ozone concentration prediction.

Keywords: ozone concentration prediction; LSTM; TCN; Transformer; Kepler optimization algorithm; ensemble forecast

收稿日期:2024-05-23; 修回日期:2024-07-03。

基金项目:教育部供需对接就业育人项目(2023122798868);教育部产学研合作协同育人项目(220902070162538);海南省自然科学基金(623RC455,623RC457,425QN244);海南大学科研启动基金项目[KYQD(ZR)-22096,KYQD(ZR)-22097]。

作者简介:曹君乾(2002-),男,学士。

李欢(1990-),女,博士,讲师,博士生导师。

通讯作者:莫欣岳(1991-),男,博士,讲师,硕士生导师。

引用格式:曹君乾,李欢,莫欣岳. 基于深度学习和元启发式算法的臭氧浓度集合预报[J]. 计算机测量与控制, 2025, 33(7): 195-202.

0 引言

随着我国工业化和城市化步伐的加快,大气污染问题日趋严重,频发的污染天气给社会带来诸多负面影响,同时也造成了巨大的经济损失^[1-2]。为了降低损失,研究并实施有效的空气污染预报显得尤为重要^[3]。特别是近年来 O₃ 对健康和环境的危害日益凸显,精准预报臭氧浓度对空气污染防控具有重要意义^[4]。

空气污染物浓度受多种污染物和气象要素共同影响,具有高度非线性、非平稳特征,因此对空气污染物浓度的准确预报存在难度大、稳定性差、预报精度低等问题。目前,国内外许多学者在空气污染预测方面做了大量研究。其中,数据驱动的空气污染预报已成为当前一个新的研究方向,主要分为统计模型和机器学习模型两类。统计模型简单易用,预测的速度较快,对短期预测效果良好,但在长期预测和非线性序列分析等方面存在局限性,例如多元线性回归模型(Multiple Linear Regression)^[5]。为了克服这些缺点,机器学习方法被引入,如支持向量回归^[6]、极限学习机^[7]能有效处理非线性序列。近年来随着深度学习在多个领域取得了显著成功,有研究者将其应用于空气污染时间序列数据的预测中,特别是长短期记忆神经网络(LSTM)^[8-9]在这方面得到广泛应用。

文献[10]提出将时空相似性引入 LSTM 中,前序模块为时空相似选择输入模块,中序模块使用 LSTM 作为深度学习网络进行训练,后序模块根据目标站点特征选择不同的输出组合进行集成,该研究有效提升了预测精度,显示出对有效数据的选取能提高预测结果精度。文献[11]介绍了一种联合 Transformer 注意力机制的循环预测网络,使用 LSTM 联合 Transformer 注意力机制对时间变化和全局空间特征统一建模形成记忆信息并对 PM_{2.5} 浓度进行预测,虽然在一定程度上提高了预测的准确度但忽视了对其他污染物和气象要素对 PM_{2.5} 浓度的影响。文献[12]介绍了一种基于前馈神经网络的预测方法,该方法基于三层前馈神经网络逼近非线性函数并采用梯度下降算法改进学习率,能够有效预测未来时刻的 API,但该神经网络模型仍存在参数初始值对预测结果影响大、计算损失函数耗时等不足。文献[13]通过引入自编码神经网络的“编码-解码”构架与 LSTM 结合建立了空气污染预测模型,其中编码器用于提取出时间序列污染物浓度数据分布特征而解码器利用提取的特征预测未来时间内污染物浓度数据并且两者均采用多层 LSTM 结构,该模型虽然提高了对长时间序列预测的准确度,但随着预测时间的增加预测精度下降且训练集和测试集上的预测效果存在较大差异。文献[14]提出了基于 CEEMDAN 和排列熵的 GRU 空气污

染物预测模型,通过使用 CEEMDAN 算法分解和排列熵算法重新组合空气污染物浓度数据,减小了模型的总体运算规模并提高了预测度,但模型的构建和运行过程相对复杂。文献[15]提出将图卷积网络(GCN)嵌入到 LSTM 中,使用 GCN 学习污染物数据的时空相关性,并使用 LSTM 提取时间特征,从而预测 PM_{2.5} 污染物浓度数据。

目前,虽然已有学者基于深度学习方法进行空气污染预报研究,但也存在一些不足需要改进和完善。例如,深度学习模型超参数的选择和调整往往被忽视,进而影响了预报精度;对于预报因子的选择,只考虑预报量的历史数据而未考虑其他污染物和气象要素的影响;一些组合模型的架构和计算过程过于复杂,可能会限制其在实际应用中的推广。因此,为了进一步提高预报准确性和改进深度学习建模方法,本研究采用集合预报作为核心思路。集合预报^[16]是一种结合多个单一预测模型的方法,通过将各种预测模型的预测结果进行合理的融合,发挥不同模型的优势和特点,以提高预测的准确性和稳定性。本研究针对当前空气污染中突出的臭氧问题,采用先进的深度学习模型 LSTM、TCN 和 Transformer 作为集成员,并引入先进的元启发式算法开普勒优化算法作为融合算法,将集成员的预测结果融合起来得到最终的预报结果,并对模型在北京地区臭氧小时浓度的多步预报进行验证。本研究采用集合预报方式,充分应用了大数据和人工智能的技术优势,提供准确、可靠的臭氧浓度预报,有效服务于空气污染防控工作,为政府的环境管理决策和企业公众的生产生活提供有力支持。

