

基于多变量LSTM与注意力机制的成都市空气质量指数预测研究

张 艳

广西师范大学数学与统计学院, 广西 桂林

收稿日期: 2026年6月24日; 录用日期: 2026年7月16日; 发布日期: 2026年7月29日

摘 要

针对成都市空气质量波动剧烈、极端污染频发且现有研究缺乏分层对比的现状, 本文基于2020~2025年逐日空气质量监测数据, 构建了单变量LSTM、多变量LSTM及融合注意力机制的多变量LSTM三种递进式预测模型, 并引入AQI > 150的极端污染专项误差评价。实验以21天为时间步长进行单步预测。结果表明: 多变量LSTM引入PM_{2.5}、PM₁₀、O₃、SO₂、CO、NO₂六项污染物后, MAE、RMSE、R²、MAPE均显著优于单变量模型; 融合注意力机制进一步将R²提升至0.7981、MAPE降至23.63%, 极端污染预测误差较单变量模型降低约25%。三种模型均有效收敛且无过拟合, 注意力机制在多变量基础上显著增强了对关键时段与峰值的捕捉能力。综合来看, 融合注意力机制的多变量LSTM模型在常规预测与极端污染场景中均表现出较好的适用性, 可为成都市空气质量预测与重污染天气应对提供一定的参考。

关键词

空气质量指数预测, 长短期记忆网络, 注意力机制, 多变量时间序列, 极端污染预警

Research on Air Quality Index Prediction in Chengdu Based on Multivariate LSTM with Attention Mechanism

Yan Zhang

School of Mathematics and Statistics, Guangxi Normal University, Guilin Guangxi

Received: June 24, 2026; accepted: July 16, 2026; published: July 29, 2026

Abstract

In response to the severe fluctuations of air quality and frequent extreme pollution events in Chengdu,

文章引用: 张艳. 基于多变量 LSTM 与注意力机制的成都市空气质量指数预测研究[J]. 统计学与应用, 2026, 15(7): 175-187. DOI: 10.12677/sa.2026.157160

as well as the lack of layered comparative studies in existing research, this paper constructs three progressive prediction models based on daily air quality monitoring data from 2020 to 2025: univariate LSTM, multivariate LSTM, and multivariate LSTM with attention mechanism. A specific error evaluation for extreme pollution ($AQI > 150$) is also introduced. The experiment uses a 21-day time step for one-step-ahead prediction. The results show that after introducing six pollutants ($PM_{2.5}$, PM_{10} , O_3 , SO_2 , CO , NO_2), the multivariate LSTM significantly outperforms the univariate LSTM in terms of MAE, RMSE, R^2 , and MAPE. The attention mechanism further improves R^2 to 0.7981 and reduces MAPE to 23.63%, while the extreme pollution prediction error decreases by about 25% compared with the univariate model. All three models converge effectively without overfitting, and the attention mechanism significantly enhances the ability to capture key periods and peaks on top of the multivariate model. Overall, the multivariate LSTM model with attention mechanism demonstrates promising applicability in both routine prediction and extreme pollution scenarios, and may serve as a useful reference for air quality forecasting and heavy pollution response in Chengdu.

Keywords

Air Quality Index Prediction, Long Short-Term Memory Network, Attention Mechanism, Multivariate Time Series, Extreme Pollution Warning

Copyright © 2026 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

伴随城市化与工业化快速推进,大气污染已成为全球性环境难题。空气污染严重危害人体呼吸系统与心血管健康,世界卫生组织统计显示,每年因室外空气污染导致的过早死亡人数达数百万人。精准的空气质量预报是开展公共健康防护与污染管控的重要技术支撑。空气质量指数(AQI)综合 $PM_{2.5}$ 、 PM_{10} 等六项污染物浓度,是衡量区域大气环境优劣的通用量化指标,其分级结果可直观反映污染等级与健康风险。

成都市地处四川盆地腹地,四面环山,大气扩散条件受限。秋冬季逆温频发、空气湿度偏高,极易促使气态污染物二次转化生成细颗粒物,造成污染物累积。虽然近年成都大气治理成效显著,全年空气质量优良率稳步提升,但受地形与季节性气象约束,短时污染反弹、AQI 突发性升高现象仍频繁出现,极端污染预警压力较大。

现有区域空气质量预测多聚焦模型整体精度优化,缺少单变量、多变量及注意力改进 LSTM 的分层对比实验,对盆地城市极端重污染场景的针对性验证不足。据此,本文依托成都市逐日环境监测数据,构建三类递进式 LSTM 预测模型,兼顾常规拟合精度与重污染($AQI > 150$)专项误差评价。从理论角度看,该研究丰富了盆地城市 AQI 时序预测的模型对比思路;从实践角度看,预测结果可作为成都市大气管控与健康预警的辅助参考,对西南同类型盆地城市的空气质量预报也有一定的参考价值。

1.2. 国内外研究现状

国外空气质量预测历经传统统计、机器学习、深度学习三个发展阶段。早期 ARIMA、多元线性回归等统计模型仅适配平稳时序,难以刻画大气污染非线性演化规律;随机森林等机器学习模型优化了多因子拟合能力,但对时序长期依赖捕捉有限。Mishra 和 Gupta 验证 LSTM 小时尺度 AQI 预报精度显著优于传统时序模型[1];为兼顾时空特征,Jin 等构建 GCN-LSTM 混合模型,依托空间关联特征有效降低 $PM_{2.5}$ 、

NO₂ 的预测误差, 当前国外研究逐步形成时序建模、空间耦合、注意力赋能、智能参数寻优的成熟研究范式[2]。

国内研究同样由 ARIMA、灰色预测等传统方法逐步转向深度学习建模, LSTM 凭借长时序记忆优势成为空气质量预测主流框架。胡彦军证实多变量 LSTM 较单变量模型误差更低、泛化性更强, 证明多特征输入可显著提升预测效果[3]; 曹天垚验证了多变量 LSTM 在多因子时序预测中的有效性[4]; 杨雨、刘小燕等研究证实, 注意力机制可自动聚焦突变时段关键信息, 有效降低重污染峰值预测误差, 为极端 AQI 预报提供新思路[5] [6]。

