

基于AQI-Transformer的生态环境大气污染浓度预测研究

范建¹, 胡兴元^{1*}, 何君², 刘京¹, 武建双¹

¹中检集团天帷信息技术(安徽)股份有限公司, 人工智能研究中心, 安徽 合肥

²滁州学院外国语学院, 安徽 滁州

收稿日期: 2025年6月16日; 录用日期: 2025年6月22日; 发布日期: 2025年7月31日

摘要

工业化和城市化的快速发展导致大气污染加剧, 对公众健康及生态环境造成显著危害。本文主要介绍了人工智能与大数据技术在大气污染防治中的创新应用及效果分析。文章探讨了人工智能在构建空气质量预测模型以及优化决策策略等方面的作用, 指出其能为决策者提供定量分析依据, 提高污染防治的科学性和有效性。同时, 还分析了数据分析与挖掘以及智慧环保建设等方面的融合创新, 强调了人工智能与大数据、FaaS函数技术在提高监测效率、发现污染规律和支持政策制定等方面的重要作用以及对未来发展趋势的展望。通过实际数据集上的实验验证, 本文提出的AQI-Transformer模型在短期(4小时)与中长期(24小时)的污染浓度预测任务中, 均展现出了更低的均方根误差(RMSE)和平均绝对误差(MAE), 证明了其在提升大气污染防治效率与效果方面的潜力。

关键词

生态环境, 大数据, 人工智能, AQI-Transformer, 污染浓度预测

Prediction of Atmospheric Pollution Concentration in Ecological Environment Based on AQI-Transformer

Jian Fan¹, Xingyuan Hu^{1*}, Jun He², Jing Liu¹, Jianshuang Wu¹

¹China Inspection Group Tianwei Information Technology (Anhui) Co. Ltd., Artificial Intelligence Research Center, Hefei Anhui

²School of Foreign Languages, Chuzhou University, Chuzhou Anhui

Received: Jun. 16th, 2025; accepted: Jun. 22nd, 2025; published: Jul. 31st, 2025

*通讯作者。

文章引用: 范建, 胡兴元, 何君, 刘京, 武建双. 基于 AQI-Transformer 的生态环境大气污染浓度预测研究[J]. 人工智能与机器人研究, 2025, 14(4): 1064-1076. DOI: 10.12677/airr.2025.144101

Abstract

The rapid development of industrialization and urbanization has led to the aggravation of air pollution, which has caused significant harm to public health and ecological environment. This paper mainly introduces the innovative application and effect analysis of artificial intelligence and big data technology in the prevention and control of air pollution. This paper discusses the role of artificial intelligence in building air quality prediction models and optimizing decision-making strategies, and points out that it can provide quantitative analysis basis for decision-makers and improve the scientificity and effectiveness of pollution prevention and control. At the same time, it also analyzes the integration and innovation of data analysis and mining and smart environmental protection construction, and emphasizes the important role of AI, big data and FAAS function technology in improving monitoring efficiency, discovering pollution laws and supporting policy-making, as well as the prospect of future development trends. Through the experimental verification on the actual data set, the AQI transformer model proposed in this paper shows lower root mean square error (RMSE) and mean absolute error (MAE) in the short-term (4 hours) and medium and long-term (24 hours) pollution concentration prediction tasks, which proves its potential in improving the efficiency and effect of air pollution prevention and control.

Keywords

Ecological Environment, Big Data, Artificial Intelligence, AQI-Transformer, Pollution Concentration Prediction

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

1.1. 研究背景与意义

研究背景：全球气候变化和城市化加速导致大气中污染物排放增加，空气质量急剧下降，包括颗粒物、二氧化硫、氮氧化物及温室气体等，大气污染已成为全球性环境问题。在中国，尽管政府已采取措施，但污染问题依旧严峻，冬季雾霾频发，影响能见度、交通及人体健康。为解决此问题，中国政府 and 科研机构正积极探索新技术手段[1]-[3]。

随着大数据和人工智能技术的快速发展，它们在环境监测、污染源识别等方面展现出巨大潜力。大数据技术对海量数据的存储和分析，挖掘数据背后趋势和规律；人工智能技术则通过机器学习、深度学习等算法，自动识别和预测复杂事件，提高监测准确性和效率。

研究意义：人工智能与大数据技术在环保领域的应用至关重要。它们能精准监测和预警大气污染，及时发现污染物变化趋势，动态识别污染源，为预防和控制提供支持。有助于有效控制和管理环境污染，减少排放和扩散，提升环境质量。未来，随着技术不断进步，它们将在环保领域发挥更重要作用。

1.2. 国内外研究现状

在中国，作为人口众多、经济发展迅速的国家，大气污染防治面临巨大挑战。政府及相关机构借助人工智能与大数据技术，开展创新研究，如深度分析空气污染数据，揭示污染物分布、来源及传播途径，

预测污染趋势，优化控制策略，降低成本，提高效率。

与国外相比，后者在此应用领域更加成熟。发达国家已建立完善的大气污染防治体系，利用人工智能技术精准分析气象、污染源等复杂数据，为防治提供强大技术支持。同时，深度挖掘和分析空气污染数据，为制定精准监控和治理策略提供科学依据。

总之，本研究旨在通过构建基于 AQI-Transformer 的大气污染浓度预测模型，探索人工智能与大数据技术在大气污染防治中的创新应用。研究内容包括模型设计、实验验证、性能评估及未来发展趋势展望等。通过实际数据集上的实验，验证 AQI-Transformer 模型在短期和中长期污染浓度预测任务中的性能优势，为大气污染防治提供新的技术思路和方法[4]-[6]。

