

基于多模型协同优化的空气质量指数预测研究 ——以福州市为例

李小英

福建师范大学数学与统计学院, 福建 福州

收稿日期: 2026年6月15日; 录用日期: 2026年8月19日; 发布日期: 2026年9月15日

摘 要

空气质量指数(Air Quality Index, AQI)的精准预测对于空气污染防治和环境管理具有重要意义。本文以福州市为研究对象, 收集2021年1月至2025年3月的AQI、六项主要空气污染物及气象数据, 采用XGBoost和Pearson相关系数进行特征筛选, 构建XGBoost、BP神经网络、BiLSTM、Prophet等单一预测模型, 以及STL-XGBoost、BiLSTM-STL-XGBoost、BP-BiLSTM-Stacking等多模型融合预测模型, 并利用支持向量机(SVM)开展空气质量等级分类研究。结果表明: BP神经网络在单一模型中预测性能最佳; BP-BiLSTM-Stacking融合模型综合性能最优, 在AQI预测任务中取得最高的预测精度; 优化后的SVM模型能够有效提升空气质量等级分类效果, 尤其提高了轻度污染类别的识别准确率。研究结果表明, 多模型融合方法能够充分发挥不同模型的互补优势, 提高AQI预测的准确性和稳定性, 可为空气质量预测及相关环境管理研究提供参考。

关键词

AQI, 空气质量预测, 多模型融合, BP神经网络, 支持向量机

Air Quality Index Prediction Based on Multi-Model Collaborative Optimization —A Case Study of Fuzhou

Xiaoying Li

School of Mathematics and Statistics, Fujian Normal University, Fuzhou Fujian

Received: June 15, 2026; accepted: August 19, 2026; published: September 15, 2026

Abstract

Accurate prediction of the Air Quality Index (AQI) is of great importance for air pollution prevention,

文章引用: 李小英. 基于多模型协同优化的空气质量指数预测研究[J]. 可持续发展, 2026, 16(9): 69-81.

DOI: 10.12677/sd.2026.169317

environmental management, and public health protection. Taking Fuzhou as the study area, this study collected daily AQI data, concentrations of six major air pollutants, and meteorological data from January 2021 to March 2025. XGBoost and Pearson correlation analysis were employed for feature selection. Four individual prediction models, including XGBoost, Back Propagation Neural Network (BPNN), Bidirectional Long Short-Term Memory (BiLSTM), and Prophet, as well as several multi-model fusion frameworks, including STL-XGBoost, BiLSTM-STL-XGBoost, and BP-BiLSTM-Stacking, were developed and evaluated. In addition, a Support Vector Machine (SVM) was applied to classify air quality levels. The results indicate that the BPNN achieved the best performance among the individual models, while the BP-BiLSTM-Stacking model demonstrated the highest overall prediction accuracy for AQI forecasting. Moreover, the optimized SVM significantly improved the classification performance of air quality levels, particularly in identifying the minority class of lightly polluted days. The findings suggest that multi-model fusion can effectively exploit the complementary strengths of different models, thereby improving the accuracy and robustness of AQI prediction and providing a useful reference for air quality forecasting and environmental management.

Keywords

Air Quality Index, Air Quality Prediction, Multi-Model Fusion, Back Propagation Neural Network (BPNN), Support Vector Machine (SVM)

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. 绪论

1.1. 研究背景与意义

近年来,随着经济社会的快速发展和城市化进程的不断推进,大气污染问题受到广泛关注。空气污染不仅影响居民身体健康,还会对生态环境、交通运输以及社会经济活动产生不利影响。为改善生态环境质量,我国相继出台了一系列政策文件,对绿色低碳发展和环境治理提出了更高要求。在这一背景下,提高空气质量预测的准确性,对于污染防治、环境管理和公众健康保障具有重要意义。

福州市作为我国东南沿海重要中心城市,近年来空气质量总体保持较好水平,但仍存在部分时段空气质量波动和轻度污染现象。福州市空气质量等级为“优”和“良”的天数占绝大多数,但轻度污染天气仍偶有发生,说明空气质量仍具有进一步改善空间。因此,对福州市空气质量指数 AQI 进行准确预测,不仅有助于及时掌握空气质量变化趋势,也能够为环境监测、污染预警及相关管理提供数据支持。