聚焦成都及四川盆地, 钟垚等通过多年站点监测数据证实, 本地 PM_{2.5}、PM₁₀ 呈冬高夏低的 U 型季节分布, 盆地静稳气象是污染物堆积的核心诱因[7]; 柳霄钰等揭示成都 PM_{2.5} 浓度与人口暴露风险高度耦合, 凸显精准预测与健康预警的现实必要性[8]; 郑梓玲基于 Attention-LSTM 实现成都 PM_{2.5} 浓度预报, 验证注意力结构对本地极端污染峰值的优化效果[9]。现有盆地相关研究多聚焦单一模型精度优化, 缺少单变量 LSTM→多变量 LSTM→注意力增强 LSTM 三层递进对比试验, 同时缺乏针对 AQI > 150 极端污染样本的专项误差评估。基于此, 本文搭建三类逐级改进模型, 从常规精度与极端预警双维度开展对比试验。

2. 相关理论基础

2.1. 长短期记忆网络(LSTM)

传统前馈神经网络难以处理时序依赖。循环神经网络(RNN)虽能保留历史信息, 但易出现梯度消失或爆炸问题。为此, Hochreiter 与 Schmidhuber 在 1997 年提出 LSTM [10], 通过引入细胞状态和三门结构(遗忘门、输入门、输出门)有效解决了长序列训练的梯度问题, 广泛用于空气质量等非线性时序预测。

LSTM 单元的计算过程如下:

遗忘门决定上一时刻细胞状态的保留程度:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f), \quad (1)$$

输入门控制新信息的写入:

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i), \quad (2)$$

候选单元状态:

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C). \quad (3)$$

细胞状态更新:

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t. \quad (4)$$

输出门决定当前隐藏状态:

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o), \quad (5)$$

$$h_t = o_t \cdot \tanh(C_t), \quad (6)$$

其中, h_{t-1} 为上一时刻的隐藏状态, x_t 为当前时刻输入, W_f 为权重矩阵, b_f 为偏置, σ 为 Sigmoid 函数, 输出范围[0, 1]。

LSTM 的门控机制使其能长期保留关键信息, 适合处理空气质量这类非平稳时序数据。在实际预测中, 引入 PM_{2.5}、PM₁₀、O₃、SO₂、CO、NO₂ 等多变量作为输入, 可使网络学习不同污染物间的交叉依赖, 进一步提升预测精度。

2.2. 注意力机制

注意力机制最初由 Bahdanau 等提出用于机器翻译任务,其核心是让模型动态聚焦输入序列中的重要部分[11]。在 AQI 预测中,不同时刻的污染物浓度对未来 AQI 影响不同,注意力机制可为每个时间步学习权重,强化对峰值、突变点和极端污染的捕捉。

给定 LSTM 输出的隐状态序列 $H = \{h_1, h_2, \dots, h_T\}$, 注意力机制的计算分为三步:

1. 计算对齐分数: 定义一个评分函数 $\text{score}(h_i, q)$, 其中 q 为查询向量(本文取最后时间步隐状态或可学习全局向量), h_i 为第 i 个输入。

2. 归一化: 将对齐分数通过 Softmax 函数转换为概率分布的权重:

$$a_i = \frac{\exp(\text{score}(h_i, q))}{\sum_{j=1}^T \exp(\text{score}(h_j, q))}. \quad (7)$$

3. 加权求和: 将权重与输入进行加权求和, 得到上下文向量:

$$c = \sum_{i=1}^T a_i h_i. \quad (8)$$

在 LSTM+ 注意力模型中, 注意力层连接在 LSTM 层之后, 对全时间步隐状态加权求和生成上下文向量, 再输入全连接层进行回归预测。相比传统 LSTM 仅用最后时间步或简单平均, 注意力机制通过自适应加权提升了模型对峰值和转折点的预测能力。

2.3. 评价指标

为全面评估模型预测效果, 本文采用常规指标与极端污染专项指标相结合的方式。常规指标包括:

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|, \quad (9)$$

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}, \quad (10)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}, \quad (11)$$

$$\text{MAPE} = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \times 100\%. \quad (12)$$

其中, MAE 反映平均预测偏差, RMSE 对较大误差赋予更高惩罚, R^2 衡量模型对方差的解释程度, MAPE 反映相对误差比例。此外, 为评估模型对重污染天气(AQI > 150)的预警能力, 本文在测试集中单独计算 AQI > 150 样本上的 MAE 和 RMSE (称为极端 MAE、极端 RMSE), 公式与常规指标一致。通过常规指标与极端指标的组合评价, 可全面反映模型在日常预测与重污染预警两方面的真实能力。

3. 数据预处理与实验设计

3.1. 数据来源与数据预处理

本文以成都市为研究区, 数据源自 GBQ 大数据平台(<https://citydev.gbqyun.com/index/chengdu>)发布的空气质量监测数据集, 包含 2020 年 1 月 1 日至 2025 年 12 月 31 日共 2192 条日尺度记录。选取 AQI、PM_{2.5}、PM₁₀、O₃、SO₂、NO₂、CO 七项指标, 其中 AQI 为预测目标, 其余为辅助特征。单变量 LSTM 仅基于 AQI 建模; 多变量及注意力改进模型联合全部特征, 挖掘污染物间的耦合规律。原始数据缺失值占

比低于 1%，采用线性插值法填补，以保持序列连续性。经清洗后，全部记录可用于模型训练与评估。

3.2. 数据基本统计特征分析

3.2.1. AQI 时间序列变化分析

如图 1 所示，2020~2025 年成都市 AQI 序列呈现显著的“双峰型”季节波动：冬季 AQI 偏高，污染累积明显；夏季 AQI 整体偏低，曲线较平缓。这一规律与四川盆地冬季静稳天气、扩散条件差密切相关。此外，序列中存在多次超过 250 的突发性 AQI 高值，对应中度至重度污染过程。整体上，序列表现出非平稳特征，均值与方差随时间变化，传统线性模型难以有效拟合，凸显了采用 LSTM 等深度学习模型进行预测的必要性。

