

用于空气质量预测的集成时间序列模型

乔慧梁*, 韩素芳#, 吴梓建

云南民族大学数学与计算机科学学院, 云南 昆明

收稿日期: 2025年9月28日; 录用日期: 2025年10月22日; 发布日期: 2025年10月29日

摘要

空气质量指数(AQI)是衡量空气质量的关键指标, 具有随机性和非平稳性, 构建一个合理的预测系统对整体社会有一定意义。本文整合集成学习策略, 提出一种新型的用于提高模型的预测精度和稳健性的集成模型框架。通过对中国云南省的空气质量指数的预测, 从而证明基于堆叠随机森林时间序列的预测模型实现了对比传统的LSTM、GRU和Transformer模型的性能较为显著的稳健性提升和精度提升。在对云南省共计13个样本点的预测中, 本文模型平均R2值达到0.801, 表明了模型在数值预测方面有较高的准确性。

关键词

深度学习, 集成学习, 空气质量指数

Ensemble Time Series Forecasting Model for Air Quality Prediction

Huiliang Qiao*, Sufang Han#, Zijian Wu

School of Mathematics & Computer Science, Yunnan Minzu University, Kunming Yunnan

Received: September 28, 2025; accepted: October 22, 2025; published: October 29, 2025

Abstract

The air quality index (AQI) serves as a crucial metric for gauging air quality. Given the intrinsic randomness and non-stationarity within, constructing a rational analysis-prediction system for AQI remains a meaningful task. This study minimizing Integrating ensemble learning strategies and effectively improving both the prediction accuracy and the robustness of the model, is applied to air quality monitoring in Yunnan Province, China. Experimental results demonstrate that the proposed based on stacked random forest time series prediction model achieves significant accuracy while

*第一作者。

#通讯作者。

outperforming traditional LSTM, GRU, and Transformer or some common other models that used for AQI prediction. In the prediction of a total of 13 sample points in Yunnan Province, the evaluation report indicates that the model's average coefficient of determination R^2 reaches 0.801. On this basis, its high accuracy in numerical prediction is demonstrated.

Keywords

Deep Learning, Ensemble Learning, Air Quality Index

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

空气质量指数(AQI)为一个无量纲指数,通过简化常规空气监测中的几种空气污染物,尤其是颗粒物、臭氧和氮氧化物等浓度,通过不同权重将其转化为一个单一的概念性指数值用以描述当前的空气质量状况。在以往的研究中,可以实证的是当上述污染物超过正常环境水平,会对人类健康、生态系统和环境造成一定量的损害。因此,悉心监测空气质量信息,以便直观、简洁地向公众传达空气质量信息至关重要。

在过往的研究中,出现了多种被用于预测各种常见的大气污染物指标的浓度的模型,大致可分为化学传输模型[1]-[3]、统计模型[4]-[8]以及机器学习模型[9]-[12]三类。其中,随着算力的发展,诸如 CNN [13]、LSTM [14]-[16]和 Transformer [17]等通过分析历史数据来预测未来的数据的深度学习模型在预测空气污染水平领域的普遍应用也证明了其有效性。

随着研究的深入以及算力设备的更迭,传统的单个模型在时序预测的精度和稳健性上开始暴露出短缺,同时集成模型(EL 模型)则展示出更强的表现,在各式情况下取得了相较于单个模型更为卓越的性能[18]-[21],本文认为以往作者过于关注模型本身性能,导致未能充分挖掘综合 AQI 指数相关数据的潜力,如地理位置,大气压力和气候相关信息。本文通过收集到的关于空气报告的丰富数据,涵盖连续时间、降雨量、站点的经纬度、海拔以及气候类型等,认为可以开发出一个通过关注更多特征,且结合长短期记忆网络(LSTM)、门控循环单元(GRU)和 Transformer,以有效且稳定地预测 AQI 的集成学习模型。

2. 相关工作

随着深度学习技术的进步,应用于非平稳时间序列的深度学习模型数量不断增加[10] [16] [22]-[24],预测精度和稳健度也随之提高。其中,林志烨等人基于元线性回归技术与门控循环单元(GRU)设计了有效的集成方案,与其他集成学习方法相比取得了更好的预测性能和精度[22];同期,毛文静提出将新型二进制黑猩猩优化算法(BChOA)与长短期记忆网络(LSTM)相结合以提高模型的准确性和效率的办法[16];比赫特·达斯采用了混合三种深度学习模型(LSTM、RNN 和 MLP)的方法,通过修改误差项和补充数据的方法提高了空气质量预测的准确性[24]。通过引入额外特征,深度学习模型性能得以提升。以下小节将回顾通过采用额外特征来提高预测准确性的相关研究。

2.1. 采用额外特征的 AQI 预测框架

通过引入观测点的具体天气情况或经纬度之间的关联,模型性能通常获得一定量的提升,这使得

额外添加地理特征成为一种常用于数据分析或模型预测的方法[25]-[28]。肖通过空间聚类开发了一种广义加法集成模型，通过不同的权重结合不同算法的预测结果，解释了各特征关于空气污染水平的关联性[27]。秦运用关联规则预测了极端颗粒物(PM10 和 PM2.5)的水平，通过揭示不同城市间 PM 污染的独特关联规则，指出南方城市的 PM 浓度与北方城市相比，尤其是 PM2.5，呈现出更强的时空相关性[28]。库尔特提出一种将地理模型与神经网络相结合(GFM-NN)的使用前馈反向传播神经网络来处理气象预报数据和空气污染物指标值的方法[26]。