2. 基于 AQI-Transformer 的大气污染浓度预测

2.1. 实验环境与数据集

1. 实验环境准备：采用了云计算平台，以确保系统高可用性和弹性扩展能力。为 FaaS 函数部署和执行提供了坚实基础，能够灵活构建和管理大气污染防治大数据 AI 分析平台。以及采用了分布式存储系统 (HDFS) 和大数据处理框架 (Spark)，为数据建模和分析提供支持。整个人工智能模型训练过程是在配备 2.20 GHz 的 Intel(R) Xeon(R) E5-2699 v4 CPU、64GB RAM 以及 4 块 48GB 显存 NVIDIA GeForce RTX 4090 GPU 工作站上进行，使用 FSDP 分布式训练策略。超参数设置：初始学习率 $3e-4$ (采用 cosine 衰减)，batch_size = 256，训练 epoch = 300。数据划分比例：训练集:验证集:测试集 = 7:2:1，严格保证时序连续性。

2. 数据预处理：空气污染研究的数据来源气象数据(涵盖了不同地区、不同时间段的大气污染物浓度、气象条件等信息)、污染排放数据(企业排放数据、现场监测数据)、空气质量监测以及历史记录数据(包含 O₃、SO₂、CO、NO₂、颗粒物、AQI 等信息)。这些数据在时间和空间上具有高度真实多样性，有助于提高模型预测精度和泛化能力[7]-[9]。其数据样例见表 1。

过程中充分利用了 FaaS 函数服务架构的优势。通过 FaaS 函数能够弹性地扩展数据处理能力，以应对不同时间段数据量的波动。当数据采集量激增时，FaaS 能够自动调配更多资源来处理数据，确保数据采集与整合的高效进行。

在数据处理初期，仔细筛选并核实所收集信息，确保其真实准确。整合多源数据，通过匹配提升全面性和精确度。实施标准化与规范化处理，去除异常值、处理缺失值，统一数据格式和单位，填补数据缺失，确保数据集完整可靠。运用统计学或机器学习算法去除噪声，保证输入数据准确有效。将环境数据划分为训练集、测试集和验证集，训练集用于模型学习，验证集实时验证调整模型，监控泛化性能，识别过拟合或欠拟合现象。测试集独立评估模型性能，确保模型在未见数据上的鲁棒性[10]-[12]。其中污染排放监控数据、污染因子数据示例如下表所示：

Table 1. Statistics of pollution factors

表 1. 污染因子数据统计表

优良天数比例 (%)	SO ₂ 浓度均值 (ug/m ³)	NO ₂ 浓度均值 (ug/m ³)	PM25 浓度均值 (ug/m ³)	PM10 浓度均值 (ug/m ³)	CO 浓度均值 (mg/m ³)	03 日最大 8h 值 90%位数数值浓度均值(ug/m ³)
2024/09-2024-12	2024/09-2024-12	2024/09-2024-12	2024/09-2024-12	2024/09-2024-12	2024/09-2024-12	2024/09-2024-12
85.0	20	30	35	50	0.9	120
81.0	25	36	35	53	0.8	115

续表

83.0	22	33	32	52	0.9	123
89.0	24	34	34	51	0.8	118
88.0	29	32	31	52	0.7	124
88.0	21	35	33	57	0.6	124
86.0	23	31	35	52	0.7	120
.....	20	38	35	50	0.8	120