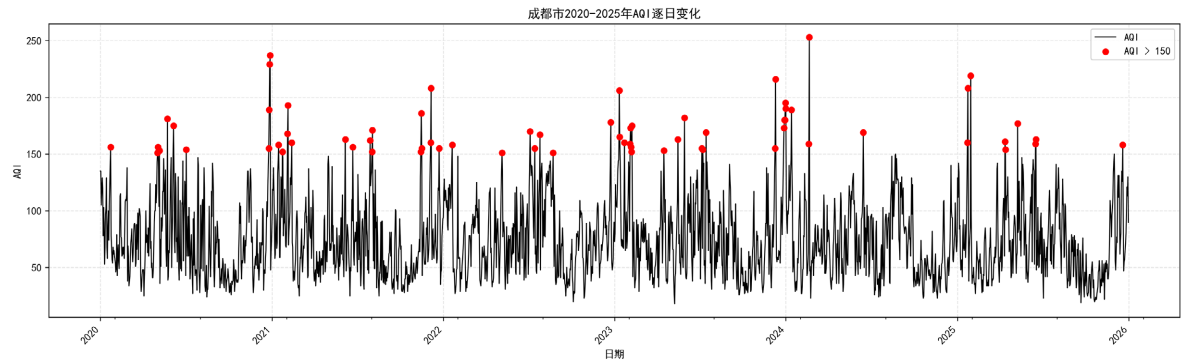


Figure 1. Daily AQI variation curve of Chengdu from 2020 to 2025
图 1. 成都市 2020~2025 年 AQI 逐日变化曲线

3.2.2. 各指标箱线图分布分析

箱线图可直观反映数据的中位数、四分位数及异常值分布。如图 2 所示，各指标在分轴坐标系下展现出清晰的内部特征：AQI 中位数约为 70，箱体主要集中在 50~100 区间，表明大部分时段空气质量处于优良水平，但上界外存在大量极端高值点，对应重污染天气；PM_{2.5} 与 PM₁₀ 的箱体分布与 AQI 高度同步，异常值数量多且数值高，说明颗粒物是影响 AQI 波动的主要因素；O₃ 箱体跨度较大，反映出夏季臭氧污染对空气质量的显著影响；SO₂、CO 与 NO₂ 箱体整体偏低、分布紧凑，极端值较少，表明成都市硫化物与一氧化碳排放控制效果较好，对 AQI 的整体贡献有限。

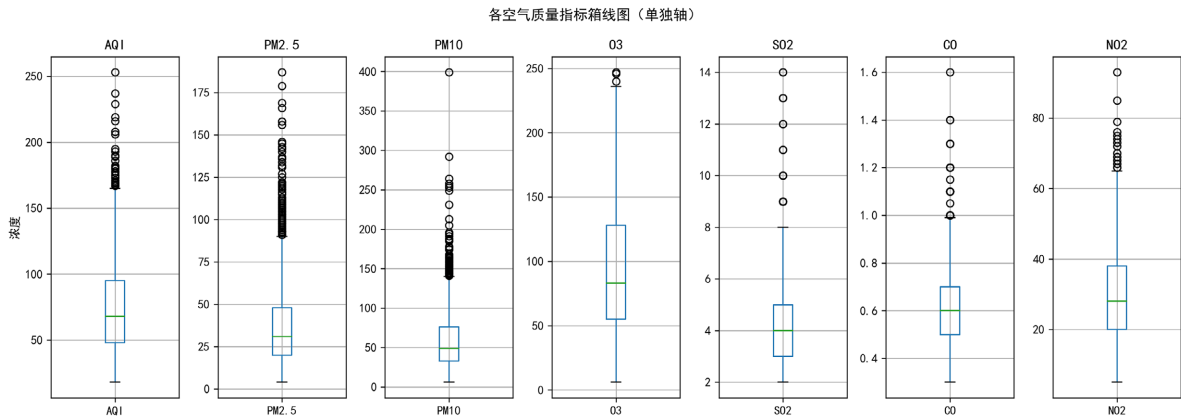


Figure 2. Axis-separated boxplots of air quality indicators
图 2. 各空气质量指标分轴箱线图

3.2.3. 污染物相关性分析

相关性热力图 3 显示，AQI 与 PM_{2.5}、PM₁₀ 的相关系数分别为 0.67 和 0.68，呈显著正相关，表明颗粒物是影响 AQI 的主要因素；PM_{2.5} 与 PM₁₀ 相关系数高达 0.92，说明二者来源与变化高度一致；O₃ 与颗粒物呈弱负相关，反映臭氧与颗粒物污染的季节错峰特征；CO、NO₂ 与 PM_{2.5}、PM₁₀ 的相关系数均高于 0.6，说明机动车尾气等排放源对多种污染物具有协同贡献。以上结果验证了多变量建模的必要性，仅用 AQI 单变量难以捕捉多污染物协同影响，引入多污染物特征有助于提升预测精度。