尽管上述工作结合了额外的特征来设计预测模型，但这些模型仍然采用单一模型，而没有结合其他具有不同学习模式的模型以复合模型的特点用来提高预测精度。

2.2. 用于空气质量预测的集成学习模型

在之前的研究中，集成学习利用各种模型的优势用来提高预测准确性的优势被实证，成为一种关于在提升非平稳时间序列预测效果的一种常用且证实有效的方法[22] [29]-[33]。Ksibi 提出了采用包括 LightGBM、XGBoost、CatBoost 和 Extra Tree 回归的两层堆叠结构的集成框架，表现出优于各类基线模型和现有模型的性能[29]。Yu 通过调整模型架构并整合多种机器学习和深度学习方法[30]，开发了深度集成机器学习框架 DEML，在不同气候区、季节和建筑类型中都保持着高精度和稳定性。Aggarwal 采用模拟鸟群行为的随机优化算法来寻找最优超参数解，并籍此开发了混合的深度学习的框架，从而改善了长短期记忆(LSTM)模型的时空不稳定性[32]。

3. 基础模型

如上一节所述，集成学习融合了不同的基学习器，并引入了每个模型的特点。与传统的单一模型相比，它具有更好的性能。本文将分别使用 LSTM、GRU 和 Transformer 这三种模型框架作为基学习器。在实际应用上，两种 RNN 架构模型的融合度高，且原在 transformer 的基学习器内添加了具有 TCN 特征的因果卷积层以用于解决两变种 RNN 模型的长序列遗忘问题。

3.1. 数据集描述

在本研究使用的数据集为收集的 2016 年 1 月 1 日至 2024 年 12 月 31 日期间云南省环境监测站共计 46 个监测站点的监测数据。地表空气质量测量数据来自中国生态环境部(MEE)最初发布数据的镜像版本。此外，2 米气温、蒸发量、地面气压、总降水量和 10 米风 v 分量数据来自欧洲 era 数据集。训练数据的类型与单位见表 1。

Table 1. Data Type and Unit used in train
表 1. 训练数据的类型与数据单位

Type	Data Type	Unit
AQI	实时 AQI 值	
PM2.5	一小时平均 PM2.5 值	μg/m ³
PM10	一小时平均 PM10 值	μg/m ³
SO ₂	一小时平均 SO ₂ 值	μg/m ³
NO ₂	一小时平均 NO ₂ 值	μg/m ³
O ₃	一小时平均 O ₃ 值	μg/m ³
CO	一小时平均 CO 值	μg/m ³
2 m 露点温度	一小时内观测点的平均温度	K
蒸发量	一小时内观测点的总蒸发量	m of water equivalent
总降雨量	一小时内观测点的总降雨量	m
大气压力	观测点处的大气压强	kPa

3.2. CNN 数据补足

前文中提到收集的数据以表格 CSV 格式的形式存在, 存在各种缺失值。为了补充缺失值以及原数据不足连续这一问题, 本文使用了专门的卷积神经网络(CNN)技术, 通过舍弃全连接层并采用特征数量 $\times$ 双通道配置, 架构依赖于卷积层的局部建模能力, 可以区别于传统的图像 CNN 依赖空间位置来利用数据, 利用特征间的相关性学习数据的内在特征和模式来预测缺失位置的值, 以保证在适应当前任务功能的前提下, 更好地保留局部特征。

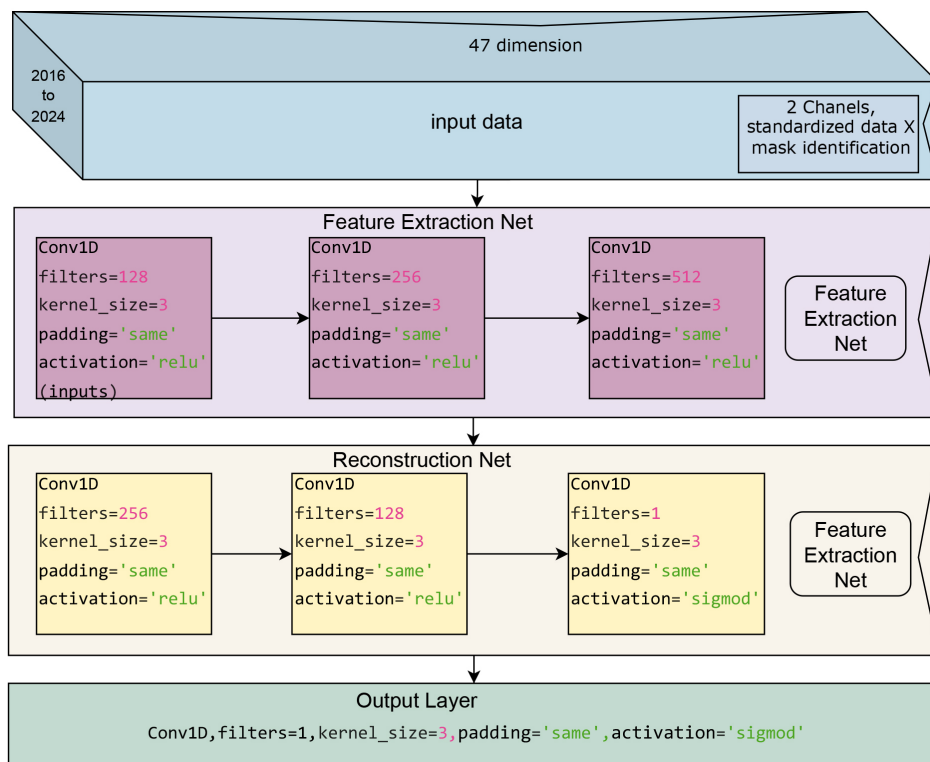


Figure 1. CNN model architecture

图 1. CNN 模型架构

如图 1 描述了本节提出的 CNN 具体架构, 由特征提取网络和重建网络两项组成。通过在编码器部分逐渐增加滤波器的数量, 而在解码器部分减少滤波器的数量, 模型可以做到捕获更复杂的特征和逐步重建输出。其中, 选择修正线性单元(ReLU)作为激活函数的目的是缓解模型中梯度消失的问题, 并加快收敛速度。仅在输出层将激活函数改为 sigmoid, 可以实现输出范围可控。通过上述设计, 该模型更适用于数据中常见的部分观测值缺失的问题。