1 数据与方法

1.1 数据描述

本文使用的数据来源于美国国家气候数据中心(NCDC, national climatic data center)。这数据包含了北京市自 2019 年 1 月 1 日至 2023 年 9 月 30 日期间的空气污染物浓度和气象要素的小时监测数据。空气污染数据包括 PM_{2.5}、PM₁₀、SO₂、NO₂、O₃ 和 CO,气象数据包括温度、气压、风速和风向。数据集被划分为训练集(60%)、验证集(20%)和测试集(20%)。表 1 为样本统计特征。

针对预报变量北京市臭氧浓度数据进行时间序列分析,可以得到臭氧浓度在不同时间尺度上的变化特征,包括其变化的周期性和规律性^[17]。图 1 显示,臭氧浓度的日变化呈现单峰模态,大约在 16 点左右,此时段紫外线辐射强度和温度较高,有利于加快大气中臭氧的生成速度,体现了日间光化学反应速率随太阳辐射强度变化所引起的臭氧浓度的规律性波动。臭氧浓度的年变

化具有明显的季节性, 从冬季至夏季, 随着气温升高和日照时间增加, 臭氧浓度逐渐升高, 在夏季达到全年的峰值; 然后随着气温降低和日照减少, 浓度又逐渐下降, 在冬季处于年度最低水平。

表 1 样本统计特征

变量	最小值	最大值	平均数	中位数	标准差
PM _{2.5}	1.25	421.63	36.20	25.06	35.94
PM ₁₀	3.91	6 167.78	67.14	52.78	88.93
SO ₂	1.41	30.72	3.37	2.63	2.17
NO ₂	1.84	133.88	27.91	23.81	17.67
O ₃	1.69	312.41	62.86	53.83	47.18
CO	0.13	3.60	0.59	0.52	0.35
温度	-23	42	13.94	15.4	12.04
气压	988.10	1 047.90	1 016.26	1 016	10.84
风速	0	17	2.56	2	1.92
风向	0	360	177.94	170	114.25

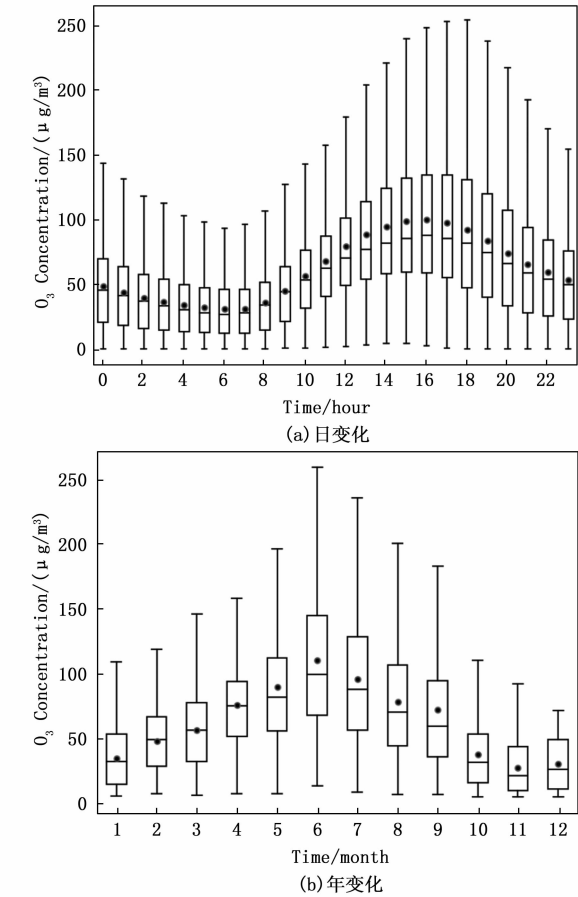


图 1 臭氧浓度变化的箱线图

1.2 研究方法

1.2.1 长短期记忆网络

LSTM 是一种特殊的循环神经网络 (RNN, recurrent neural network), 与 RNN 相比, LSTM 将 RNN 的隐含层用记忆单元来代替, 可以实现文本的长距离依赖编码, 缓解了传统循环神经网络梯度消失的问题, 它可

以有效地学习长序列数据中的长期依赖关系。LSTM 通过引入门控机制, 控制信息的流入、流出和内部状态更新。LSTM 模型包含 3 个门: 输入门、遗忘门和输出门。输入门控制新输入信息的流入, 遗忘门控制旧状态信息的保留, 输出门控制当前状态的输出。

LSTM 的更新公式可以表示为:

f_t = σ(W_f · [(h_t t - 1), x_t] + b_f) (1)

i_t = σ(W_i · [(h_t t - 1), x_t] + b_i) (2)

C_t = tanh(W_c · [(h_t t - 1), x_t] + b_c) (3)

C_t = f_t ⊙ C_t t - 1 + i_t ⊙ C_t (4)

o_t = σ(W_o · [(h_t t - 1), x_t] + b_o) (5)

h_t = o_t ⊙ tanh(C_t) (6)