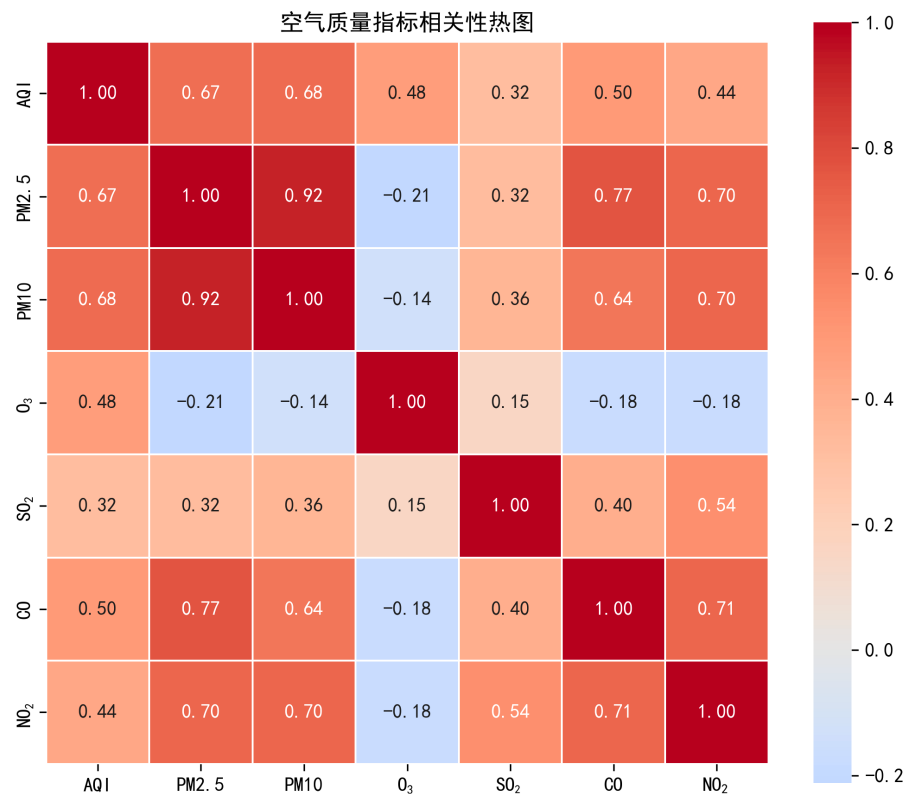


Figure 3. Correlation heatmap of air quality indicators
图 3. 空气质量指标相关性热力图

3.2.4. 成都市空气质量指标描述性统计

Table 1. Descriptive statistical results of air quality indicators in Chengdu
表 1. 成都市空气质量指标描述性统计结果

	AQI	PM _{2.5}	PM ₁₀	O ₃	SO ₂	CO	NO ₂
mean	74.99	37.53	58.41	93.91	4.44	0.65	29.87
std	34.56	25.91	35.69	49.44	1.7	0.17	13.18
min	18	4	6	6	2	0.3	5
25%	48	20	33	55	3	0.5	20
50%	68	31	49	83	4	0.6	28
75%	95	48	76	128	5	0.7	38
max	253	187	399	247	14	1.6	93

由表 1 可知, 研究期内成都市 AQI 均值为 74.99, 最小 18.00、最大 253.00, 标准差 34.56, 波动大、稳定性差, 呈现非线性、非平稳特征, 与盆地污染易累积特点一致。50%与 75%分位数为 68.00 和 95.00, 多数时段优良但仍有部分接近轻度污染, 凸显精准预警必要性。

PM_{2.5}与 PM₁₀ 均值分别为 37.53 μg/m³、58.41 μg/m³, 标准差显著高于其他污染物, 是影响 AQI 的首要污染物; O₃ 均值 93.91 μg/m³, 最大值 247.00, 标准差 49.44, 反映夏季臭氧快速累积; SO₂ (4.44 μg/m³) 与 CO (0.65 mg/m³) 均值低、波动小, 工业与燃煤控制良好; NO₂ 均值 29.87 μg/m³, 标准差 13.18, 波动主要来自交通排放。各指标最大值远高于均值, 存在重污染极值样本, 为设置 AQI > 150 的极端评价提供数据基础, 验证了模型兼顾常规与极端预警的必要性。

3.3. 特征选择与归一化

为评估模型在空气质量预测中的适用性, 本研究设计了三种递进模型: ① 单变量 LSTM (仅以 AQI 历史值为输入); ② 多变量 LSTM (输入 AQI 及 PM_{2.5}、PM₁₀、O₃、SO₂、CO、NO₂ 六项污染物); ③ 多变量 LSTM + 注意力机制(在前者基础上对历史关键时间步自适应加权)。三种模型采用相同的数据划分, 通过递进对比量化多变量输入和注意力机制的贡献。

由于各污染物量纲不同、数值范围差异显著(如 CO 常低于 2 mg/m³, PM_{2.5} 可达数百 μg/m³), 直接使用原始数值会导致大数值特征主导损失函数, 影响训练稳定性与收敛速度。为此, 采用 Min-Max 归一化将数据线性变换至[0, 1]区间, 公式如式(13)所示:

$$x' = \frac{x - x_{\min}}{x_{\max} - x_{\min}}, \quad (13)$$

其中, $x_{\min}$ 和 $x_{\max}$ 分别为训练集上相应特征的最小值和最大值, 测试集使用同一组极值, 避免数据泄露。归一化使各特征处于同一数量级, 提升训练稳定性与收敛效率。

3.4. 滑动窗口样本构造

LSTM 模型要求输入为固定长度序列, 本研究采用滑动窗口法构造监督学习样本。设窗口长度为 L , 特征维度为 d (单变量为 1, 多变量为 7), 则第 t 个样本由连续 L 个时间步的特征构成, 标签为后续时刻的 AQI 值, 如式(14)所示:

$$\begin{aligned} X_t &= \{x_{t-L+1}, x_{t-L+2}, \dots, x_t\}, \\ y_t &= \text{AQI}_{t+1}, \end{aligned} \quad (14)$$

结合成都市空气污染过程通常持续 2~3 周, 本文设置 $L = 21$ 天, 21 天窗口可覆盖一次重污染全过程; 初步实验也表明该窗口下模型验证性能最优。滑动步长设为 1, 以最大化利用样本量。

3.5. 数据集划分

为客观评估模型泛化能力, 按时间顺序将样本集划分为训练集(85%)、验证集(10%)、测试集(5%), 避免未来信息泄露, 符合真实预测场景。样本总数为 $N = 2171$, 各子集样本数如式(15)所示:

$$\begin{aligned} N_{\text{train}} &= \lfloor N \times 85\% \rfloor = 1845 \\ N_{\text{val}} &= \lfloor N \times 10\% \rfloor = 217 \\ N_{\text{test}} &= N - N_{\text{train}} - N_{\text{val}} = 109 \end{aligned} \quad (15)$$