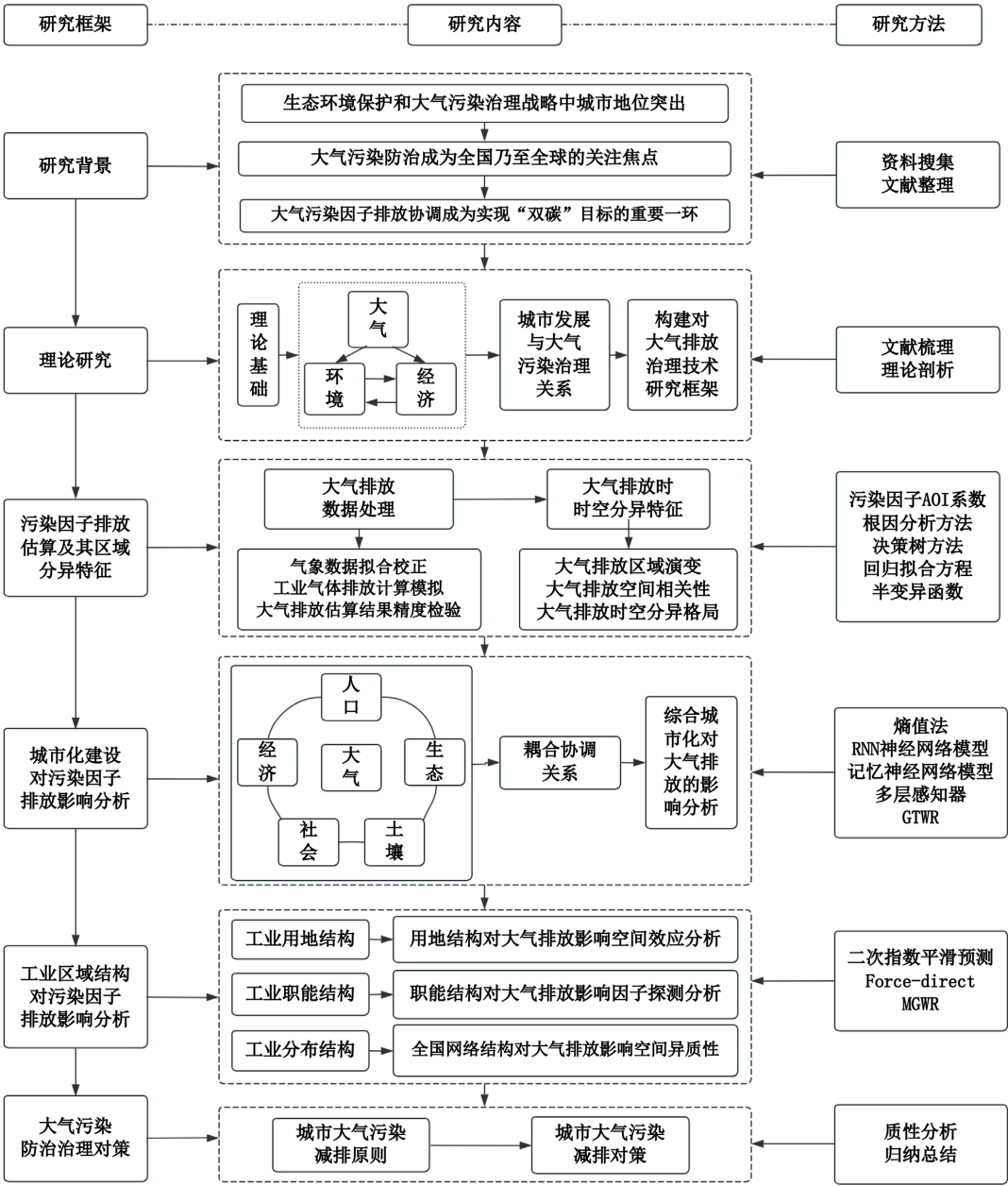


Figure 1. Research technology roadmap
图 1. 研究技术线路图

2.2. 基于 AQI-Transformer 的预测模型设计

本模型基于 Transformer 架构创新性整合时空特征分析与动态属性筛选机制, 专为污染物浓度预测任务优化。如图 1 所示, 系统框架由动态特征筛选层(基于 AQI 相关系数的多源属性优化选择)、时空表征层(融合位置编码与时间语义的复合编码器)、深度编解码网络层(含 Na 层编码器与 Nb 层解码器的级联处理架构)、预测输出模块层(掩码时序解码与线性映射机制) [13]-[15]。如图 2 所示。

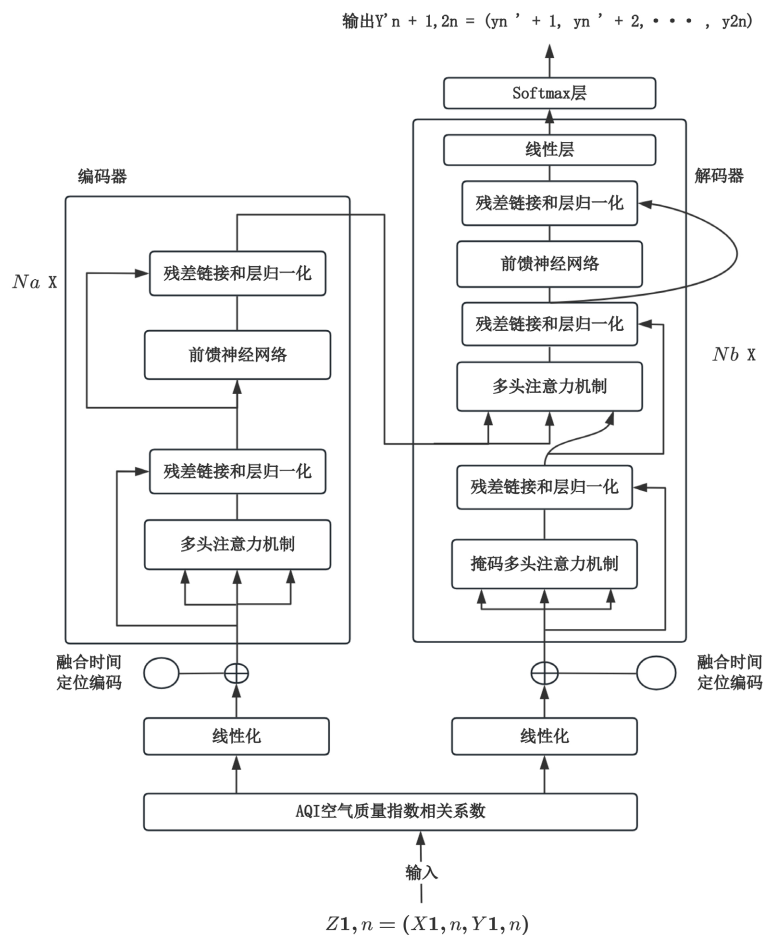


Figure 2. Structure diagram of AQI-Transformer

图 2. AQI-Transformer 模型结构

其中中心化浓度矩阵协方差计算公式为

$$\text{Cov}(X_{1,n}^u, Y_{1,n}) = \frac{1}{n^2} \sum_{a=1}^n \sum_{b=1}^n x_a' y_{a,b}'. \quad (1)$$

其中 AQI 空气质量指数相关系数计算公式

$$\text{Cor}(X_{1,n}^u, Y_{1,n}) = \frac{\text{Cov}(X_{1,n}^u, Y_{1,n})}{\sqrt{\text{Cov}(X_{1,n}^u, X_{1,n}^u) \text{Cov}(Y_{1,n}, Y_{1,n})}}. \quad (2)$$