3.3. 基学习器

如相关工作节所述, 深度学习框架已在此领域到应用并日趋成熟, 特别是长短期记忆网络和门控循环单元这两种特殊的循环神经网络模型, 在时间序列预测中展现出极高的效率和性能。此外, 为了在长时间序列上有良好的预测效果, 学者们创造出具有多头注意力机制的 Transformer 模型用于处理长期依赖关系。

本节将分别介绍使用长短期记忆网络(LSTM)、门控循环单元(GRU)和 Transformer 网络的三个基础模

型的设计。

3.3.1. LSTM 模型描述

长短期记忆(LSTM)模型是一种特殊的循环神经网络(RNN)架构。通过引入细胞状态, 以及如遗忘门(公式 1)、输入门(公式 2、3)和输出门(公式 5、6)等门控循环单元, 有效地解决了基础的 RNN 模型在处理长序列数据时遇到的严重梯度消失问题。

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

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

$$C_t^{(l)} = \tanh\left(W_C^{(l)}\left[h_{t-1}^{(l)}, x_t^{(l)}\right] + b_C^{(l)}\right) \quad (3)$$

$$C_t^{(l)} = f_t^{(l)} \odot C_{t-1}^{(l)} + i_t^{(l)} \odot C_t^{(l)} \quad (4)$$

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

$$h_t^{(l)} = o_t^{(l)} \odot \tanh\left(C_t^{(l)}\right) \quad (6)$$

图 2 展示了长短期记忆(LSTM)的神经单元, 其中 $f_t^{(l)}$ 决定要丢弃的历史信息的比例, $i_t^{(l)}$ 决定要添加的新信息的比例, $C_t^{(l)}$ 通过逐元素乘法实现信息的选择性保留和更新, $o_t^{(l)}$ 决定从细胞状态中传递到当前输出的信息。

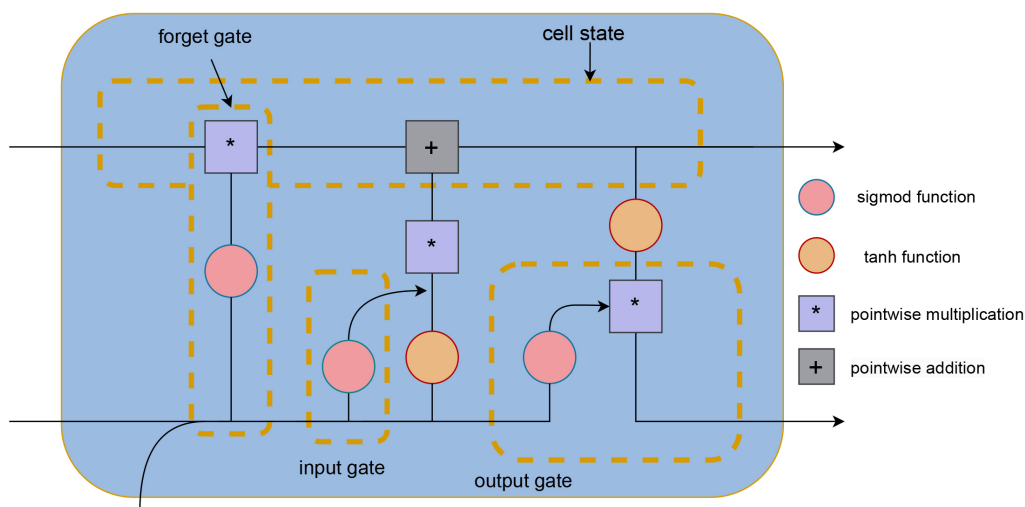


Figure 2. LSTM Neural unit

图 2. LSTM 神经单元

具体的说, 本模型通过使用双层 LSTM 层, 有效保留时间维度, 且通过堆叠 LSTM 层做到压缩时间信息的设计的形式成功捕捉污染物浓度的长期依赖变化。此外, 不同的随机失活(Dropout)率的区分, 和单元数量的减少, 使模型能够逐步压缩特征维度。这种方法增强了模型对高级特征的泛化能力, 并保留了更多原始信息。根据上述操作模型具体约束时间记忆能力, 而不过度限制输入特征的变换, 达成了“编码器 - 预测器”的结构从而实现端到端的预测。

图 3 展示了上述 LSTM 模型的架构, 其总共包含三层: 输入层、特征编码网络和预测输出网络。

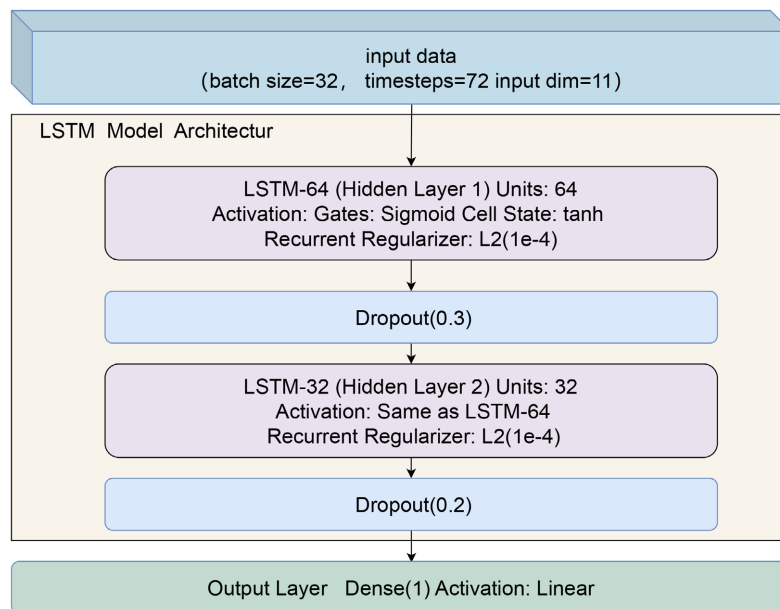


Figure 3. LSTM model architecture

图 3. LSTM 模型架构

3.3.2. GRU 模型描述

门控循环单元(GRU)是长短时记忆(LSTM)模型的一种变种。通过将 LSTM 的遗忘门和输入门合并为一个单一的更新门，并去除独立的细胞状态，减少了参数数量。更新门的数学公式如以下公式 7 所示。

$$Z_t = \sigma(w_z[h_{t-1}, x_t] + b_z) \quad (7)$$

与长短期记忆(LSTM)模型相比, GRU 的优势在于平衡性能和效率, 在快速迭代、资源有限或涉及小规模数据的场景表现优异。图 4 展示 GRU 模型的一个具体神经元。