训练集位于序列前部, 用于学习模型参数; 验证集紧随其后, 不参与训练过程, 仅用于早停监控, 当验证损失连续 12 轮未下降时, 将终止训练并恢复至最佳权重, 以此防止过拟合; 测试集位于序列末尾,

仅用于对模型最终泛化能力进行独立评估。该按时间顺序划分的策略，模拟了“利用过去预测未来”的实际应用模式，保障了实验的可靠性与可重复性。

4. 模型构建

4.1. 单变量 LSTM 模型构建

单变量 LSTM 模型仅以历史 AQI 序列为输入，作为基准预测模型。网络结构有两层 LSTM，其中第一层 128 单元、第二层 64 单元，均使用 tanh 激活、两层 Dropout (比例 0.15)及两层全连接层(32 个神经元 + ReLU，输出节点线性激活)组成。输入为标准化后的 21 天时序片段(维度 21×1)，输出未来一日 AQI 预测值。训练采用 MSE 损失、Adam 优化器，批次大小 8，最大轮次 100，配合早停及学习率衰减，详细参数配置见表 2。

Table 2. Parameter configuration of the univariate LSTM model
表 2. 单变量 LSTM 模型参数配置

层级	参数	参数值
输入层	输入节点数	1
LSTM 层 1	隐藏单元数目	128
	激活函数	tanh
Dropout 层 1	Dropout 比例	0.15
LSTM 层 2	隐藏单元数目	64
	激活函数	tanh
Dropout 层 2	Dropout 比例	0.15
全连接层	神经元数目	32
	激活函数	ReLU
输出层	输出节点数	1
训练参数	时间步长	21
	优化器	Adam
	损失函数	MSE
	最大迭代次数	100
	批次大小	8
	早停轮数	12

注：LSTM 层默认激活函数为 tanh，训练过程中使用学习率衰减策略。

4.2. 多变量 LSTM 模型构建

在单变量 LSTM 基础上，本文构建多变量 LSTM 模型，同时引入 PM_{2.5}、PM₁₀、O₃、SO₂、CO、NO₂ 六项污染物作为辅助特征，与 AQI 共同构成多维度输入，以学习污染物间的内在关联，提升预测精度与稳定性。

为保证对比公平，模型网络结构、层数、神经元数量、激活函数、正则化等与单变量 LSTM 完全一致，仅将输入维度调整为 21×7 ，即七项指标，21 天窗口。通过多特征融合，模型可降低单一变量波动干扰，尤其在 AQI 快速变化阶段提升明显。训练配置(批次大小、优化器、损失函数、Dropout、早停、学习率衰减)亦与单变量模型保持一致，确保性能差异真实反映多特征输入的贡献，模型参数配置见表 3。

Table 3. Parameter configuration of the multivariate LSTM model
表 3. 多变量 LSTM 模型参数配置

层级	参数	参数值
输入层	输入节点数	7

注：表中未列出的网络层与训练参数均与表 2 完全相同。

4.3. 融合注意力机制的多变量 LSTM 模型构建

为进一步提升关键时序信息的捕捉能力，本文在多变量 LSTM 基础上引入注意力机制，让模型自主学习不同时间步的重要性，提高关键时段权重、降低噪声干扰。模型保留两层 LSTM 结构，在第二层 LSTM 输出后加入注意力模块：计算各时间步得分，经 Softmax 归一化后加权求和生成全局特征向量，聚焦于 AQI 快速上升和污染高峰等关键阶段，再经全连接层输出预测结果。训练配置与前两种模型一致(Adam、MSE、批次 8、Dropout 0.15、早停及学习率衰减)，保证对比公平。注意力模块在不显著增加训练负担下提升了关键时段特征提取能力，适用于成都市 AQI 波动剧烈、极端污染频发的预测场景，模型参数配置见表 4。

Table 4. Parameter configuration of the multivariate LSTM model with attention mechanism
表 4. 多变量 LSTM + 注意力机制模型参数配置

层级	参数	参数值
输入层	输入节点数	7
注意力层	权重激活函数	Softmax

注：表中未列出的网络层与训练参数均与表 2 完全相同。

5. 实验结果与分析

5.1. 模型训练过程与收敛性分析

模型训练过程中的损失曲线能够直观反映网络的学习效率、收敛速度与泛化能力，是判断模型拟合效果的重要依据。本章通过训练损失与验证损失的变化趋势，对三种模型的收敛性能进行对比分析，结果如图 4~6 所示。