$X_{1,n}^u$ 和 $Y_{1,n}$ 的 AQI 空气质量指数相关系数越大, 说明 $X_{1,n}^u$ 和 $Y_{1,n}$ 的相关性越强。通过 AQI 空气质量指数相关系数筛选属性, 保留 $|\rho| > 0.7$ 的强相关特征生成优化序列 X' 。

1. AQI-Transformer 模型架构

(1) 底层: Transformer-Xception 时空特征提取网络空间特征处理: 采用改进的 Xception 网络(和其他卷积模型的参数对比见表 2 所示), 将城市划分为 $1 \text{ km} \times 1 \text{ km}$ 网格(输入尺寸: $256 \times 256 \times 3$), 每个通道对应 PM2.5、NO₂、O₃ 浓度热力图时间特征处理: 集成 Transformer 模块, 设置 12 个注意力头, 处理 72 小时时序数据(时间窗口滑动步长 = 1 h)。

特征融合: 通过空间 - 时间交叉注意力机制, 实现特征矩阵拼接:

$$F_{\text{output}} = T\left(\left\{x(I_t)\right\}_{t=1}^{72}\right) \quad (3)$$

式(3)说明:

- X 表示改进的 Xception 网络, 通过深度可分离卷积提取多尺度空间特征。
- T 表示 Transformer 模块, 通过多头自注意力捕获 72 小时时序依赖。
- 最终输出 F_{output} 融合了时空联合语义, 可用于 PM2.5 等污染物浓度预测任务。

式(3)分步解释:

1) 空间特征提取(改进的 Xception 网络):

$$X: \mathbb{R}^{256 \times 256 \times 6} \rightarrow \mathbb{R}^{C \times H \times W}$$

输入为 $I_t \in \mathbb{R}^{256 \times 256 \times 6}$ (第 t 小时的 PM2.5、NO₂、O₃、PM10、CO、SO₂ 热力图, 输出为空间特征图 $F_{\text{spatial},t} \in \mathbb{R}^{C \times H \times W}$ 。

2) 时空特征融合编码:

AQI-Transformer 通过设计融合式时空编码系统, 将局部位置感知(基于正弦/余弦函数的位置编码)与全局时间语义(如季节、昼夜等宏观时间规律)相结合, 突破了传统 Transformer 仅依赖序列位置编码的局限, 显著提升了污染物浓度时序特征的捕获能力。规则巧妙地使用了余弦和正弦函数特性[16]-[18], 对污染浓度数据的局部位置编码(positional encoding, PE)具体定义为

$$\text{PE}(p, 2j) = \sin\left(\frac{p}{10000^{2j/d_{\text{model}}}}\right) \quad (4)$$

$$\text{PE}(p, 2j+1) = \cos\left(\frac{p}{10000^{2j/d_{\text{model}}}}\right) \quad (5)$$

公式(4)~(5)定义了基于正弦 - 余弦交替的位置编码方法: 其中 p 代表序列位置, d 为特征维度, 偶数维度($2j$)采用正弦编码, 奇数维度($2j+1$)使用余弦编码, 共同构成完整的局部位置表征。

AQI-Transformer 创新性地引入全局时间编码机制, 通过将月份、日期、星期和小时等时间参数进行归一化处理(式 14), 有效补充了局部位置编码在时间相关性表征上的不足, 显著提升了污染物浓度预测的时序建模能力。

$$\tilde{M} = \frac{M-6.5}{12}, \quad \tilde{H} = \frac{H-11.5}{24} \quad (6)$$

$$\text{GE}(Z'_i) = (t'_{\text{month}}, t'_{\text{week}}, t'_{\text{day}}, t'_{\text{hour}}) W_t \quad (7)$$

$$\text{GP}(Z'_{1,n}) = \text{PE}(Z'_{1,n}) + \text{GE}(Z'_{1,n}) \quad (8)$$

其中编码融合机制: 通过式(7)对时间参数进行线性映射生成全局时间编码, 再与式(4)~(5)定义的局部位置编码按式(8)叠加, 形成融合时空特征向量, 最终与原始污染数据相加后输入编解码模块进行特征学习[19][20]。

3) 时间特征建模(多头 Transformer 编码器):

$$T = \text{MultiHeadAttention}(Q, K, V) + \text{FFN} \quad (9)$$

- 输入序列: $\{F_{\text{spatial},t}\}_{t=1}^{72}$ (72 个时间步的空间特征);
- 查询(Q)、键(K)、值(V)投影:

$$\begin{aligned} Q &= W_Q \cdot \text{Flatten}(F_{\text{spatial}}), \\ K &= W_K \cdot \text{Flatten}(F_{\text{spatial}}), \\ V &= W_V \cdot \text{Flatten}(F_{\text{spatial}}) \end{aligned} \quad (10)$$

- 输出时空特征 $F_{\text{output}} \in \mathbb{R}^{L \times d}$, 其中 L 为序列长度, d 为特征维度。
- 多头注意力计算:

$$\text{Attention}(Q, K, V) = \text{Softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (11)$$

多头注意力机制将查询 Q 、键 K 和值 V 分别以不同的线性投影多次投影到 d_k 、 d_k 和 d_v 维, 假设头数为 b , d_k 和 d_v 的计算公式为:

$$d_k = d_v = \frac{d_{\text{model}}}{b} \quad (12)$$

式(11)~(12)中: d_k , d_v 为投影维度; d_{model} 为输入维度。在查询 Q 、键 K 和值 V 的每个投影层次上, 并行地执行注意力函数, 生成 d_v 维输出值, 将它们连接起来并再次进行投影, 得到最终值, 多头注意力机制如图 3 所示。