近年来,随着人工智能技术的发展,机器学习和深度学习方法逐渐应用于空气质量预测领域。相比传统统计模型,这类方法能够更好地挖掘空气质量数据中的非线性关系和复杂时序特征,提高预测精度。然而,不同模型各有优势和局限,单一模型难以兼顾预测准确性和稳定性。因此,本文以福州市 AQI 预测为研究对象,综合利用 Prophet、XGBoost、BP 神经网络、BiLSTM 等模型,构建多模型融合预测框架,并结合支持向量机开展空气质量等级分类研究,以期提高 AQI 预测性能,为空气质量预测研究提供参考。

1.2. 文献综述

早期研究多采用自回归积分滑动平均(ARIMA/SARIMA)模型对 AQI 进行短期预测,但其对非线性和高阶时序依赖难以充分捕捉。李子熠等提出一种 XGBoost-LSTM 组合模型用于短期 PM_{2.5} 浓度预测,该方法通过 XGBoost 提取特征生成预测值作为增强特征,输入 LSTM 捕捉时序依赖关系[1]。王子航等构建了 LSTM

多步预测模型,并结合斯皮尔曼相关性和灰色关联分析完成特征筛选,在长期 AQI 预测中表现优异[2]。

随着计算资源的普及,支持向量回归(SVR)、随机森林(RF)及梯度提升树等方法逐渐成为 AQI 预测的主流。有研究表明,SVR 通过核函数映射高维特征空间可有效建模非线性关系[3],而 XGBoost 凭借残差迭代优化策略显著提升了突变事件预测精度[4]。针对空气质量时序的长时依赖性,长短期记忆网络(LSTM)及其变体(如 BiLSTM、GRU)通过门控机制捕捉动态特征[5]。为克服单一模型局限性,多模型融合技术(如 Stacking)通过异质模型互补提升稳健性[6]。

总体来看,现有研究取得了较为丰富的成果,但仍存在以下不足:一是多数研究主要关注单一模型或单一融合策略,对不同类型融合模型的系统比较相对较少;二是对于传统时间序列模型、机器学习模型和深度学习模型协同优化机制的研究仍不够充分;三是针对福州市空气质量预测的相关研究相对有限,多模型融合框架在该地区的适用性仍有待进一步验证。

基于此,本文以福州市 AQI 预测为研究对象,综合运用 XGBoost、BP 神经网络、BiLSTM、Prophet 等模型,构建多种协同优化预测框架,并结合支持向量机开展空气质量等级分类研究,对不同模型的预测性能进行系统比较,以期提高 AQI 预测精度,并为城市空气质量预测研究提供参考。

1.3. 研究创新

本研究通过查阅空气质量研究方面的文献资料,在全面了解空气质量研究预测方面的研究情况下,总结出了两个创新点。

(1) 特征筛选的协同策略应用

在变量筛选阶段,采用了结合机器学习模型(XGBoost)内在重要性评估与传统统计学(Pearson 相关系数)分析的双重筛选方法,综合两种特征选择方法的结果,最终选取 PM2.5、PM10、NO₂、SO₂、O₃、湿度六个变量建立预测模型,这种方法兼顾了变量的线性和非线性影响,筛选出更具信息量和稳健性的特征变量,使模型具有更良好的预测能力。

(2) 多模型协同优化框架探索

构建并系统性地比较了多种不同机制的 AQI 预测协同优化框架。不同于许多研究仅聚焦于单一类型的组合模型,本研究探索并对比了:基于时间序列分解的组合(如 STL-XGBoost 及其增强版)[7],结合时序分解与深度学习的混合策略(如 BiLSTM-STL-XGBoost),基于残差修正的组合(如 Prophet-RF)等,较全面的比较为选择最优的 AQI 预测策略提供了实证依据。

2. 数据与方法介绍

2.1. 数据来源与介绍

1) 研究区域概况

福州市是福建省的省会城市,位于中国东南部沿海地区,地处闽江下游河谷盆地,地貌以丘陵山地为主。市区坐落于福州盆地内,四周被鼓山、旗山、五虎山等低山环抱,闽江穿城而过。福州市气候属于典型的亚热带季风气候,雨季集中在 4~6 月,常出现梅雨天气,夏季高温频繁,冬季霜雪罕见。

2) 数据来源与缺失处理

本文选取的数据为 2021 年 1 月 1 日至 2025 年 3 月 31 日逐日的数据,共 1551 组,为遵循时间序列建模中“未来不可见”的预测逻辑,并更贴近实际应用场景,其中 2021 年 1 月 1 日至 2025 年 2 月 28 日的数据作为训练集,用于建立模型,2025 年 3 月的数据作为测试集,用于模型的效度分析。