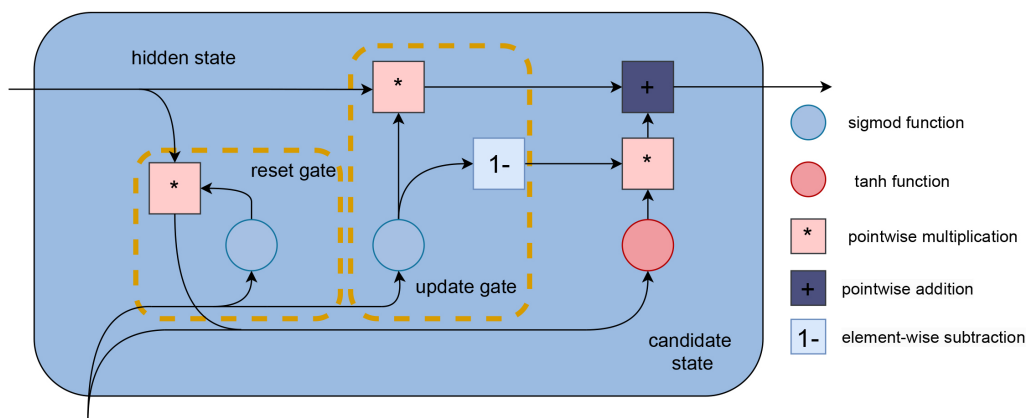


Figure 4. GRU Neural unit

图 4. GRU 神经元

因为 GRU 在计算时间和性能方面优于 LSTM 网络, 将 GRU 模型引入上述 LSTM 模型中, 可以确保快速迭代并增强模型多样性以提升集成模型的泛化能力, 同时提高模型的稳健性, 并降低单一模型存在偏差的风险。为了减少模型融合过程中拟合度的差异以及不恰当融合带来的风险, 本文中门控循环单元

(GRU)的设计与前文提到的长短期记忆网络(LSTM)模型基本一致。仅将相应的 LSTM 层替换为 GRU 层。

3.3.3. Transformer 模型描述

Transformer 模型具有独特的多头注意力机制,在处理序列数据时,能够并行计算每个位置与所有其他位置之间的关联程度,从而直接捕捉长序列中的长距离依赖关系,解决了随着序列长度增加,上述两种模型难以在长序列中有效利用远距离信息的问题。

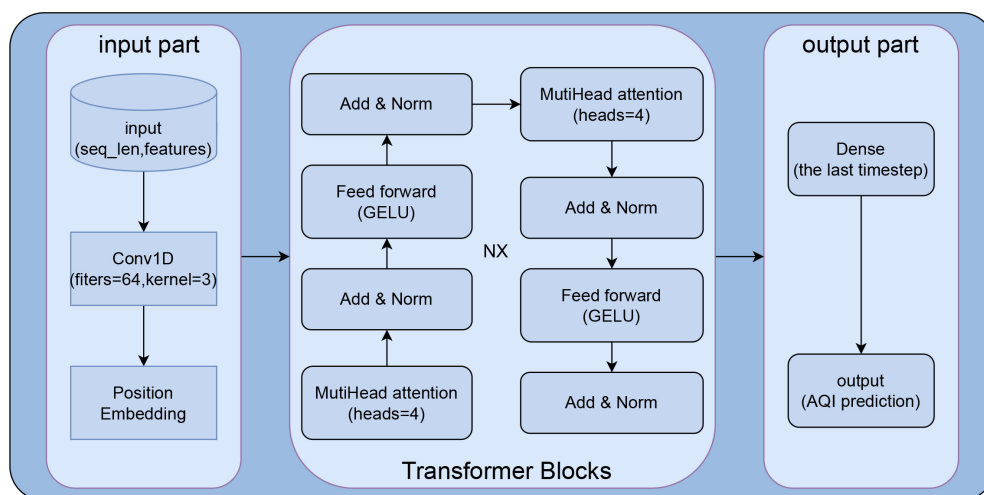


Figure 5. Transformer model architecture

图 5. Transformer 模型架构

图 5 展示为实现上述功能而引入的 Transformer 模型的网络架构,通过融合了因果卷积和 Transformer 模型的混合神经网络架构,在输入阶段使用因果卷积层作为主要特征提取器,时间因果关系得以保持完整,且模型在时空特征融合层中引入了的可学习的位置编码系统可以保证嵌入层在时间步上生成位置向量的动态编码方式使得模型能够自适应地学习不同时间尺度上的周期性特征。

本研究使用了 11 个数据特征,包括 PM2.5、SO₂、NO₂ 等,为保障能够解决在处理多变量空气质量数据,更好地捕捉多变量之间的相互作用关系以及解决数据中的噪声或局部数据缺失问题,基模型引入了对异常值的敏感性更低得 transformer 模型。模型中的多变量相互作用关系提供的互补信息保证了当变量的数据出现异常时,可以通过其他变量的关联模式推断出合理的值,从而提高模型的鲁棒性和准确性。