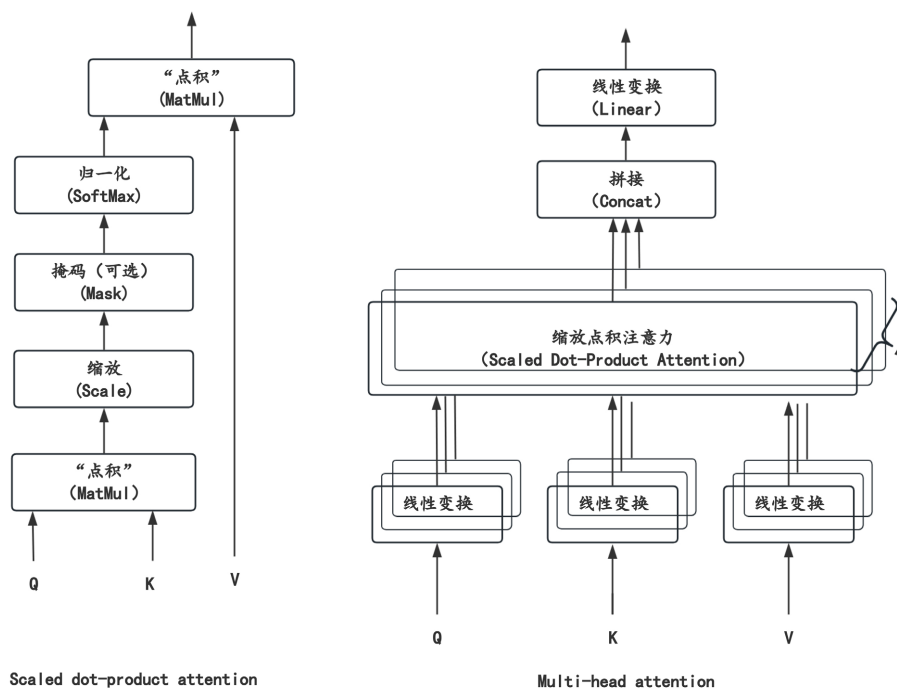


Figure 3. Diagram of multi-head attention mechanism

图 3. 多头注意力机制

4) 残差与层规范化

在多头注意力机制和前馈网络后面，都包含一个残差与规范化操作，残差机制通过将前一层的输入与输出相加实现，对 x 残差的计算公式为：

$$x_{\text{out}} = \text{LayerNorm}(G(x) + x) \quad (13)$$

式(13)中： x 为残差机制前一层的输入； $G(x)$ 为残差前一层的输出； x_{out} 表示对 x 进行残差操作。

5) 前馈网络

编解码器中的前馈神经网络模块(FFN)采用双层线性变换结构(式 14)，通过 ReLU 激活函数构建非线性映射关系，其数学表达式计算公式为：

$$F_{\text{fn}}(Z) = \text{ReLU}(ZW_1 + B_1)W_2 + B_2 \quad (14)$$

其中 W_1 、 W_2 和 B_1 、 B_2 分别对应两层网络的权重矩阵与偏置向量。

6) 解码器和输出机制

在多头注意力层、前馈网络层和掩码多头注意力层这三层的周围使用残差连接，掩码多头注意力层(防止未来信息泄露)、常规多头注意力层和前馈网络层，各层均采用残差连接和层归一化设计，最终通过线性变换输出预测的空气质量指数序列。

$$Y'_{n+1,2n} = (y'_{n+1}, y'_{n+2}, y'_{n+3}, \dots, y'_{2n})。$$

Table 2. Comparison of parameters between Xception model and other convolution models

表 2. Xception 模型和其他卷积模型的参数对比

模型	大小(MB)	Top-1 准确率	Top-5 准确率	参数数量	深度
Xception	88	0.790	0.945	22,910,480	126
VGG16	528	0.713	0.901	138,357,544	23
VGG19	549	0.713	0.900	143,667,240	26
ResNet50	99	0.749	0.921	25,636,712	168
InceptionV3	92	0.779	0.937	23,851,784	159
InceptionResNetV2	215	0.803	0.953	55,873,736	572
MobileNet	16	0.704	0.895	4,253,864	88
MobileNetV2	14	0.713	0.901	3,538,984	88
DenseNet121	33	0.750	0.923	8,062,504	121
DenseNet169	57	0.762	0.932	14,307,880	169

(2) 中间层：因果推断模块

混杂因子控制：构建因果图识别 6 类混杂因素。因果图是一种用于表示变量之间因果关系的图形模型，用于识别和分析潜在的混杂因素。在大气污染预测的场景中，构建因果图有助于理解污染物浓度、气象条件、排放源等多种因素之间的复杂关系，从而更准确地预测污染物浓度。其构建过程首先确定变量，其中目标变量有污染物浓度(如 PM2.5、NO₂ 等)；潜在原因变量有气象条件(温度、湿度、风速等)、排放源(工业排放、交通排放等)、地理因素(地形、城市布局等)。如图 4 所示：