式中, f_t、i_t 和 o_t 分别表示遗忘门、输入门和输出门的输出, C_t 表示候选细胞状态, C_t 表示当前细胞状态, h_t 表示当前隐藏状态, W_f、W_c 和 W_o 分别表示权重矩阵, b_f、b_i、b_c 和 b_o 分别表示偏置向量, σ 表示激活函数, ⊙ 表示元素乘积。

LSTM 能有效地捕捉时间序列数据中的长期依赖关系, 这有助于预测臭氧浓度数据中的趋势和周期性。其记忆单元可以保留和更新历史信息适用于处理具有记忆性质的时间序列数据。相对于传统的 RNN, LSTM 具有更好的长期记忆能力, 能够更好地处理时间序列数据中的梯度消失和梯度爆炸问题。

1.2.2 时间卷积网络

时间卷积网络 (TCN, temporal convolutional network) 是一种基于卷积神经网络 (CNN) 设计的特殊模型, 用于处理一维时间序列数据。与 LSTM 通过循环和门控机制来解决长距离依赖问题的方式不同, TCN 利用了一维卷积和空洞卷积技术, 以并行方式高效捕捉时序数据中的长期依赖关系。

TCN 的核心原理在于其层级结构的设计, 每一层通过一维卷积对输入序列的局部上下文信息进行编码, 并通过空洞卷积在不增加参数数量的前提下扩大感受野, 从而能够处理远距离的依赖。序列元素的空洞卷积运算定义为如式 (7) 所示:

F(s) = ∑_{i=0}^{k-1} f(i) · x_{s-d·i} (7)

式中, · 为卷积计算, d 为扩张数, k 为卷积核, f(i) 为卷积核中的第 i 个数据, x_{s-d·i} 表示仅对过去的数据进行卷积的结果, s-d·i 为过去的方向。

时间卷积网络摒弃了循环神经网络 (如 LSTM) 中的循环反馈结构, 转而依靠一维卷积和空洞卷积来逐步扩大其对时间序列数据的感受野, 以便于捕捉并处理数据中的长期依赖性。在预测臭氧浓度的任务上, 时间卷积网络展示了独特的优点。首先, 通过逐层递增的感受野设计, TCN 能够高效地捕获并学习时间序列数据中的长期趋势和周期性模式, 这对于理解并预测臭氧浓度

的波动走势至关重要。其次,与长短期记忆网络相比,TCN 的计算流程更利于并行化处理,这意味着在面对大规模时间序列数据时,TCN 的训练和预测速度得到显著提升。此外,TCN 在处理极其庞大的序列数据或预测远期臭氧浓度时,资源利用效率更高,不会像 LSTM 那样随着序列长度增加而遭遇训练难度加大和计算复杂度显著攀升的问题。

在包括臭氧浓度预测这样的应用场景中,TCN 因其对时间序列内部隐含趋势和周期性特征的强大识别能力,以及对长距离依赖的有效处理,配合其较低的计算复杂度和相对较低的超参数敏感度,使得 TCN 的结构和参数调整更为便捷,从而能够取得优异的预测效果。

1.2.3 Transformer

Transformer 是一种基于自注意力机制的深度学习模型,在计算机视觉中展现出巨大的潜力。它的核心思想是直接对序列的每个位置上应用自注意力而不是像循环神经网络那样依赖序列的顺序。

Transformer 模型的核心是自注意力机制,该模型在处理序列数据时会判断序列中不同位置之间的关联。自注意力机制是通过计算序列中每个位置的加权平均值从而得到每个位置的表示。自注意力机制的公式表示为:

$$\text{Attention}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = \text{softmax}\left(\frac{\mathbf{Q}\mathbf{K}^T}{\sqrt{d_k}}\right)\mathbf{V} \quad (8)$$

其中: $\mathbf{Q}$ 、 $\mathbf{K}$ 和 $\mathbf{V}$ 分别表示查询 (Query)、键 (Key) 和值 (Value) 矩阵, d_k 表示键向量维度。

由于 Transformer 没有循环结构,因此需要引入位置编码来表示序列中单词的位置信息。位置编码的公式可以表示为:

$$PE_{(pos, 2i)} = \sin(pos/10000^{2i/d}) \quad (9)$$

$$PE_{(pos, 2i+1)} = \cos(pos/10000^{2i+1/d}) \quad (10)$$

其中: $PE_{(pos, 2i)}$ 是位置函数,用于映射输入序列中 pos 处的元素到位置矩阵的 $(pos, 2i)$ 处, d 是输出嵌入空间的维度, i 是用于映射到列索引, $0 \leq i < d/2$, 单个值 i 映射到正弦和余弦函数。

Transformer 模型由编码器和解码器这两部分组成。编码器用于处理输入序列而解码器用于生成输出序列。每个编码器和解码器都包含着多个自注意力层和前馈神经网络层。

Transformer 模型能有效地处理多变量的时间序列数据。自注意力机制可以通过捕捉复杂的时间依赖关系来得到臭氧浓度变化的长期模式和规律。Transformer 能够克服非线性关系预测的局限性从而能够更全面地分析数据。