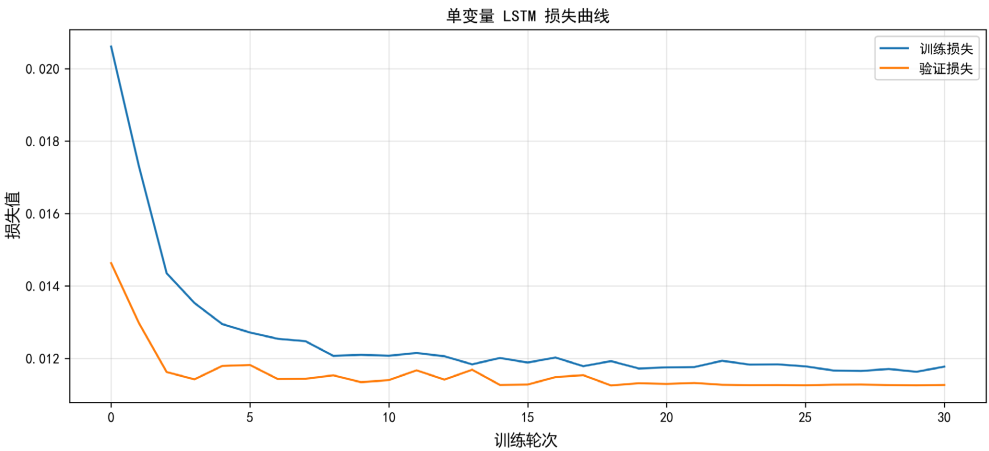


Figure 4. Loss curves of the univariate LSTM model
图 4. 单变量 LSTM 模型损失曲线图

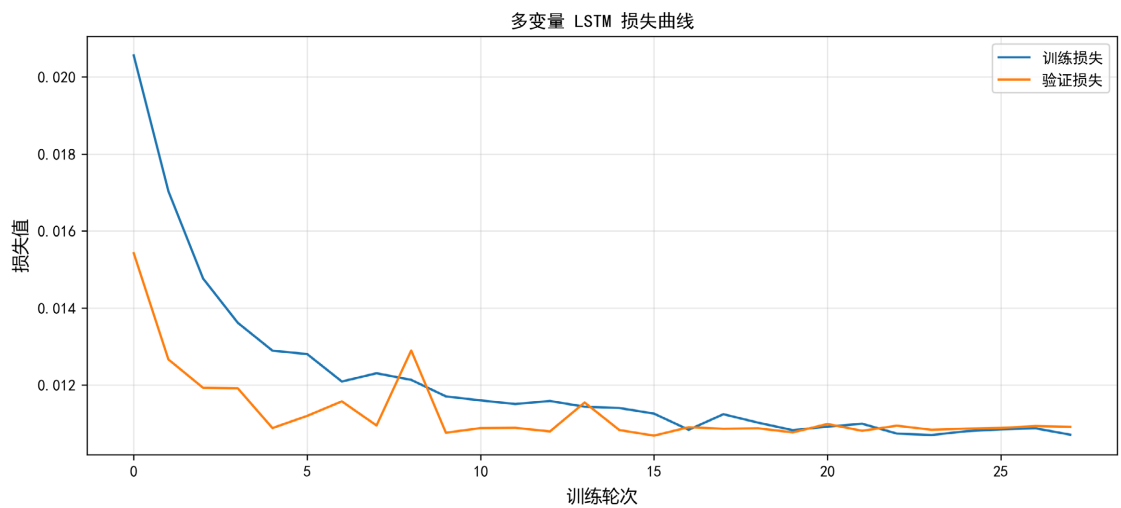


Figure 5. Loss curves of the multivariate LSTM model
图 5. 多变量 LSTM 模型损失曲线图

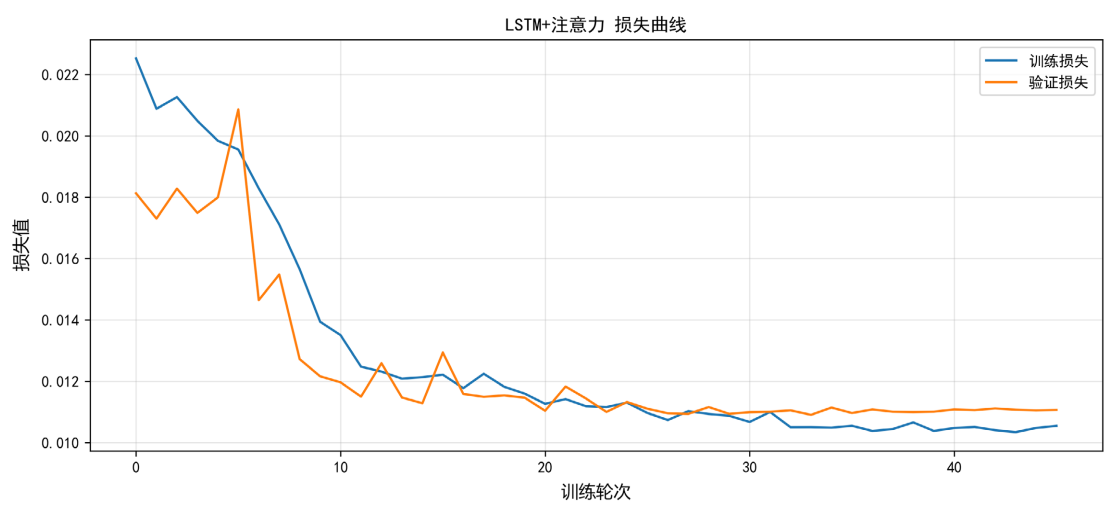


Figure 6. Loss curves of the multivariate LSTM model with attention mechanism
图 6. 多变量 LSTM + 注意力机制模型损失曲线图

单变量 LSTM 训练、验证损失前期快速回落，后期趋于重合平稳，Dropout 与早停策略有效规避过拟合，但模型最终损失偏高，单一 AQI 输入无法完整刻画污染物耦合变化规律。多变量 LSTM 凭借多污染物输入特征，收敛速度更快，稳态损失与曲线平稳度均优于单变量模型。注意力 LSTM 仅训练初期验证损失小幅震荡，后续持续下降并取得三类模型中最低损失，注意力结构筛选关键时序信息、弱化无效数据干扰。三类模型全部完成有效收敛，综合收敛速率与最终损失，注意力改进多变量 LSTM 综合性能最优。

5.2. 模型预测性能对比分析

5.2.1. 评价指标结果对比

为全面客观地衡量模型在测试集上的预测效果，采用平均绝对误差(MAE)、均方根误差(RMSE)、决定系数(R^2)与平均绝对百分比误差(MAPE)四项指标进行综合评估，三种模型的计算结果如表 5 所示。

Table 5. Overall performance comparison of the three models
表 5. 三种模型整体性能对比

模型	MAE	RMSE	R ²	MAPE
单变量 LSTM	14.08	18.34	0.7228	30.13%
多变量 LSTM	12.48	16.45	0.777	27.73%
LSTM + 注意力机制	11.74	15.65	0.7981	23.63%

由表可见，三种模型的预测精度呈逐步提升趋势。单变量 LSTM 仅依靠历史 AQI 序列，MAE 为 14.08，RMSE 为 18.34，R² 为 0.7228，MAPE 为 30.13%，误差较高，说明单一时间序列难以完整描述空气质量的非线性变化。多变量 LSTM 引入六项污染物特征后，各项指标明显改善：MAE 降至 12.48，RMSE 降至 16.45，R² 提升至 0.7770，MAPE 降至 27.73%，证明多源环境特征能有效增强模型对影响因素的捕捉能力。融合注意力机制的多变量 LSTM 表现最优，MAE 进一步降至 11.74，RMSE 降至 15.65，R² 提升至 0.7981，MAPE 降至 23.63%，在误差控制与拟合程度上均达最佳，表明注意力机制可有效提升关键时序特征权重，使预测更接近真实值。