数据来源包括两部分,第一部分是福州市空气质量指数(AQI)以及空气质量等级和六种主要污染物的浓度数据,这部分数据来源于中国空气质量在线监测分析平台历史数据网站,网址为

<http://www.aqistudy.cn/historydata>, 其中 AQI 是用于定量描述空气质量状况的无量纲指数, 主要反映空气污染程度及其对健康的影响, 它基于多种污染物浓度计算得出, 数值越高, 污染越严重, 健康风险越大, 而空气质量等级的划分如表 1 所示, 六种主要污染物分别为 PM_{2.5}、PM₁₀、NO₂、SO₂、CO、O₃; 第二部分是福州市的气象数据, 包括大气温度的、相对湿度、平均风速、降水量, 数据来源于全球天气与气象数据网站 RP5, 网址为 <https://rp5.ru>, 气象数据选取的是站点当天早上 8 时所观测的数据。

在数据采集的过程中, 由于种种原因导致污染物 O₃ 有 13 条缺失数据, 为了降低其影响, 以提高模型的准确性, 本文对缺失数据采用建立时间序列模型 SARIMA 的方式来处理, 运用 SARIMA 模型来预测缺失数据的值。

Table 1. AQI ranges and corresponding air quality grades

表 1. AQI 取值范围和空气质量等级分布

空气质量指数	空气质量级别	空气质量指数类别	颜色
0~50	一级	优	绿
51~100	二级	良	黄
101~150	三级	轻度污染	橙
151~200	四级	中度污染	红
201~300	五级	重度污染	紫
>300	六级	严重污染	褐红

2.2. 评价指标

对预测模型的性能评估, 通常综合运用多种评价指标从预测准确性和模型稳定性等维度进行全面考量, 在本研究中采用了三个常用的评价指标: 平均绝对误差(Mean Absolute Deviation, MAE), 均方根误差(Root Mean Squared Error, RMSE), 相关系数(R squared, R²)。

MAE 表示真实值和预测值之间的误差的绝对值的平均, 取值范围为[0, +∞), 计算公式如下:

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

RMSE 表示真实值和预测值两者的偏离程度, 取值范围为[0, +∞), 计算公式为:

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

R² 表示原始数据值与预测值之间的拟合程度, 计算公式如下:

$$R^2 = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y})^2} \quad (3)$$

其中, y_i 是第 i 个样本的真实值, $\hat{y}_i$ 是第 i 个样本的预测值, $\bar{y}$ 是样本的平均值。MAE、RMSE 值越小代表模型效果越好, R² 越接近 1 说明模型拟合程度越好。

3. 基于单一模型的 AQI 预测

为了增强组合模型预测结果的可比性, 本文选取了包括 XGBoost、BP 神经网络、BiLSTM 网络、Prophet 共四种单个模型作为对照组进行预测 AQI。为了提升模型的预测能力且降低冗余信息的干扰, 在模型构建前, 进行了特征筛选。

3.1. 特征选择方法与变量确定

在开展空气质量指数(AQI)预测建模之前,为确保模型输入变量具备良好的预测能力并降低冗余信息干扰,本文采用了 XGBoost 模型和 Pearson 相关系数分析对特征进行筛选。

XGBoost 模型根据每个特征在所有树结构中对模型性能提升的“增益(gain)”进行排序,本研究中基于该方法得到的重要性前六位特征依次为:

O₃、PM2.5、PM10、SO₂、湿度、温度。其次,为衡量各变量与 AQI 之间的线性相关性,运用 Pearson 相关系数的绝对值,得出相关性最高的前六个特征依次为: O₃、PM2.5、PM10、湿度、NO₂、降水量。

对比两种方法可以发现,PM2.5、PM10、O₃、湿度在两个排序中均排名靠前,说明其对 AQI 的解释能力较强,具有稳定的预测价值。SO₂ 在 XGBoost 模型中排名第四,尽管其线性相关性不高,但在非线性模型中的贡献不容忽视,因此被保留作为重要特征。NO₂ 在相关系数法中表现良好,虽然在 XGBoost 中贡献较低,但考虑其潜在影响,也将其纳入最终特征集中。