3.4. 堆叠元学习器

在时间序列预测领域,单一模型往往受自身特性限制,难以全面捕捉数据中的复杂模式与规律。集成学习整合了多个模型的优势,为提高预测性能提供了一种常规的有效方法。以随机森林(RF)为例,元模型的非线性有助于捕捉复杂的相互作用,使得其对于传统的平均方法有更好的表现。且随机森林在广泛的应用中都取得了显著成功,因而应用广泛,也促成了 RF 在各领域的案例研究中得到了证明和应用的成熟,也因其成熟性,我们选择 RF 作为集成学习的顶层元学习器。

使用随机森林的堆叠元学习器

集成模型相较于单一模型因为整合了多个模型的优点,在多层次的协同效应以及增强的系统稳健性上,为提升预测性能精度和泛用性的提升尤其是泛化性,鲁棒性和可解释性上提供了一种有效途径。

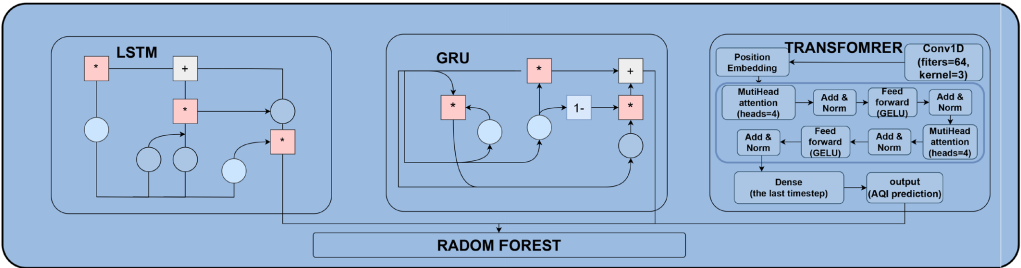


Figure 6. Ensemble random forest architecture
图 6. 集成随机森林模型架构

图 6 展示了为将 3.3 节中提到的基础模型与随机森林(RF)集成而设计的双层异构集成系统架构。通过互补的形式，引入了各模型的优势，来解决 LSTM 对短期波动敏感度不足以及 GRU 在长期依赖关系解析方面能力较弱的问题，同时，两种 RNN 变体模型的结合弥补了 Transformer 模型对局部突变响应延迟的问题。

通过分别训练三个基学习器，并通过堆叠使用随机森林(RF)作为额外训练的元学习器，整体集成框架得以保留了每个基学习器的特征提取模式和随机森林本身的复杂交互能力从而提高了框架的准确性和通用性。

4. 结果

本文旨在开发一种有效的能够在模型精度和稳健性之间保持平衡的集成空气质量指数预报模型网络，从而建立了三个基础学习器：Transformer、LSTM 和 GRU，并通过集成随机森林(RF)开发了一个集成学习框架。在模型完成后，我们使用仅进行基于 CNN 的数据插补得原始数据集来评估集成模型的性能，为评估模型在实际条件下的性能提供更真实的场景。

4.1. 评估指标

在模型的性能评估中，合理选择评估指标是准确衡量模型预测能力的关键一步。在模型指标的预测性评估中，单独使用任一指标都不足以全面准确地评估模型性能。通过均方根误差和平均绝对误差相结合的方式，同时引入 R2 指数从拟合优度的角度衡量模型对数据的解释能力，结合各评估指标达成相互验证和补充，减少单一指标的偏差，提高模型评估的可靠性和稳定性。

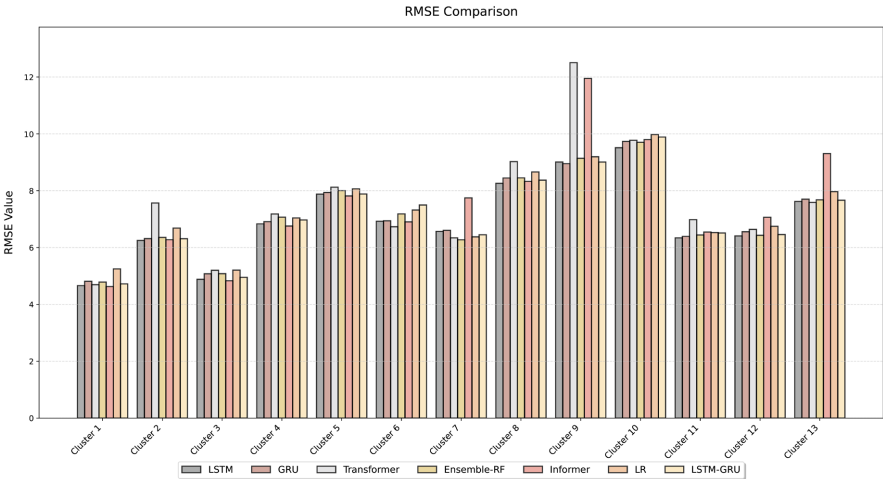


Figure 7. Bar-chart of RMSE value for models
图 7. 模型 RMSE 值柱状图

在以下三个等式中, y_i 是实际值, $\hat{y}_i$ 是预测值, n 是样本数量。

在上述提到的那些指标中, 均方根误差(RMSE)通过将误差平方然后再取平方根, 可以放大误差权重, 使得可以更好关注极端情况下的准确性。为更直观的感受, 图 7 展示了各模型的 RMSE 值。

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (8)$$

平均绝对误差(MAE)是预测值与真实值之间绝对误差的平均值, 能够直观地展示模型在整个预测范围内的平均误差水平。此外, MAE 的计算相对简单, 无需复杂的数学运算, 在本文的预测中, 可以更简便的了解到模型预测结果在总体的偏离情况, 避免因个别极端值的影响而高估或低估模型的性能。图 8 为各模型的 MAE 值柱状图。