1.2.4 开普勒优化算法

开普勒优化算法 (KOA, kepler optimization algorithm)^[18] 是一种受开普勒行星运动定律启发的新型物理

基础启发式算法,其基本原理和核心公式借鉴了天体力学中行星围绕太阳运动的规律。在 KOA 中,每一个候选解被视为一颗行星,其位置在优化过程中随着搜索进程随时间随机更新,并且参考至目前为止找到的最佳解 (相当于“太阳”)。该算法结合了全局搜索和局部搜索的优势,能够在广阔的搜索空间内有效地探索和发掘高质量解,并能快速收敛到全局最优解。

1) 初始化:随机生成一组候选解 (行星),每个候选解代表问题可能解的一种状态,并赋予初始速度矢量。

2) 评估:对每个候选解 (行星) 进行评估,使用适应度函数衡量各个候选解的好坏,适应度高的候选解对应着更好的解质量。

3) 行星运动更新:根据开普勒定律,行星位置的更新反映了对太阳引力作用下的运动轨迹,核心公式可以抽象为:

$$x(t+1) = x(t) + v(t+1) \times \Delta t \quad (11)$$

其中: $x(t+1)$ 表示候选解在 $t+1$ 时刻的位置, $v(t+1)$ 是在 $t+1$ 时刻的速度矢量, Δt 为时间间隔,速度更新依赖于当前位置与太阳 (最优解) 的相对位置、质量和引力场强度等因素。

4) 速度更新则融合了局部搜索和全局搜索机制,模拟行星在不同阶段加速或减速的特性,可表达为:

$$v(t+1) = w \times v(t) + c_1 \times \text{rand}() \times$$

$$[p_{\text{best}} - x(t)] + c_2 \times \text{rand}() \times [g_{\text{best}} - x(t)] \quad (12)$$

其中: w 是惯性权重, c_1 和 c_2 是加速常数, $\text{rand}()$ 是 $[0, 1]$ 范围内的随机数, p_{best} 是个体历史最优位置, g_{best} 是全局最优解的位置。

5) 搜索过程模拟:KOA 并未直接描述遗传操作,而是通过类似力学系统的动力学规则模拟搜索过程。行星在其椭圆轨道上运动时,不仅受到太阳引力的影响,还会结合局部最优点和个人最优点进行动态调整,从而实现搜索空间的有效覆盖。

6) 迭代与终止:如同其他优化算法, KOA 也会经历迭代过程,每一轮迭代都根据前一轮的搜索结果更新所有候选解的位置和速度。算法会在达到预设的迭代次数上限或者适应度收敛到满足阈值的情况下终止,最后输出适应度最高的候选解作为问题的最优解。

综上所述,开普勒优化算法巧妙地运用了开普勒行星运动定律,通过模拟行星运动的过程,实现了对复杂优化问题的有效求解,在众多领域尤其是工程设计和参数估计等问题上展现出强大的优化能力和较高的搜索效率。开普勒优化算法通过模拟太阳系中行星运动规律,构造了一套完整的优化框架,相较于仅基于开普勒第一定律的简化模型, KOA 利用更为复杂的动力学机制,让搜索过程更加灵活且具有更强的全局优化能力。在实践中, KOA 被成功应用于多个具有挑战性的连续优化

问题,并在与其他多种最新发表的、已广泛研究的和高性能的随机优化算法对比中,显示出更高的性能潜力和更出色的搜索效果。

1.2.5 集合预报

集合预报技术,作为一种在统计学和机器学习领域被广泛采用的先进策略,通过集成多个不同模型的预报结果,以期达到超越任何单一模型的预报精度,实现预测效果的提升^[19]。其核心理念在于突出模型多样性的综合优势而非追求单一模型的极致精确性。集合预报不仅能够有效降低模型预报的偏差和方差,而且通过分散风险避免对单一模型的过度依赖和过拟合现象,这使其在对复杂系统预测例如高度非线性非平稳时间序列预测时的优势所在。

这一方法通过鼓励模型间的异质性,促使不同模型在各自擅长的数据子集或特征空间中深入挖掘信息进而强化整个预测体系的效能。集合预报系统内的成员相互补充,一个模型可能遗漏或误判的信息往往能被其他模型准确捕捉或修正,这种机制极大增强了模型系统的鲁棒性,使得即便在面对数据噪声或异常值时,集合预报依然能维持较为稳定和准确的预测性能。

综上所述,针对高度振荡的臭氧浓度时间序列,采用集合预报策略并融合多种先进的深度学习模型及元启发式算法,可以提升预报模型对臭氧污染预报的准确性,还能增强模型在复杂环境变化中的适应性和预报稳定性,为空气质量预报预警和空气污染防控策略的制定提供科学、有效的支撑,这也标志着智能预测技术在未来生态环保领域的创新应用和发展。