综上所述,本文综合两种特征选择方法的结果,最终确定以下六个变量作为 AQI 预测模型的输入特征: PM2.5、PM10、NO₂、SO₂、O₃、湿度。该特征组合兼顾了线性和非线性特征影响,有助于提升模型的预测精度与稳健性。

3.2. 基于单一模型的 AQI 预测结果与分析

1. XGBoost 模型

本研究利用 Scikit-learn 提供 XGBoost 库实现 XGBoost 模型,并利用网格搜索方法寻找各个参数的最优值。该模型的预测结果如图 1 所示,且其 MAE 值为 3.32, RMSE 值为 7.62, R² 为 0.9295,模型的预测效果较好。但需要指出的是,较高的 R² 并不意味着模型具有完全的泛化能力。由于本研究仅以福州市数据进行模型训练与验证,模型在其他地区或更长时间尺度上的适用性仍需进一步检验。未来可利用跨城市、多年份数据开展外部验证,以进一步评估模型的稳定性和泛化能力。

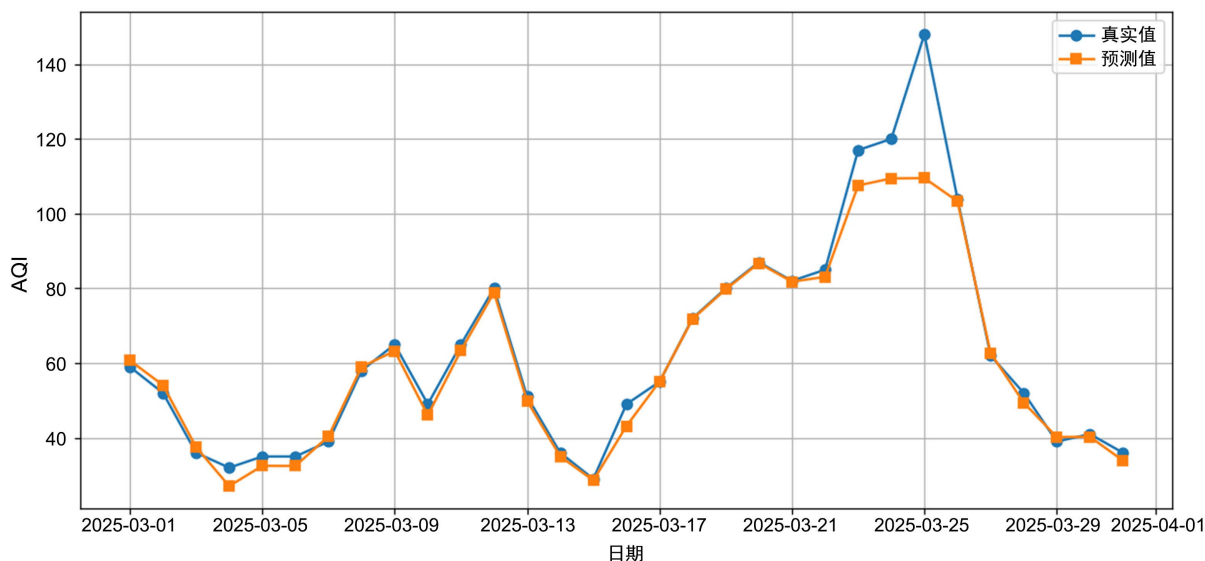


Figure 1. Prediction results of the XGBoost model

图 1. XGBoost 模型预测结果

2. BP 神经网络

对于 BP 神经网络采用 TensorFlow Keras 中的 Sequential 神经网络模型来构建一个多层感知机(MLP),

实现 BP 神经网络的功能。BP 神经网络的预测结果图 2 所示, 且 MAE 值为 1.98, RMSE 值为 2.77, R^2 为 0.9907, 该模型的预测效果非常好。