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

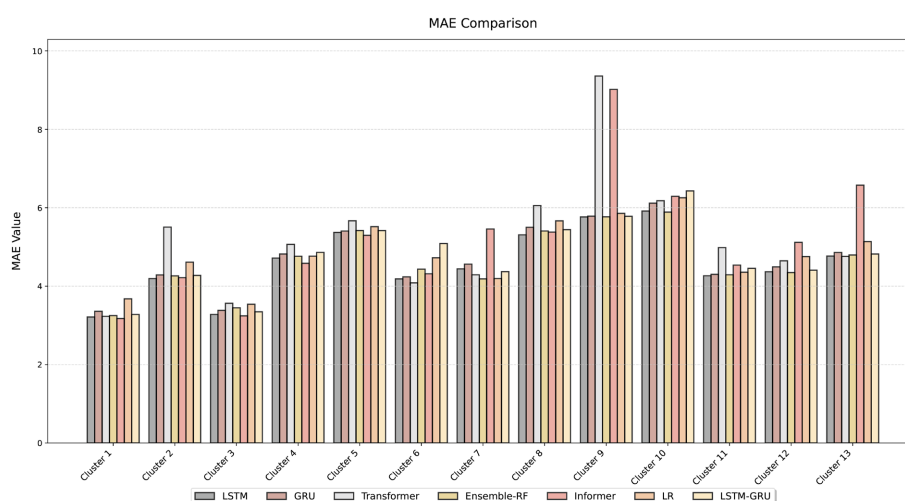


Figure 8. Bar-chart of MAE value for models

图 8. 模型 MAE 值柱状图

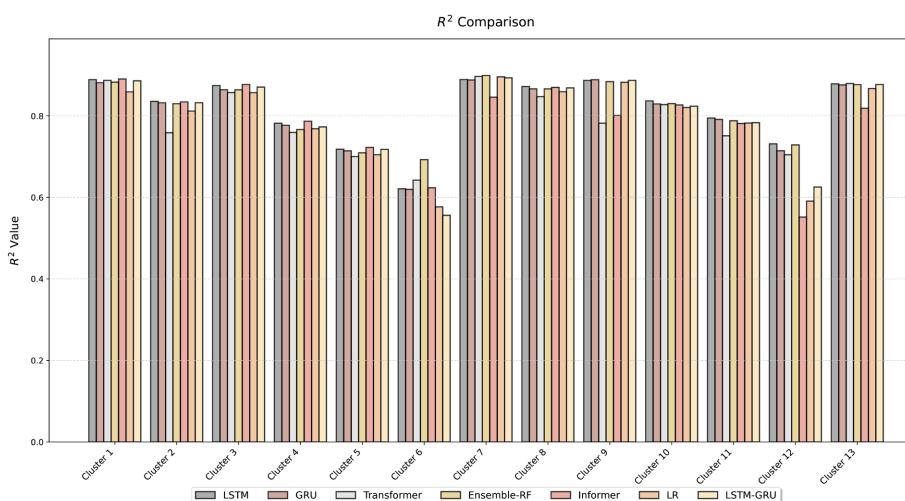


Figure 9. Bar-chart of R2 value for models

图 9. 模型 R2 值柱状图

R2 的值域定义明确(在 0 到 1 之间), 可以反映模型对数据的整体拟合优度。一般来说, R2 越接近 1, 模型的拟合效果越好, 模型就越能有效地捕捉空气质量变化的主要趋势和模式。各模型关于 R2 指标的柱状图见图 9。

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

4.2. 评估总结

本文的实验结果将对所提出的集成模型与三种基学习器 LSTM、GRU、Transformer 以及三种常用于时间序列预测的模型(LSTM-GRU、线性回归、Informer), 使用平均绝对误差 MAE、均方根误差 RMSE、R2 作为评估指标进行全面比较。

Table 2. Comprehensive collection of model evaluation metrics

表 2. 模型评估指标总汇

预测类	预测模型	Max	Min	Avg
MAE	LSTM	5.9150	3.2137	4.599
	GRU	6.1148	3.3595	4.7005
	Transformer	9.3577	3.2302	5.1835
	Ensemble -RF	5.8883	3.2484	4.6348
	LSTM-GRU	6.4291	3.2781	4.7670
	LR	6.2551	3.5381	4.8500
	Informer	9.0170	3.1726	5.1691
RMSE	LSTM	9.5143	4.6618	7.0129
	GRU	9.7337	4.8151	7.1075
	Transformer	12.5056	4.6956	7.5665
	Ensemble -RF	9.7048	4.7864	7.124
	LSTM-GRU	9.8888	4.7227	7.1316
	LR	9.9758	5.2072	7.31
	Informer	11.9515	4.6289	7.536
R2	LSTM	0.8888	0.621	0.8081
	GRU	0.8882	0.6139	0.8028
	Transformer	0.8964	0.6041	0.7838
	Ensemble -RF	0.8986	0.6923	0.8010
	LSTM-GRU	0.8928	0.556	0.7992
	LR	0.8953	0.5766	0.7901
	Informer	0.8900	0.7865	0.5515

表 2 为常规的预测模型在本数据集上的表现的汇总结果。如上表所示, 集成模型在原始数据上的表现对于基学习器模型有一定优势, 表现最佳的单个基础模型 GRU、LSTM 和 Transformer 仅获得了平均 R2 分数为 0.803、0.808 和 0.784, 而集成学习框架则可以达到 0.8986。除了属于模型集成的基学习器外, 在各项指标上集成模型也优于 Informer、LSTM-GRU 和线性回归等常用模型。考察对各指标的最值的进一步验证中, 可以看出不同模型对数据分布表现得更为敏感, 而集成模型因其集成优势能证明更稳定的性能的优势, 因此容易在特定的集群中会出现显著波动的各模型区分开来。这种增强的稳定性, 再加上更高的准确性, 使其成为目标预测任务中更可靠、更稳定的解决方案。