1.3 构建预报模型

本研究的预报模型架构和计算主要步骤如下:
根据预报因子分别构建 LSTM、TCN 和 Transformer 三种独立的预报模型,通过模型的训练和超参数

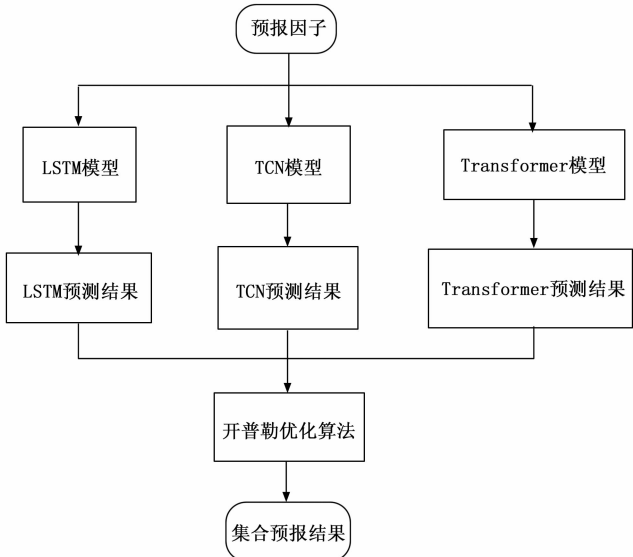


图 2 预报模型架构

的调整以得到最佳的预测结果;
将 LSTM、TCN 和 Transformer 三种模型独立预测得出的结果融合成最终的预报结果,其中采用先进的元启发式算法对不同模型的预测结果进行赋权从而得到更加精准和稳定的集合预报结果。

本研究采用了 3 种在时间序列分析方面具有显著优势的深度学习模型,即 LSTM 在揭示时序数据的长期依赖关系上有独到之处,TCN 在解析时间序列数据时擅长提炼局部时间和空间关联特征,而 Transformer 则在感知全局依赖结构方面表现出卓越性能。此外创新性地引入了先进的开普勒优化算法融合 LSTM、TCN 和 Transformer 模型进行集合预报,充分利用了开普勒优化算法在全局搜索、高效并行处理和自动适应复杂优化问题等方面的强大功能,同时发挥了不同预测模型的优势。总体来说,集合预报机制结合了各模型在处理非线性时间序列、存储和利用长短期记忆信息以及把握全局语境的优势,通过这种多模型融合的方式,集合预报模型能有效提升预测的鲁棒性。

2 实验过程

2.1 预报因子筛选

为了探究预报因子 PM_{2.5}、PM₁₀、SO₂、NO₂、CO、温度、气压、风向和风速与预报量 O₃ 之间的相关性,本文采用了斯皮尔曼等级相关系数^[20]进行研究。斯皮尔曼等级相关系数是统计学中用于衡量两个变量间单调等级相关强度的方法,不受数据分布类型限制,特别适用于非线性相关场景。

若有两个长度为 n 的序列 $X = \{x_1, x_2, \dots, x_n\}$ 和 $Y = \{y_1, y_2, \dots, y_n\}$,将每组数据进行排序,并用排名来表示。假设 X 的排名为 $\{x_1, x_2, \dots, x_n\}$, Y 的排名为 $\{y_1, y_2, \dots, y_n\}$ 。斯皮尔曼等级相关系数 ρ 的计算公式可以表示为:

$$\rho = 1 - \frac{6 \sum [r_{x_i} - r_{y_i}]^2}{n(n^2 - 1)}$$

(13)

其中: n 表示数据点的数量, r_{x_i} 和 r_{y_i} 分别是 X 和 Y 中第 i 个数据点的排名。 ρ 的取值范围介于-1到1之间,负值表示负相关,0表示无相关,正值表示正相关。

依据斯皮尔曼等级相关系数公式,本文计算了上述各项空气污染物以及气象因素与 O₃ 之间的相关性,并将结果可视化相关性热力图(图 3)。通过该热力图,可以直观地观察并分析不同因子与 O₃ 浓度的相关程度。

图 3 显示,在气象要素中,气温与臭氧浓度呈现最大的正相关性。夏季高温有利于大气中的氮氧化物(NO_x)和挥发性有机化合物(VOCs)的光化学反应,进而促进臭氧的生成,导致臭氧浓度升高^[21]。相反,气压与臭氧浓度之间呈现最强的负相关性。在高压条

件下,臭氧浓度的增加与气压的下降密切相关^[22]。但在此样本中,风速和风向与臭氧的相关性最低,远低于其他气象要素。PM_{2.5}、PM₁₀、SO₂、NO₂、CO 和 O₃之间存在复杂的相互作用和转化机制,彼此具有一定的相关性^[23]。总体来说,臭氧与其他大气污染物呈负相关,尤其是与 NO₂ 的负相关性最为显著。这是因为 NO₂ 和 CO 可以通过光化学反应生成臭氧。而 PM_{2.5}、PM₁₀ 与臭氧呈负相关是因为它们增加了气溶胶的光学厚度,从而阻碍了臭氧的生成。综上分析,为了提高模型的准确性和计算效率,剔除了相关性最小的风向、风速和 SO₂ 这几个因子,并考虑到温度、气压、NO₂、PM_{2.5}、PM₁₀ 和 CO 对臭氧浓度的影响,将它们作为预报因子纳入模型中,作为特征输入进行臭氧浓度的预测^[24]。