5.2.2. 预测曲线对比分析

为更直观地展示模型对 AQI 变化趋势的拟合效果，利用测试集 109 个样本的真实值与预测值绘制对比曲线，三种模型的对比结果如图 7~9 所示。

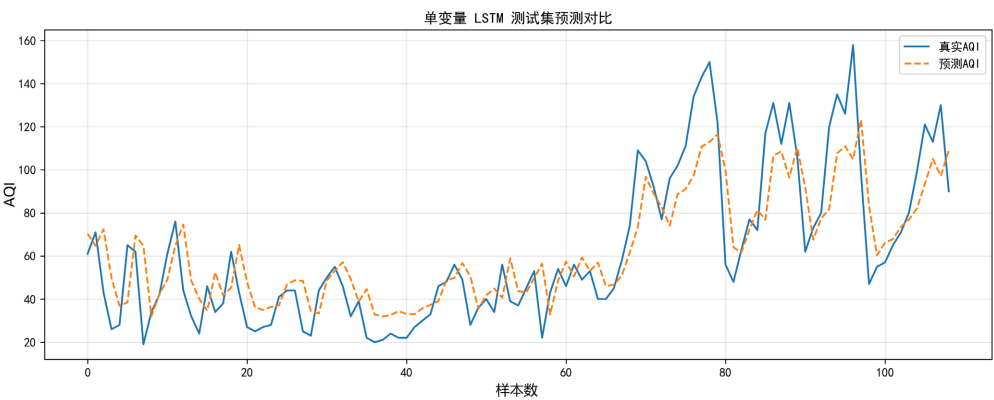


Figure 7. Prediction comparison on the test set of the univariate LSTM model
图 7. 单变量 LSTM 模型测试集预测对比图

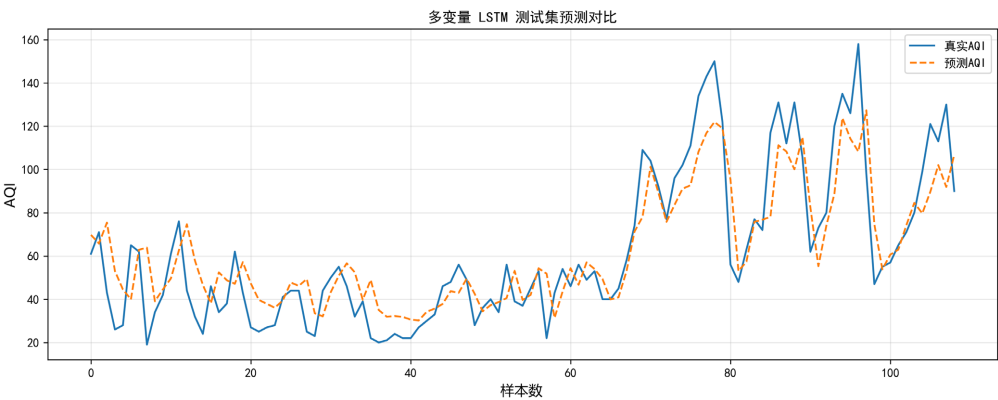


Figure 8. Prediction comparison on the test set of the multivariate LSTM model
图 8. 多变量 LSTM 模型测试集预测对比图

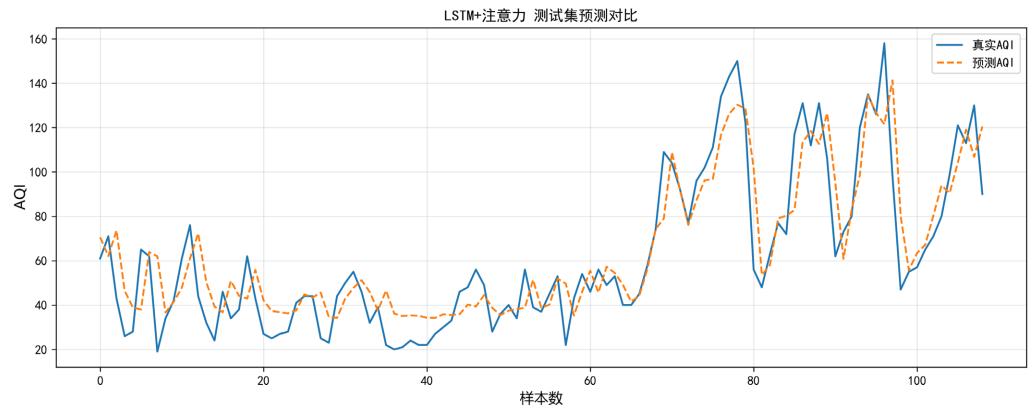


Figure 9. Prediction comparison on the test set of the multivariate LSTM model with attention mechanism
图 9. 多变量 LSTM + 注意力机制模型测试集预测对比图

从曲线对比结果可以看出，单变量 LSTM 模型能够大致捕捉 AQI 的整体变化趋势，但在局部波动、拐点及峰值附近存在明显滞后与偏差，对突变点的跟踪能力较弱。多变量 LSTM 模型的预测曲线与真实曲线贴合程度更高，在平稳变化区间的误差显著降低，对趋势转折的响应更为及时，但在极端高值附近仍存在一定偏差。融合注意力机制的多变量 LSTM 模型预测曲线与真实曲线一致性最高，不仅在常规区间拟合效果良好，在 AQI 快速上升或下降的突变阶段也能保持较高精度，对峰值的跟踪能力明显强于前两种模型，整体预测效果更为稳定可靠。