5. 讨论

本研究开发了一种集成时间序列预测模型，在实际应用中，该模型在数据集上各个不同的集群涵盖的 13 个测试样本中，其性能优于基础模型的最佳性能。实验验证了该集成策略在捕捉复杂数据模式方面的有效性，与其他用于空气质量指数(AQI)预测的模型网络架构相比，这种优势表明，该集成模型在处理具有异质性特征的数据时具有更强的适应性和稳健性，在实际应用场景中的精准预测更加有力。

然而，尽管其平均表现令人瞩目，但本模型在特定场景中最低的 R^2 值仅为 0.6920，仍存在局限性，表明其泛化能力还有提升空间。此外，本研究使用的数据来源于中国云南省，这是一个内陆地区，拥有多样的气候类型和民族群体，本身具有地理和环境的特殊性，这样的区域独特性可能会影响其更广泛的适用性，可见能否推广到全国或大陆范围仍需进一步研究。

参考文献

- [1] Sousa, S., Martins, F., Alvimferraz, M. and Pereira, M. (2007) Multiple Linear Regression and Artificial Neural Networks Based on Principal Components to Predict Ozone Concentrations. *Environmental Modelling & Software*, **22**, 97-103. <https://doi.org/10.1016/j.envsoft.2005.12.002>
- [2] Ancelet, T., Davy, P.K., Trompetter, W.J., Markwitz, A. and Weatherburn, D.C. (2014) Sources and Transport of Particulate Matter on an Hourly Time-Scale during the Winter in a New Zealand Urban Valley. *Urban Climate*, **10**, 644-655. <https://doi.org/10.1016/j.uclim.2014.06.003>
- [3] Li, Y., Peng, T., Hua, L., Ji, C., Ma, H., Nazir, M.S., et al. (2022) Research and Application of an Evolutionary Deep Learning Model Based on Improved Grey Wolf Optimization Algorithm and DBN-ELM for AQI Prediction. *Sustainable Cities and Society*, **87**, Article 104209. <https://doi.org/10.1016/j.scs.2022.104209>
- [4] Song, C. and Fu, X. (2020) Research on Different Weight Combination in Air Quality Forecasting Models. *Journal of Cleaner Production*, **261**, Article 121169. <https://doi.org/10.1016/j.jclepro.2020.121169>
- [5] Li, G., Tang, Y. and Yang, H. (2022) A New Hybrid Prediction Model of Air Quality Index Based on Secondary Decomposition and Improved Kernel Extreme Learning Machine. *Chemosphere*, **305**, Article 135348. <https://doi.org/10.1016/j.chemosphere.2022.135348>
- [6] Wei, W., Ramalho, O., Malingre, L., Sivanantham, S., Little, J.C. and Mandin, C. (2019) Machine Learning and Statistical Models for Predicting Indoor Air Quality. *Indoor Air*, **29**, 704-726. <https://doi.org/10.1111/ina.12580>
- [7] Li, H., Wang, J., Li, R. and Lu, H. (2019) Novel Analysis-Forecast System Based on Multi-Objective Optimization for Air Quality Index. *Journal of Cleaner Production*, **208**, 1365-1383. <https://doi.org/10.1016/j.jclepro.2018.10.129>
- [8] Xu, R., Deng, X., Wan, H., Cai, Y. and Pan, X. (2021) A Deep Learning Method to Repair Atmospheric Environmental Quality Data Based on Gaussian Diffusion. *Journal of Cleaner Production*, **308**, Article 127446. <https://doi.org/10.1016/j.jclepro.2021.127446>
- [9] Liu, H., Yue, F. and Xie, Z. (2022) Quantify the Role of Anthropogenic Emission and Meteorology on Air Pollution Using Machine Learning Approach: A Case Study of PM_{2.5} during the COVID-19 in Hubei Province, China. *Environmental Pollution*, **300**, Article 118932. <https://doi.org/10.1016/j.envpol.2022.118932>
- [10] Middya, A.I. and Roy, S. (2022) Pollutant Specific Optimal Deep Learning and Statistical Model Building for Air Quality Forecasting. *Environmental Pollution*, **301**, Article 118972. <https://doi.org/10.1016/j.envpol.2022.118972>
- [11] Zheng, L., Lin, R., Wang, X. and Chen, W. (2021) The Development and Application of Machine Learning in Atmospheric Environment Studies. *Remote Sensing*, **13**, Article 4839. <https://doi.org/10.3390/rs13234839>
- [12] Yazdi, M.D., Kuang, Z., Dimakopoulou, K., Barrat, B., Suel, E., Amini, H., Lyapustin, A., Katsouyanni, K. and Schwartz, J. (2019) Predicting Fine Particulate Matter (PM_{2.5}) in the Greater London AREA: an Ensemble Approach Using Machine Learning Methods. *Remote Sensing*, **3**, 355-356.
- [13] Huang, C. and Kuo, P. (2018) A Deep CNN-LSTM Model for Particulate Matter (PM_{2.5}) Forecasting in Smart Cities. *Sensors*, **18**, Article 2220. <https://doi.org/10.3390/s18072220>
- [14] Subramanian, D. and Marimuthu, P. (2024) Design and Synthesis of Time Series Forecasting and Deep Learning Prediction Model for Air Quality Index Prediction in Indian Cities. *Indian Journal of Engineering and Materials Sciences*, **31**, 589-600.
- [15] Tsai, Y., Zeng, Y. and Chang, Y. (2018) Air Pollution Forecasting Using RNN with LSTM. 2018 *IEEE 16th Intl Conf on Dependable, Autonomic and Secure Computing*, 16th *Intl Conf on Pervasive Intelligence and Computing*, 4th *Intl Conf on Big Data Intelligence and Computing and Cyber Science and Technology Congress (DASC/PiCom/DataCom/*