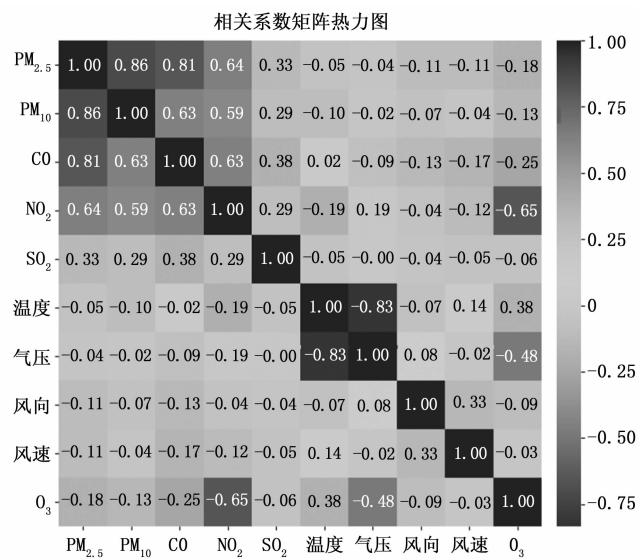


图 3 臭氧浓度与其他预报因子的相关性

2.2 模型参数设置

在搭建和训练 LSTM、TCN 和 Transformer 的时间序列预测深度学习模型时,超参数的选择和调整是整个模型构建的关键过程。其中预报时滞、学习率和批次大小是 LSTM、TCN 和 Transformer 模型的重要超参数。预报时滞决定了模型对时间序列数据历史信息的观察范围,而学习率影响着模型参数更新的速度和稳定性,批次大小影响着模型训练过程中的内存占用以及梯度估计质量。合理配置超参数对提高模型的性能有着重要作用。通过在验证集上使用 Trial-and-Error 方法探索了不同的学习率、预报时滞以及批次大小的组合,最终确定了最优的超参数配置,具体如表 2 所示。

表 2 模型参数设置

预测模型	Transformer	LSTM	TCN
学习率	0.001	0.000 1	0.001
预报时滞	24	12	24
批次大小	16	32	16

2.3 模型评价

本文使用经典的评价指标均方根误差 (RMSE)、平均绝对误差 (MAE) 和决定系数 R² 来评估不同模型的预测性能。RMSE 和 MAE 可以量化模型的预测误差大小,而 R² 则可以衡量模型对实际数据的解释能力。通过这些客观的量化指标,可以更全面地评估模型的性能。

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}}$$
 (14)

$$MAE = \frac{\sum_{i=1}^n |y_i - \hat{y}_i|}{n}$$
 (15)

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \hat{y}_i)^2}$$
 (16)

式中, y_i 和 $\hat{y}_i$ 分别表示第 i 个样本的实际值和预测值; n 为样本总数; $\bar{y}$ 表示所有样本实际值的平均值。

3 实验结果与分析

3.1 集合预报与集合成员对比

通过对集合成员 LSTM、TCN、Transformer 和集合预报的结果进行了对比分析,验证集合预报策略在提升模型预测精度和稳定性方面的效果。针对臭氧小时浓度的 3 步预报任务,在测试集上开展不同模型的统计对比分析。同时由于测试集数据过多,选取臭氧浓度最高一周的预测结果为代表进行可视化展示。由图 4 显示的各模型预报结果可以看出,本研究建立的 4 种模型均取得了良好的预报结果,预报值与观测值接近,变化趋势基本一致。然而单一的集合成员仍存在不足,例如 LSTM 对峰值的估计偏差较大且对趋势的把握不够准确,TCN 的预测值总体偏低,Transformer 的预报结果仅次于集合预报但对一些拐点的估计上误差大于集合预报,而集合预报模型则展现了更高的预测精度和稳定性。

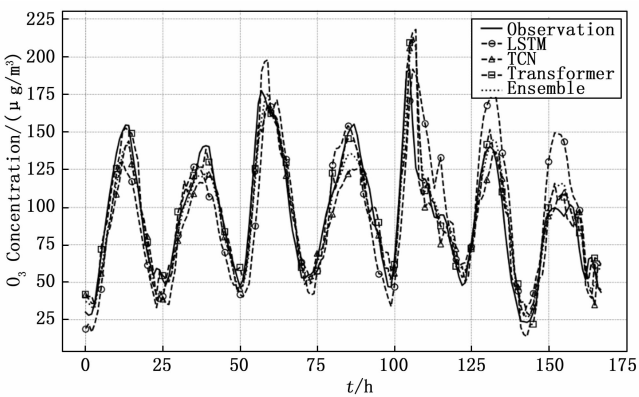


图 4 集合预报与集合成员的预报结果

从表 3 的统计结果可以得出,集合预报模型在对 O₃ 小时浓度的多步预报中显示出最低的误差水平。相较于 LSTM、TCN 以及 Transformer 模型,集合预报模型在 RMSE 上分别降低了 45.3%、12.3% 和 8.4%,

在 MAE 上分别减少了 52.9%、25.0% 和 7.5%。而在衡量模型拟合优度的 R^2 指标上分别提升了 40.4%、42.6% 和 27.1%。

表 3 各模型评价结果

预测模型	RMSE	MAE	R^2
Multiple Linear Regression	29.32	22.96	0.43
Random Forest	25.68	19.81	0.56
LSTM	23.52	18.82	0.63
TCN	15.52	11.74	0.84
Transformer	14.01	9.37	0.87
Ensemble Forecast	12.98	8.80	0.89