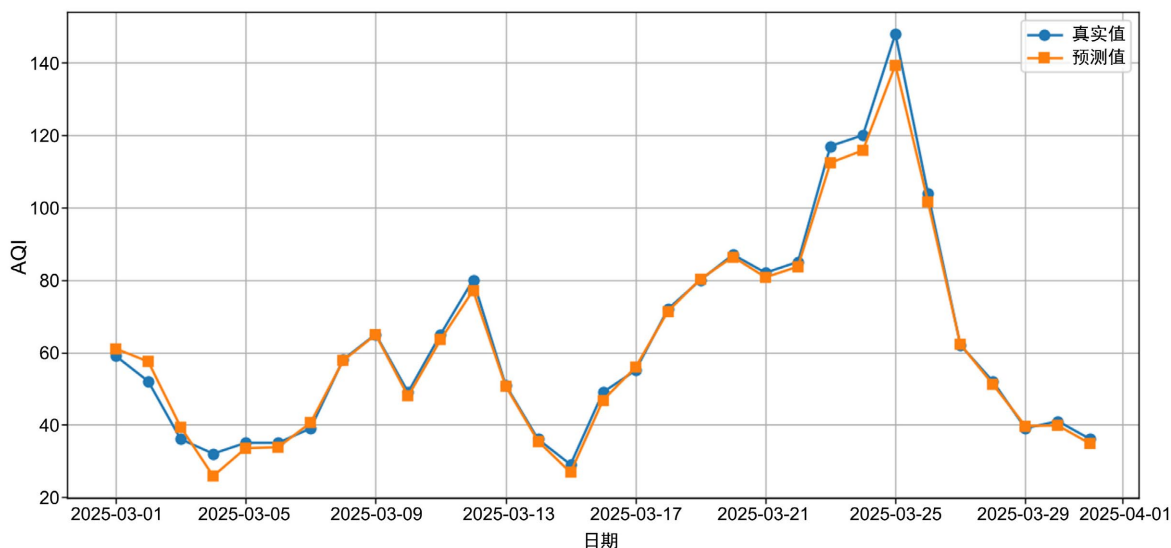


Figure 2. Prediction results of the BP neural network model

图 2. BP 神经网络预测结果

3. 双向长短期记忆网络

BiLSTM 网络是采用 TensorFlow 的 Keras API 中的 Sequential 模型和 Bidirectional (LSTM)层来实现的。BiLSTM 网络的预测结果如图 3 所示, 且其 MAE 值为 2.46, RMSE 值为 3.21, R^2 为 0.9875, 该模型的预测效果非常好。

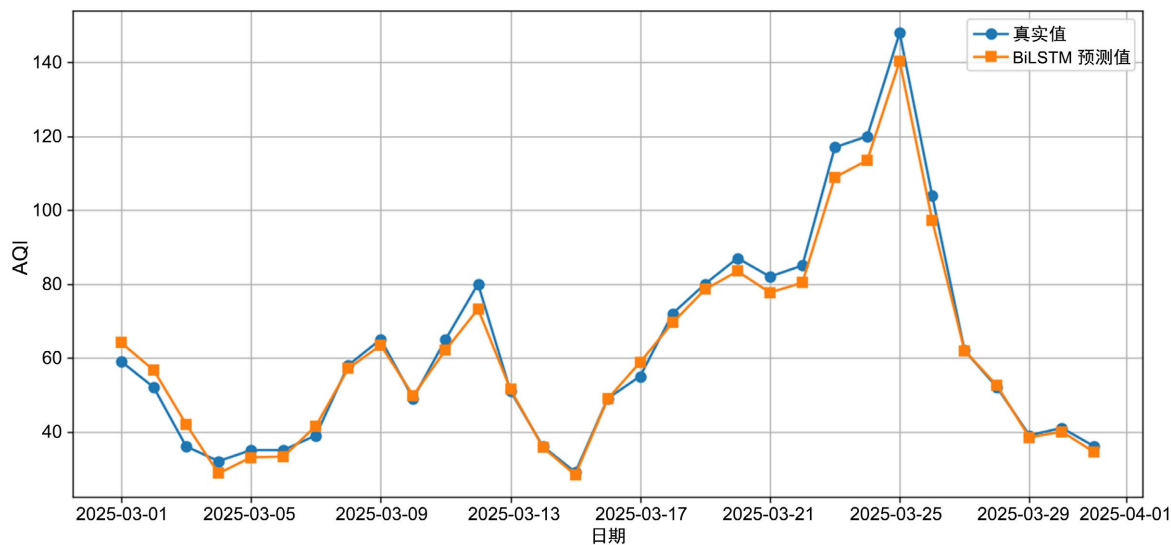


Figure 3. Prediction results of the BiLSTM network model

图 3. BiLSTM 网络预测结果

4) Peophet 模型

本文采用了 Prophet 模型对福州市的 AQI 进行预测[8], 加入了前文所提及的外部变量(PM2.5、PM10、

NO₂、SO₂、O₃、湿度)作为额外回归量。Prophet 模型在整个数据集上的拟合效果如图 4 所示,红色的点为真实观测到的 AQI 值,蓝色的实线为模型所拟合的预测值,灰色区域为 80%置信区间。从图 4 中可以看出,模型捕捉到了一定周期性波动趋势,但是真实的 AQI 值表现出显著的波动性且有极端值,这些值偏离模型的预测线,落在了置信区间以外。