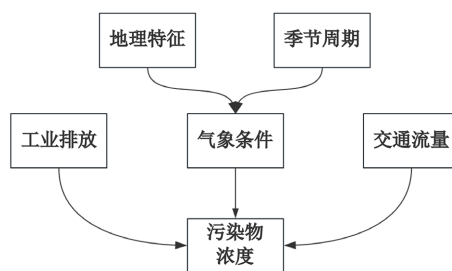


Figure 4. Causality diagram
图 4. 因果关系图

通过因果图分析,找出那些同时影响原因变量和目标变量,而混杂因素的控制通过将数据按照混杂因素的不同水平进行分层,然后在每一层内分析原因变量和目标变量的关系。找到具有相似混杂因素水平的个体或时间段进行匹配,以消除混杂因素的影响,在回归模型中加入混杂因素作为协变量,通过回归调整消除其影响[21]-[24]。

而双重稳健估计(Doubly Robust Estimation)是一种结合倾向性得分加权和回归调整的方法,为提高因果效应估计的稳健性。具体实现步骤如下:

首先进行倾向性得分估计:使用逻辑回归分类算法,根据混杂因素预测每个样本接受处理(如高排放源影响)的概率,即倾向性得分。倾向性得分反映了在给定混杂因素条件下,样本接受处理的可能性。

然后进行回归调整:建立回归模型,以污染物浓度为因变量,排放源原因变量和混杂因素为自变量,进行回归分析,通过回归模型预测在每个混杂因素水平下,原因变量对目标变量的平均影响。

接着进行双重稳健估计,结合倾向性得分加权和回归调整的结果,计算双重稳健估计量。双重稳健估计采用逆概率加权 + 回归调整方法[30] [31],消除混杂偏差,公式为:

$$ATE = \frac{1}{N} \sum_{i=1}^n \left[\frac{YT_i}{\hat{P}(X_i)} - \frac{Y(1-T_i)}{1-\hat{P}(X_i)} \right] \quad (23)$$

其中 $\hat{P}(x)$ 代表倾向性得分。

最后通过交叉验证方法评估双重稳健估计的稳健性和准确性[25]-[27]。

(3)顶层: 多任务学习模块

任务配置: 主任务: AQI 等级预测(输出维度 = 6, 对应优/良/轻度污染等); 辅助任务: 污染物浓度回归(PM2.5、NO₂、O₃、PM10、CO、SO₂六通道输出)。

其中损失函数公式为:

$$D_{KL}(p \parallel q) = \sum_{i=1}^n p(x_i) \log \left(\frac{p(x_i)}{q(x_i)} \right) \quad (24)$$

$P(x)$ 是真实分布, $Q(x)$ 是模型预测的分布。KL 散度衡量的是使用模型分布 $Q(x)$ 拟合真实分布 $P(x)$ 时产生的信息损耗[28] [29]。

2.3. 模型评价指标

对模型预测污染浓度数据的准确性进行评价,采用均方根误差(root mean square error, RMSE)和平均绝对误差(mean absolute error, MAE)评价预测的准确性,均方根误差 ERMSE 和平均绝对误差 EMAE 的计算公式分别为

$$E_{\text{RMSE}} = \sqrt{\frac{2}{n} \sum_{i=1}^n (y_i - y'_i)^2}, \quad (25)$$

$$E_{\text{MAE}} = \frac{1}{n} \sum_{i=1}^n |y_i - y'_i|. \quad (26)$$

式(25)~(26)中: y_i 为实际的污染浓度数据; y'_i 为预测的污染浓度数据。

2.4. 实验结果与分析

针对短期预测任务(4 h/24 h), AQI-Transformer 在污染物浓度预测精度上显著优于 Transformer、Informer 等主流时序模型, 对比实验的 RMSE 与 MAE 指标详见表 3, 验证了该模型在时间跨度适应性上的优势。

Table 3. Accuracy of AQI-Transformer and other models for power load forecasting

表 3. AQI-Transformer 与其他模型开展污染浓度预测的准确性

模型	RMSE 值		MAE 值	
	4 h	24 h	4 h	24 h
AQI-Transformer	1128.73	1252.13	833.93	975.18
Transformer	1420.21	2026.60	1048.64	1598.17
Informer	1290.99	1893.95	972.44	1543.28
Autoformer	5209.50	2441.57	3693.68	1826.33
Seq2seq	1844.93	1638.91	1287.57	1199.11
LSTM	1608.69	1580.64	1158.22	1122.49

由表 3 数据可见, 不同模型在 4 小时(短期)和 24 小时(中长期)预测任务中的表现差异显著, AQI-Transformer 在两项指标上全面领先, 其 4 小时 RMSE (1128.73)和 MAE (833.93)均显著低于其他模型, 24 小时预测的 RMSE (1252.13)和 MAE (975.18)也保持最优。传统 Transformer 表现较差, 尤其在 24 小时预测中, RMSE (2026.60)和 MAE (1598.17)均明显高于 AQI-Transformer。

综上可知, 跨时段预测对比实验表明, 相较于基准模型(如 Transformer、LSTM 等), AQI-Transformer 在 4 小时及 24 小时预测任务中均展现出显著的性能优势, 其 RMSE 与 MAE 指标全面领先。

2.5. 消融实验

为定量评估 AQI 相关系数模块与融合时间编码机制的关键作用, 本研究采用控制变量法系统性构建四组对照模型(基准模型/增强型模型 A/增强型模型 B/完整模型), 通过模块贡献度分析验证核心技术创新点的有效性。如下:

- 1) 基准模型(Transformer): 保留原始 Transformer 架构, 未引入污染物关联分析与时间编码模块;
- 2) 增强型模型 A (A-Transformer): 在基准模型基础上整合 AQI 污染物关联分析模块, 用于捕捉多监测点间的浓度相关性;
- 3) 增强型模型 B (T-Transformer): 在基准模型中嵌入融合时间定位编码模块, 强化时序特征表达;
- 4) 完整模型(AQI-Transformer): 协同融合上述两项技术模块的优化架构。

通过消融模型开展污染浓度预测的准确性如表 4 所示, 通过对比四组模型在污染浓度预测任务中的 RMSE 与 MAE 指标, 可定量评估各模块对预测性能的贡献度。

Table 4. Accuracy of AQI-Transformer ablation for power load forecasting
表 4. AQI-Transformer 消融模型开展污染浓度预测的准确性

模型	RMSE 值		MAE 值	
	4 h	24 h	4 h	24 h
AQI-Transformer	1129.93	1252.13	833.93	975.18
B-Transformer	1647.96	1609.35	1251.75	1243.82
A-Transformer	1155.12	1518.91	860.04	1172.20
Transformer	1420.21	2026.60	1048.64	1598.17

根据表 4 的数据分析，我们可以得出以下结论：

1) A-Transformer 模型的表现：在预测 4 小时污染浓度数据时，A-Transformer 的均方根误差(RMSE)和平均绝对误差(MAE)分别为 1,155.12 和 860.04，相较于 Transformer 模型，其 RMSE 和 MAE 分别减少了 18.67%和 17.99%。这表明，通过 AQI 污染浓度相关系数的筛选，A-Transformer 能够更精确地预测短期内的污染浓度。对于 24 小时的污染浓度预测，A-Transformer 的 RMSE 和 MAE 分别为 1518.91 和 1172.20，与 Transformer 相比，这两个指标分别降低了 25.05%和 27.42%。这进一步证实了 AQI 污染浓度相关系数筛选在提升中长期预测准确性方面的有效性。

2) B-Transformer 模型的优势：在 24 小时污染浓度预测任务中，B-Transformer 的 RMSE 和 MAE 分别为 1,609.35 和 1,243.82，与 Transformer 相比，其 RMSE 和 MAE 分别下降了 20.59%和 22.17%。这一改进归功于融合时间定位编码的引入，使得模型能够同时捕获污染浓度数据的局部位置信息和全局时间信息，从而提高了预测的准确性。

3) 消融实验表明(表 4 数据支撑)，AQI-Transformer 模型结合相关系数筛选与时间编码双模块时，其 RMSE/MAE 指标较单一模块模型(A-Transformer/QI-Transformer)提升显著，证明双技术协同可深度挖掘污染物时空演化规律，增强预测精度。

3. 结论

针对工业化进程中大气污染精准治理需求，本研究创新提出集成时空注意力机制与双重因果推断框架的 AQI-Transformer 复合模型，通过有机融合 Transformer、Autoformer 以及 FaaS 函数等架构优势，攻克多尺度时空特征融合与混杂因子干扰难题。实验验证该模型在污染物浓度预测中的 RMSE/MAE 指标显著优于主流时序模型，为智慧环保决策提供可靠技术支撑。

参考文献

- [1] 稳清, 姜琳琳, 朱元航, 等. 苏州市土地利用及生境质量变化预测[J]. 环境监测管理与技术, 2023, 35(6): 15-21.
- [2] 孔冬艳, 陈会广, 吴孔森. 中国“三生空间”演变特征、生态环境效应及其影响因素[J]. 自然资源学报, 2021, 36(5): 10-30.
- [3] 廖雨辰, 史雪威, 刘俊雁, 等. 九寨沟国家级自然保护区地震前后生态系统服务价值评估[J]. 生态学报, 2022, 42(6): 2063-2073.
- [4] 杨光宗, 吕凯, 李峰. 基于格网尺度的南昌市土地利用变化及生态系统服务价值时空相关性分析[J]. 中国土地科学, 2022, 36(8): 121-130.
- [5] 李姚旺, 张宁, 杜尔顺, 等. 基于碳排放流的电力系统低碳需求响应机制研究及效益分析[J]. 中国电机工程学报, 2022, 42(8): 2830-2842
- [6] Hmiel, B., Petrenko, V.V., Dyonisius, M.N., Buizert, C., Smith, A.M., Place, P.F., *et al.* (2020) Preindustrial ¹⁴CH₄ Indicates Greater Anthropogenic Fossil CH₄ Emissions. *Nature*, **578**, 409-412.