综上所述,集合预报模型在臭氧预测上取得了较高的预测精确度,相较于 LSTM、TCN 以及 Transformer 模型,集合预报模型拥有更强的预测性能。与此同时,实验结果也显示了 3 个集成员在时序预测上的可靠性。LSTM 模型在处理时间序列数据中的长期依赖关系上独具优势,TCN 模型则擅长通过深度卷积结构捕获时间序列中的局部和时间相关的特征,而 Transformer 模型在处理长距离依赖关系和高效抓取全局信息方面表现突出。集合预报模型通过应用元启发式算法巧妙地融合了这些模型各自的优势,进一步提升了对臭氧浓度预报的性能。

3.2 集合预报与经典模型对比

此外,进一步对比了提出的集合预报模型和经典的统计模型多元线性回归 (Multiple Linear Regression) 和机器学习模型随机森林 (Random Forest) 在臭氧浓度多步预报上的结果。实验结果显示,集合预报模型在预测性能上展现出了显著优势。从表 3 和图 5 可以清晰看出,集合预报模型在各个评价指标上都获得了出色的表现。在 $RMSE$ 上,集合预报模型仅为 12.98,明显低于随机森林模型的 25.68 和多元线性回归模型的 29.32。同样在 MAE 上,集合预报模型为 8.80,相比随机森林模型的 19.81 和多元线性回归模型的 22.96,集合预报模型的误差显然更小,体现了更高的精度水

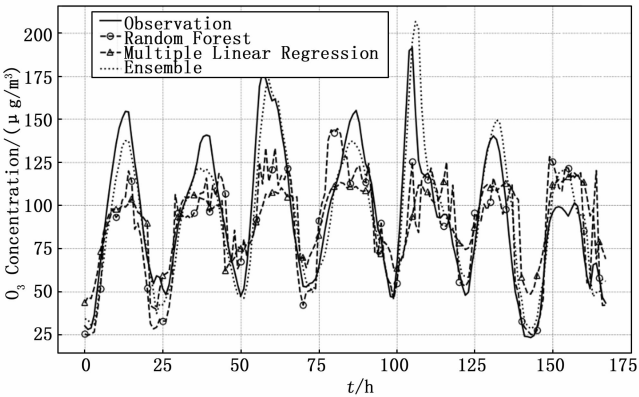


图 5 集合预报与经典模型的预报结果

平。在 R^2 上,集合预报模型达到了 0.89,这一数值远高于随机森林模型的 0.56 和多元线性回归模型的 0.43,揭示了集合预报模型在预测数据和变化趋势方面具备更强的能力。

集合预报模型融合了多种深度学习模型的优势,有效降低了预测误差,提高了模型的拟合度和预测准确性,从而在各项评价指标上均优于经典模型和集成员。这也意味着使用集合预报技术建立深度学习组合模型,对于空气污染的预报预警相关工作是一种有效、可靠的方案。

4 结束语

本研究针对日益严峻的臭氧污染问题,提出了一种创新的预报方法,结合了 LSTM、TCN 和 Transformer 三种深度学习模型的优势,并采用开普勒优化算法进行集合预报,以提高预测的准确性和可靠性,同时降低单一模型的局限性和误差,从而在 O_3 浓度的多步预报中获得更高的预报精度和稳定性。

实验证明,集合预报模型在臭氧浓度预报上展现出显著的优势。与 LSTM、TCN 和 Transformer 等集成员相比,集合预报模型在 $RMSE$ 、 MAE 以及 R^2 这 3 个评价指标上均有明显提升,降低了预报误差并提高了模型的拟合优度。此外,与多元线性回归和随机森林等经典统计模型和传统机器学习模型相比,基于深度学习的集合预报在预测性能上优势显著。

综上,集合预报模型通过融合多种模型的优点,减少了单个模型预测的不确定性,能够适应复杂的环境变化,提供更为全面和可靠的预测。基于深度学习和元启发式算法的集合预报为今后的空气污染预报预警提供了一个新思路、新方案,能有效服务于政府空气污染防控和智慧环保城市建设。

参考文献:

[1] MO X, LI H, ZHANG L, et al. Environmental impact estimation of $PM_{2.5}$ in representative regions of China from 2015 to 2019: policy validity, disaster threat, health risk, and economic loss [J]. Air Quality, Atmosphere & Health, 2021, 14: 1-15.

[2] MO X, LI H, ZHANG L, et al. A novel air quality evaluation paradigm based on the fuzzy comprehensive theory [J]. Applied Sciences, 2020, 10 (23): 8619.

[3] 莫欣岳. 基于人工智能的空气质量决策支持系统建模研究 [D]. 兰州: 兰州大学, 2021.

[4] 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.