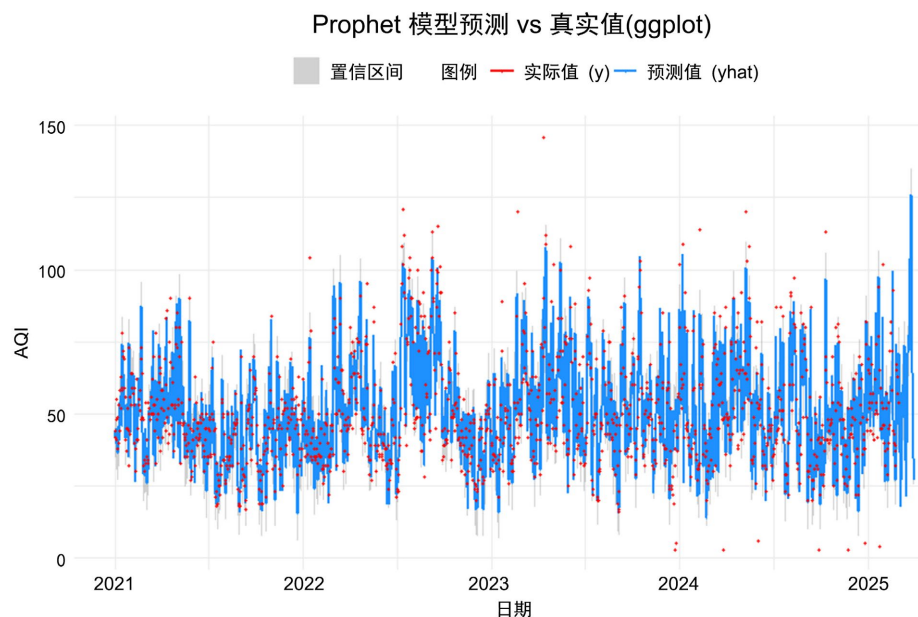


Figure 4. Prediction results and confidence intervals of the prophet model

图 4. Prophet 模型的预测结果与置信区间

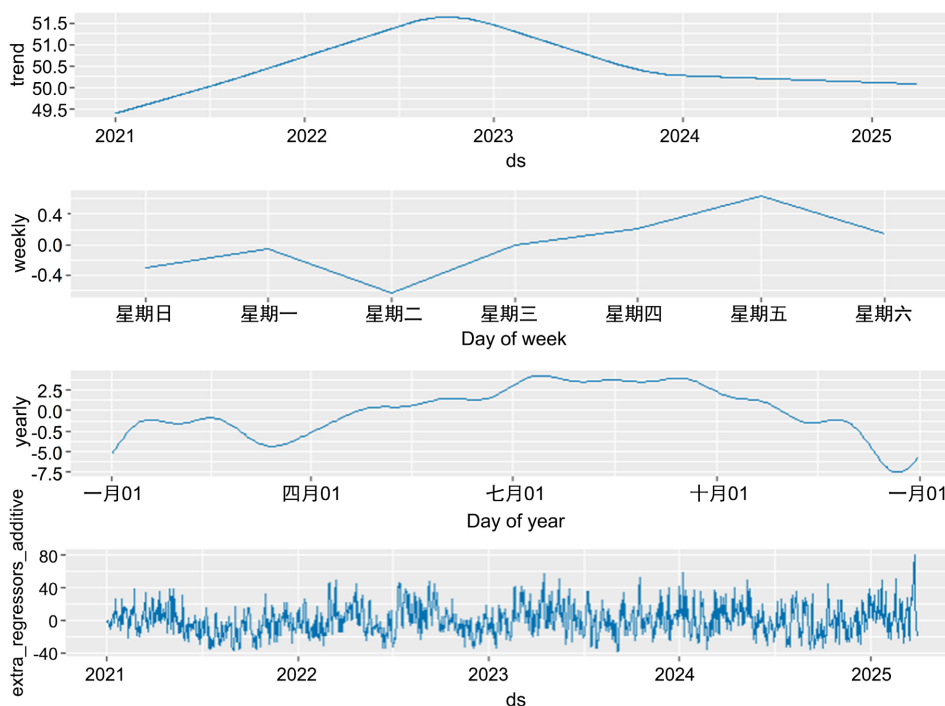


Figure 5. Main components of AQI time series decomposition

图 5. AQI 时间序列分解的主要成分

由图 5 可以看出, Prophet 模型将时间序列分解为四个主要成分: (1) 趋势, 图 5 上部分显示了模型识别出的长期趋势, 可以看出 2021 年至 2025 年 AQI 先上升后略有下降; (2) 周季节性, 图 5 第二部分展示了 AQI 一周内的模式, 星期二最低, 星期五最高; (3) 年季节性, 图 5 第三部分呈现了 AQI 一年内的周期性变化, 在一月和四月附近达到了低谷, 在七月和九月左右达到了顶峰; (4) 额外回归量的加性效应, 图 5 第四部分展示了外部回归变量对目标序列中未被趋势项和周期性成分解释的残差波动的解释能力, 可以看出这部分贡献了大量波动, 说明外部因素对 AQI 的每日变化有显著影响。

图 6 体现的是 Prophet 在测试集上的预测值(蓝色线)与真实的 AQI(红点)的对比, 浅蓝色区域是 80% 置信区间。可以看出, 模型大致捕捉到了测试集 AQI 的波动趋势, 大部分真实值都落在了置信区间。计算该模型的 MAE 为 7.30, RMSE 为 8.79, R^2 为 0.9062。

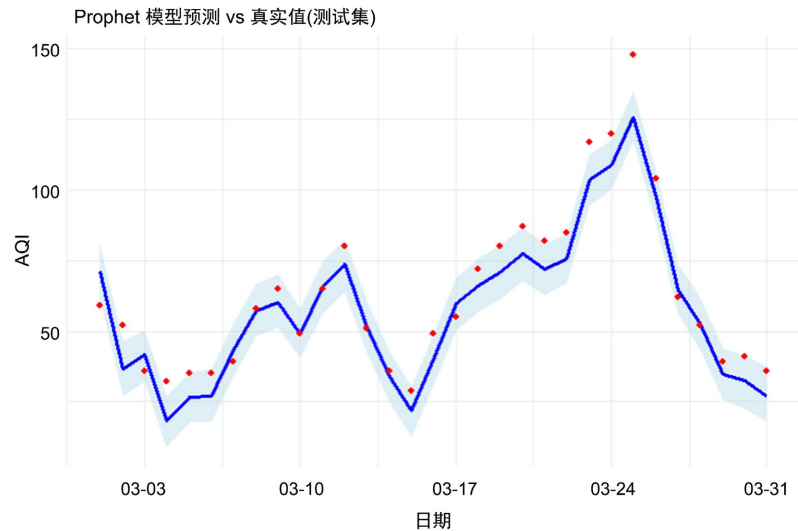