5.3. 极端污染天气预测效果分析

在实际空气质量预警工作中，中度及以上污染天气的准确预测具有重要的应用价值。因此，本章以 AQI > 150 作为极端污染阈值，单独统计三种模型在该类样本上的预测误差，结果如表 6 所示。

Table 6. Extreme pollution prediction performance comparison of the three models
表 6. 三种模型极端污染预测性能对比

模型	极端值样本数(AQI > 150)	极端值 MAE	极端值 RMSE
单变量 LSTM	1	48.38	48.38
多变量 LSTM	1	43.09	43.09
LSTM + 注意力	1	36.11	36.11

由上表可知，单变量 LSTM 对该重污染样本的预测误差达到 48.38，说明仅靠 AQI 历史值难以捕捉突发性高污染事件。引入多变量信息后，预测误差降至 43.09，相对降低了 10.9%。加入注意力机制后，预测误差进一步下降至 36.11，相比单变量模型降低了 25.4%，相比多变量模型降低了 16.2%。这一结果有力证明了注意力机制在提升极端值预测能力方面的显著作用，对于重污染预警具有明确的现实意义。

5.4. 模型综合对比与分析

综合收敛性能、预测精度、极端值预测及可视化对比可知：第一，多变量 LSTM 显著优于单变量 LSTM，引入多项污染物特征可充分挖掘污染物间的耦合关系，大幅提升预测准确性。第二，加入注意力机制后，模型在误差、拟合曲线及极端值预测上进一步提升，尤其在 AQI 波动剧烈和峰值预测方面优势突出。第三，融合注意力机制的多变量 LSTM 模型整体性能最优、稳定性最强、泛化能力最好，最适用于成都市空气质量预测与污染预警任务。

6. 结论与展望

6.1. 研究结论

本文以成都市为研究区域，基于 2020~2025 年逐日空气质量监测数据，构建了单变量 LSTM、多变量 LSTM 及融合注意力机制的多变量 LSTM 三种预测模型，通过多维度实验对比得到以下结论：数据预处理与统计分析表明成都市空气质量呈明显季节性与非平稳特征，AQI 波动大且污染物间耦合显著，为多变量建模提供了依据；三种模型均有效收敛，无过拟合，多变量 LSTM 收敛效率优于单变量模型，融合注意力机制的模型最终损失最低、泛化性能最优；常规预测中多变量 LSTM 的 MAE、RMSE、 R^2 、MAPE 均显著优于单变量模型，融合注意力机制进一步将 R^2 提升至 0.7981、MAPE 降至 23.63%；极端污染(AQI>150)预测中，单变量模型偏差较大，多变量模型有所改善，融合注意力机制的极端值误差较单变量模型降低约 25%，在重污染预警中优势明显；综合来看，融合注意力机制的多变量 LSTM 模型在常规与极端污染场景中均表现出相对较好的预测效果，在本文所对比的三种模型中具有一定优势，可为成都市空气质量预测与污染预警提供一定的参考。

6.2. 研究不足与展望

尽管本文取得一定研究成果，但仍存在以下不足：数据方面，仅采用常规污染物监测数据，未引入气象、交通、工业等外部因素；模型方面，仅使用 LSTM 及其改进模型，未来可引入 CNN、Transformer、GNN 或多模型融合框架；预测任务方面，仅实现单步 AQI 预测，未来可拓展至多步或多站点联合预测；极端样本方面，数据集中极端污染样本占比较低，测试集仅含 1 个 AQI>150 的样本，统计稳健性不足，后续可采用数据增强、过采样或收集更长时间跨度的数据，以更可靠地评估模型的重污染预警能力。

参考文献

- [1] Mishra, A. and Gupta, Y. (2024) Comparative Analysis of Air Quality Index Prediction Using Deep Learning Algorithms. *Spatial Information Research*, **32**, 63-72. <https://doi.org/10.1007/s41324-023-00541-1>
- [2] Jin, Y.M., Ren, G., Hu, Y.X., Wang, W.N. and Zhang, J.T. (2024) A Study on Air Quality Prediction with Multiple Features Based on GCN-LSTM. *Journal of Physics: Conference Series*, **2816**, Article 012074. <https://doi.org/10.1088/1742-6596/2816/1/012074>
- [3] 胡彦军. 基于多变量 LSTM 模型的大蒜种植面积预测研究[J]. 河南科技学院学报(自然科学版), 2026, 54(1): 75-83.
- [4] 曹天焱. 基于多变量 LSTM 的配电网调度预测[J]. 电子元器件与信息技术, 2026, 10(1): 32-35.
- [5] 杨雨, 王永千, 胥娇, 等. 融合多头自注意力机制和 LSTM 的风电塔筒倾角预测模型[J]. 风机技术, 2026, 68(2): 49-56.
- [6] 刘小燕, 邵长虹, 李瑞, 等. 融合 PMV 物理方程和 Attention-LSTM 神经网络的铁路客站旅客舒适度模型研究[J]. 中国铁路, 2024(5): 16-24.
- [7] 钟垚, 曾胜兰, 宋雨润. 四川盆地大气污染物质量浓度时空变化特征[J]. 成都信息工程大学学报, 2023, 38(1): 107-115.
- [8] 柳霄钰, 梁蓝元, 何雨璐, 等. 四川盆地 PM_{2.5} 浓度与人口暴露水平的时空变化特征[J/OL]. 地球环境学报, 2026: 1-17. <https://link.cnki.net/urlid/61.1482.x.20260401.1359.008>, 2026-05-22.
- [9] 郑梓玲. 基于 Attention-LSTM 模型的 PM_{2.5} 浓度预测: 以成都市为例[D]: [硕士学位论文]. 广州: 暨南大学, 2023.
- [10] Hochreiter, S. and Schmidhuber, J. (1997) Long Short-Term Memory. *Neural Computation*, **9**, 1735-1780. <https://doi.org/10.1162/neco.1997.9.8.1735>
- [11] Bahdanau, D., Cho, K. and Bengio, Y. (2015) Neural Machine Translation by Jointly Learning to Align and Translate. arXiv: 1409.0473. <https://arxiv.org/abs/1409.0473>