- <https://doi.org/10.1038/s41586-020-1991-8>
- [7] Hensen, A., Velzeboeri, Frumau, K.F.A., *et al.* (2019) Methane e-Mission Measurements of Offshore Oil and Gas Platforms: TNO Report R10895. TNO, 32-35.
 - [8] Denil, M., Shakibi, B., Dinh, L., *et al.* (2013) Predicting Parameters in Deep Learning. In: *Proceedings of the 26th International Conference on Neural Information Processing Systems*, Curran Associates Inc., 2148-2156.
 - [9] Girshick, R. (2015) Fast R-CNN. 2015 *IEEE International Conference on Computer Vision (ICCV)*, Santiago, 13-16 December 2015, 1440-1448. <https://doi.org/10.1109/iccv.2015.169>
 - [10] Ren, S., He, K., Girshick, R. and Sun, J. (2017) Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **39**, 1137-1149. <https://doi.org/10.1109/tpami.2016.2577031>
 - [11] Chen, Y., Huo, J., Liu, Y., Zeng, Z., Zhu, X., Chen, X., *et al.* (2020) Development of a Novel Flow Cytometry-Based Approach for Reticulocytes Micronucleus Test in Rat Peripheral Blood. *Journal of Applied Toxicology*, **41**, 595-606. <https://doi.org/10.1002/jat.4068>
 - [12] Saunio, M., Stavert, A.R., Poulter, B., *et al.* (2020) The Global Methane Budget 2000-2017. *Earth System Science Data*, **12**, 1561-1623.
 - [13] Bai, J., *et al.* (2023) Qwen Technical Report.
 - [14] Glm, T., *et al.* (2024) ChatGLM: A Family of Large Language Models from GLM-130b to GLM-4 All Tools.
 - [15] Rajpurkar, P., Chen, E., Banerjee, O. and Topol, E.J. (2022) AI in Health and Medicine. *Nature Medicine*, **28**, 31-38. <https://doi.org/10.1038/s41591-021-01614-0>
 - [16] Mao, Y., *et al.* (2024) A Survey on Lora of Large Language Models.
 - [17] Yang, A., *et al.* (2023) Baichuan 2: Open Large-Scale Language Models.
 - [18] Devlin, J. (2018) Bert: Pre-Training of Deep Bidirectional Transformers for Language Understanding.
 - [19] Yao, H., Liu, H. and Zhang, P. (2018) A Novel Sentence Similarity Model with Word Embedding Based on Convolutional Neural Network. *Concurrency and Computation: Practice and Experience*, **30**, 1-12. <https://doi.org/10.1002/cpe.4415>
 - [20] Tyner, D.R. and Johnson, M.R. (2021) Where the Methane Is—Insights from Novel Airborne LiDAR Measurements Combined with Ground Survey Data. *Environmental Science & Technology*, **55**, 9773-9783. <https://doi.org/10.1021/acs.est.1c01572>
 - [21] Zhou, H., Zhang, S., Peng, J., Zhang, S., Li, J., Xiong, H., *et al.* (2021) Informer: Beyond Efficient Transformer for Long Sequence Time-Series Forecasting. *Proceedings of the AAAI Conference on Artificial Intelligence*, **35**, 11106-11115. <https://doi.org/10.1609/aaai.v35i12.17325>
 - [22] Wu, H.X., Xu, J.H., Wang, J.M., *et al.* (2021) Autoformer: Decomposition Transformers with Auto-Correlation for Long-Term Series Forecasting. *Proceedings of the Annual Conference on Neural Information Processing Systems*, 6-14 December 2021, 22419-22430.
 - [23] Souza, F., Nogueira, R. and Lotufo, R. (2020) BERTimbau: Pretrained BERT Models for Brazilian Portuguese. In: Cerri, R. and Prati, R.C., Eds., *Intelligent Systems*, Springer International Publishing, 403-417. https://doi.org/10.1007/978-3-030-61377-8_28
 - [24] Plant, G., Kort, E.A., Floerchinger, C., Gvakharia, A., Vimont, I. and Sweeney, C. (2019) Large Fugitive Methane Emissions from Urban Centers along the U.S. East Coast. *Geophysical Research Letters*, **46**, 8500-8507. <https://doi.org/10.1029/2019gl082635>
 - [25] Sapci, A.H. and Sapci, H.A. (2020) Artificial Intelligence Education and Tools for Medical and Health Informatics Students: Systematic Review. *JMIR Medical Education*, **6**, e19285. <https://doi.org/10.2196/19285>
 - [26] Tsourounis, D., Kastaniotis, D., Theoharatos, C., Kazantzidis, A. and Economou, G. (2022) SIFT-CNN: When Convolutional Neural Networks Meet Dense SIFT Descriptors for Image and Sequence Classification. *Journal of Imaging*, **8**, 256-291. <https://doi.org/10.3390/jimaging8100256>
 - [27] Eyring, V., Collins, W.D., Gentine, P., Barnes, E.A., Barreiro, M., Beucler, T., *et al.* (2024) Pushing the Frontiers in Climate Modelling and Analysis with Machine Learning. *Nature Climate Change*, **14**, 916-928. <https://doi.org/10.1038/s41558-024-02095-y>
 - [28] Zhang, H., Ren, R., Gao, X., Wang, H., Jiang, W., Jiang, X., *et al.* (2025) Synchronous Monitoring Agricultural Water Qualities and Greenhouse Gas Emissions Based on Low-Cost Internet of Things and Intelligent Algorithms. *Water Research*, **268**, Article ID: 122663. <https://doi.org/10.1016/j.watres.2024.122663>
 - [29] Chen, S., Huang, J., Wang, P., Tang, X. and Zhang, Z. (2024) A Coupled Model to Improve River Water Quality Prediction towards Addressing Non-Stationarity and Data Limitation. *Water Research*, **248**, Article ID: 120895.

- <https://doi.org/10.1016/j.watres.2023.120895>
- [30] Liu, R., Zayed, T. and Xiao, R. (2024) Advanced Acoustic Leak Detection in Water Distribution Networks Using Integrated Generative Model. *Water Research*, **254**, Article ID: 121434. <https://doi.org/10.1016/j.watres.2024.121434>
- [31] Hamed, S.Z., Emami, H., Khayamzadeh, M., Rabiei, R., Aria, M., Akrami, M., *et al.* (2024) Application of Machine Learning in Breast Cancer Survival Prediction Using a Multimethod Approach. *Scientific Reports*, **14**, Article No. 30147. <https://doi.org/10.1038/s41598-024-81734-y>