Figure 6. Prediction results and confidence intervals of the prophet model (Test Set)
图 6. Prophet 模型的预测结果与置信区间(测试集)

3.3. 单模型预测结果对比

对比前文所做的单模型的 MAE、RMSE、 R^2 , 如表 2 所示, 可以看出单模型中, BP 神经网络的拟合优度最高, R^2 高达 0.9907, MAE 和 RMSE 也比较小, 而 Prophet 的拟合优度比较一般, R^2 仅有 0.9062, RMSE 也高达 8.97。

Table 2. Comparison of prediction results of single models
表 2. 单模型预测结果对比表

模型	MAE	RMSE	R^2
XGBoost 模型	3.32	7.62	0.9295
BP 神经网络	1.98	2.77	0.9907
双向长短期记忆网络	2.46	3.21	0.9875
Peophet 模型	7.30	8.97	0.9062

4. 基于多模型协同优化的 AQI 预测

在单一模型预测基础上, 为进一步提升空气质量预测的准确性与稳定性, 本文引入多种模型集成与

组合策略。通过融合时间序列分解、机器学习与深度学习方法，构建多个协同优化的组合模型，探索其在福州市空气质量预测中的应用效果与性能差异。

4.1. 基于 STL 分解与 XGBoost 的协同预测

1) 模型构建

在传统机器学习模型中，输入变量往往仅包括污染物浓度、气象因素等静态特征，未能充分考虑空气质量数据本身所具有的时序结构特征，如周期性与趋势性波动。为提升预测精度，本文引入 STL 分解，将原始 AQI 序列拆分为趋势项(trend)、季节项(seasonal)与残余项(residual)，并据此构建两种与 XGBoost 模型融合的协同优化方案：