[5] 郑 红, 程云辉, 胡阳生, 等. 基于 MLP&ST 模型的空

- 气质量预测 [J]. 应用科学学报, 2022, 40 (2): 302 - 315.
- [6] 郑洋洋, 白艳萍, 续 婷. 基于 SARIMA-SVR 组合模型的空气质量指数预测 [J]. 河北工业科技, 2019, 36 (6): 436 - 441.
- [7] MO X, LI H, ZHANG L, et al. A novel air quality early-warning system based on artificial intelligence [J]. International Journal of Environmental Research and Public Health, 2019, 16 (19): 3505.
- [8] 朱菊香, 谷 卫, 罗丹悦, 等. 基于 CEEMD-SSA-LSTM 的空气质量预测 [J]. 国外电子测量技术, 2022, 41 (11): 120 - 129.
- [9] MO X, LI H, ZHANG L, et al. Design a regional and multistep air quality forecast model based on deep learning and domain knowledge [J]. Frontiers in Earth Science, 2022, 10: 995843.
- [10] 方 伟, 朱润苏. 基于时空相似 LSTM 的空气质量预测模型 [J]. 计算机应用研究, 2021, 38 (9): 2640 - 2645.
- [11] 刘恩海, 付英健, 张 智, 等. 联合 Transformer 注意力机制的 $PM_{2.5}$ 浓度预测网络研究 [J]. 安全与环境学报, 2023, 23 (10): 3760 - 3768.
- [12] 韩 广, 徐 璐, 孙晓云, 等. 空气污染指数的前馈神经网络预测方法 [J]. 计算机与应用化学, 2016, 33 (2): 147 - 151.
- [13] 秦东明, 丁志军, 金玉鹏, 等. 基于自编码网络的空气污染物浓度预测 [J]. 同济大学学报 (自然科学版), 2019, 47 (5): 681 - 687.
- [14] 凌德森, 王晓凯, 朱 涛. 基于自适应时序分解的空气污染物浓度预测 [J]. 测控技术, 2023, 42 (1): 83 - 91, 98.
- [15] CHEN Q, DING R, MO X, et al. An adaptive adjacency matrix-based graph convolutional recurrent network for air quality prediction [J]. Scientific Reports, 2024, 14: 4408.
- [16] 刘莘昱, 郭志民, 王 超, 等. 基于集合预报的输电线路舞动概率预测方法 [J]. 制造业自动化, 2023, 45 (6): 175 - 179.
- [17] 蒋伯琪, 浮 天, 程映璇, 等. 沈阳市臭氧污染特征及其影响因素 [J]. 生态环境学报, 2024, 33 (1): 72 - 79.
- [18] MOHAMED A B, REDA M, SHAOMAA A, et al. Kepler optimization algorithm: A new metaheuristic algorithm inspired by Kepler's laws of planetary motion [J]. Knowledge-Based Systems, 2023, 268: 110454.
- [19] GAO Z, MO X, LI, H. Prediction of $PM_{2.5}$ concentration based on deep learning, multi-objective optimization, and ensemble forecast [J]. Sustainability 2024, 16: 4643.
- [20] 贾 科, 杨 哲, 魏 超, 等. 基于斯皮尔曼等级相关系数的新能源送出线路纵联保护 [J]. 电力系统自动化, 2020, 44 (15): 103 - 111.
- [21] 王宇蝶, 滕泽宇, 陈智文, 等. 环渤海地区空气质量时空变化特征及动态预测 [J]. 中国环境监测, 2024, 40 (1): 68 - 78.
- [22] 张小玲, 雷 雨, 王聪聪, 等. 四川盆地城市群臭氧污染时空精细演变特征的观测与模拟 [J]. 中国环境科学, 2024, 44 (1): 1 - 14.
- [23] 杨丽蓉, 徐学哲, 季 群, 等. 银川市夏季臭氧污染特征及敏感性分析 [J]. 环境污染与防治, 2024, 46 (1): 14 - 19.
- [24] 杨雨佳, 肖庆来, 陈 健, 等. 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), 2023, 59 (2): 322 - 332.
- ~~~~~
- (上接第 194 页)
- [11] GHIASI G, LEE H, KUDLUR M, et al. Exploring the structure of a real-time, arbitrary neural artistic stylization network [J]. ArXiv Preprint ArXiv: 1705.06830, 2017.
- [12] DUMOULIN V, SHLENS J, KUDLUR M. A learned representation for artistic style [J]. ArXiv Preprint ArXiv: 1610.07629, 2016.
- [13] JACKSON P T G, ABARGHOUEI A A, BONNER S, et al. Style augmentation: data augmentation via style randomization [C] //CVPR Workshops, 2019, 6: 10 - 11.
- [14] ISOLA P, ZHU J Y, ZHOU T, et al. Image-to-image translation with conditional adversarial networks [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017: 1125 - 1134.
- [15] ZHU J Y, PARK T, ISOLA P, et al. Unpaired image-to-image translation using cycle-consistent adversarial networks [C] // Proceedings of the IEEE International Conference on Computer Vision, 2017: 2223 - 2232.
- [16] CHEN W, HAYS J. Sketchygan: Towards diverse and realistic sketch to image synthesis [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 9416 - 9425.
- [17] DAI Q, CHENG X, WANG Y, et al. Light-weight automatic residual scaling network for image super-resolution reconstruction [J]. Journal of Computer Applications, 2020, 40 (5): 1446.
- [18] XU B, CHEN M, GUAN W, et al. Efficient teacher: Semi-Supervised object detection for Yolov5s [J]. ArXiv Preprint ArXiv: 2302.07577, 2023.
- [19] ZHAO Q, SHENG T, WANG Y, et al. M2det: A single-shot object detector based on multi-level feature pyramid network [C] // Proceedings of the AAAI Conference on Artificial Intelligence, 2019, 33 (1): 9259 - 9266.