- CyberSciTech), Athens, 12-15 August 2018, 1074-1079.
<https://doi.org/10.1109/dasc/picom/datacom/cyberscitech.2018.00178>
- [16] Mao, W., Wang, W., Jiao, L., Zhao, S. and Liu, A. (2021) Modeling Air Quality Prediction Using a Deep Learning Approach: Method Optimization and Evaluation. *Sustainable Cities and Society*, **65**, Article 102567. <https://doi.org/10.1016/j.scs.2020.102567>
 - [17] Espinosa, R., Palma, J., Jiménez, F., Kamińska, J., Sciacicco, G. and Lucena-Sánchez, E. (2021) A Time Series Forecasting Based Multi-Criteria Methodology for Air Quality Prediction. *Applied Soft Computing*, **113**, Article 107850. <https://doi.org/10.1016/j.asoc.2021.107850>
 - [18] Méndez, M., Merayo, M.G. and Núñez, M. (2023) Machine Learning Algorithms to Forecast Air Quality: A Survey. *Artificial Intelligence Review*, **56**, 10031-10066. <https://doi.org/10.1007/s10462-023-10424-4>
 - [19] Liu, H., Yan, G., Duan, Z. and Chen, C. (2021) Intelligent Modeling Strategies for Forecasting Air Quality Time Series: A Review. *Applied Soft Computing*, **102**, Article 106957. <https://doi.org/10.1016/j.asoc.2020.106957>
 - [20] Sun, W. and Li, Z. (2020) Hourly PM_{2.5} Concentration Forecasting Based on Feature Extraction and Stacking-Driven Ensemble Model for the Winter of the Beijing-Tianjin-Hebei Area. *Atmospheric Pollution Research*, **11**, 110-121. <https://doi.org/10.1016/j.apr.2020.02.022>
 - [21] Nur'aiah, Z., Woen, E.L., Najah, A.A. and Abdul, M.M. (2021) A Systematic Literature Review of Deep Learning Neural Network for Time Series Air Quality Forecasting. *Environmental Science and Pollution Research*, **29**, 4958-4990. <https://doi.org/10.1007/s11356-021-17442-1>
 - [22] Lin, C.-Y., Chang, Y.-S. and Abimannan, S. (2021) Ensemble Multifeatured Deep Learning Models for Air Quality Forecasting. *Atmospheric Pollution Research*, **12**, Article 101045. <https://doi.org/10.1016/j.apr.2021.03.008>
 - [23] Diaz Resquin, M., Lichtig, P., Alessandrello, D., De Oto, M., Gómez, D., Rössler, C., et al. (2023) A Machine Learning Approach to Address Air Quality Changes during the COVID-19 Lockdown in Buenos Aires, Argentina. *Earth System Science Data*, **15**, 189-209. <https://doi.org/10.5194/essd-15-189-2023>
 - [24] Das, B., Dursun, Ö.O. and Toraman, S. (2022) Prediction of Air Pollutants for Air Quality Using Deep Learning Methods in a Metropolitan City. *Urban Climate*, **46**, Article 101291. <https://doi.org/10.1016/j.uclim.2022.101291>
 - [25] Zhao, Z., Wu, J., Cai, F., Zhang, S. and Wang, Y. (2022) A Statistical Learning Framework for Spatial-Temporal Feature Selection and Application to Air Quality Index Forecasting. *Ecological Indicators*, **144**, Article 109416. <https://doi.org/10.1016/j.ecolind.2022.109416>
 - [26] Kurt, A. and Oktay, A.B. (2010) Forecasting Air Pollutant Indicator Levels with Geographic Models 3days in Advance Using Neural Networks. *Expert Systems with Applications*, **37**, 7986-7992. <https://doi.org/10.1016/j.eswa.2010.05.093>
 - [27] Xiao, Q., Chang, H., Geng, G. and Liu, Y. (2018) An Ensemble Machine-Learning Model to Predict Historical PM_{2.5} Concentrations in China from Satellite Data. *ISEE Conference Abstracts*, **2018**, 13260-13269. <https://doi.org/10.1289/isesisee.2018.o01.04.07>
 - [28] Qin, S., Liu, F., Wang, C., Song, Y. and Qu, J. (2015) Spatial-Temporal Analysis and Projection of Extreme Particulate Matter (PM₁₀ and PM_{2.5}) Levels Using Association Rules: A Case Study of the Jing-Jin-Ji Region, China. *Atmospheric Environment*, **120**, 339-350. <https://doi.org/10.1016/j.atmosenv.2015.09.006>
 - [29] Ksibi, A., Salhi, A., Saleh Alluhaidan, A. and A. El-Rahman, S. (2023) Airstacknet: A Stacking Ensemble-Based Approach for Air Quality Prediction. *Computers, Materials & Continua*, **74**, 2073-2096. <https://doi.org/10.32604/cmc.2023.032566>
 - [30] Yu, W., Nakisa, B., Ali, E., Loke, S.W., Stevanovic, S. and Guo, Y. (2023) Sensor-based Indoor Air Temperature Prediction Using Deep Ensemble Machine Learning: An Australian Urban Environment Case Study. *Urban Climate*, **51**, Article 101599. <https://doi.org/10.1016/j.uclim.2023.101599>
 - [31] Wu, Q. and Lin, H. (2019) A Novel Optimal-Hybrid Model for Daily Air Quality Index Prediction Considering Air Pollutant Factors. *Science of The Total Environment*, **683**, 808-821. <https://doi.org/10.1016/j.scitotenv.2019.05.288>
 - [32] Aggarwal, A. and Toshniwal, D. (2021) A Hybrid Deep Learning Framework for Urban Air Quality Forecasting. *Journal of Cleaner Production*, **329**, Article 129660. <https://doi.org/10.1016/j.jclepro.2021.129660>
 - [33] Huang, H. and Qian, C. (2023) Modeling PM_{2.5} Forecast Using a Self-Weighted Ensemble GRU Network: Method Optimization and Evaluation. *Ecological Indicators*, **156**, Article 111138. <https://doi.org/10.1016/j.ecolind.2023.111138>