(1) STL-XGBoost 模型：该模型首先借助 STL 分解方法对原始 AQI 时间序列进行分解，将其拆解为趋势项、季节项以及残余项。为了更好地剥离周期性波动对预测模型的干扰，选取去除季节成分后的 AQI 序列进行训练。具体而言，模型以 PM2.5、PM10、NO₂、SO₂、O₃、湿度作为输入变量，使用 XGBoost 回归算法拟合去季节化的 AQI 值。在完成预测后，再将原始时间序列中的季节项加回到预测结果中，从而恢复对完整 AQI 的估计。这一建模策略的核心思想在于先行滤除周期性波动，使模型更专注于长期变化趋势与短期波动，从而提升预测的稳定性和泛化能力。

(2) 在传统 STL 分解与回归建模的基础上，本研究进一步考虑增强版 STL-XGBoost 模型策略。先对 AQI 时间序列进行 STL 分解，得到趋势项与季节项；接着将趋势、季节分量以及外部污染物特征共同构建输入变量，并作为整体输入特征集引入 XGBoost 模型，直接对原始 AQI 序列进行建模与预测。该策略的核心在于，在引入外部驱动因子的同时，充分融合了 AQI 序列自身的结构性信息，从而提升模型对长期趋势演化和季节周期波动的刻画能力，增强预测结果的稳定性与准确性。

2) 模型预测

图 7 和图 8 分别是 STL-XGBoost 模型和增强版 STL-XGBoost 模型在测试集上的效果，可以看出模型在该时间段内真实值与预测值有较高的一致性，能较好捕捉 AQI 变化趋势。STL-XGBoost 模型的 MAE 值为 6.71，RMSE 值为 8.21，R² 为 0.9182；增强版 STL-XGBoost 模型的 MAE 值为 3.42，RMSE 值为 7.39，R² 为 0.9337。通过融合外部污染物特征以及 AQI 时间序列的结构性信息，可以进一步提升预测结果的稳定性和准确性。

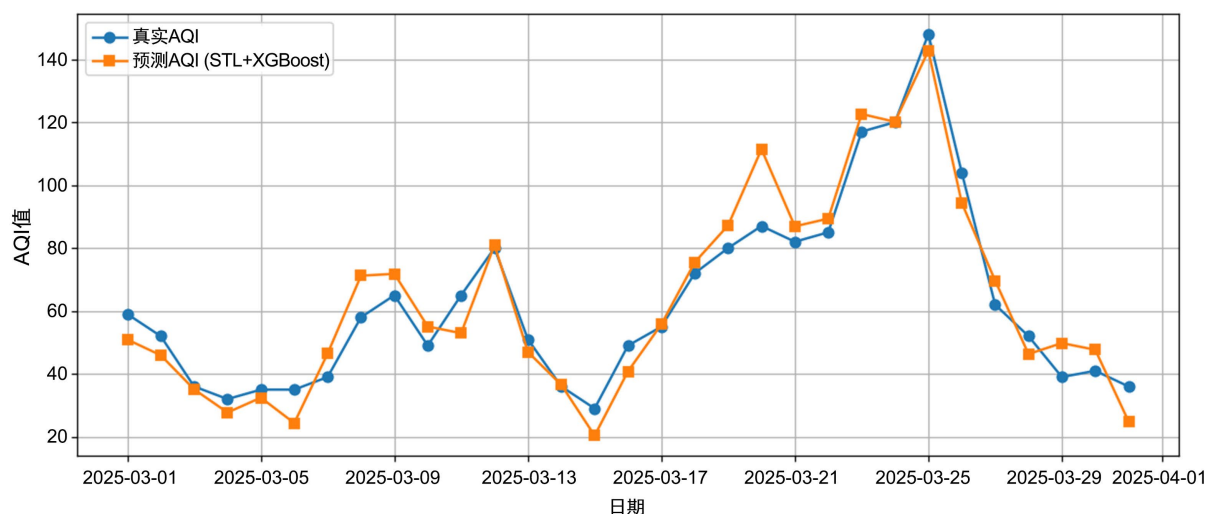


Figure 7. Prediction results of the STL-XGBoost model

图 7. STL-XGBoost 模型预测结果

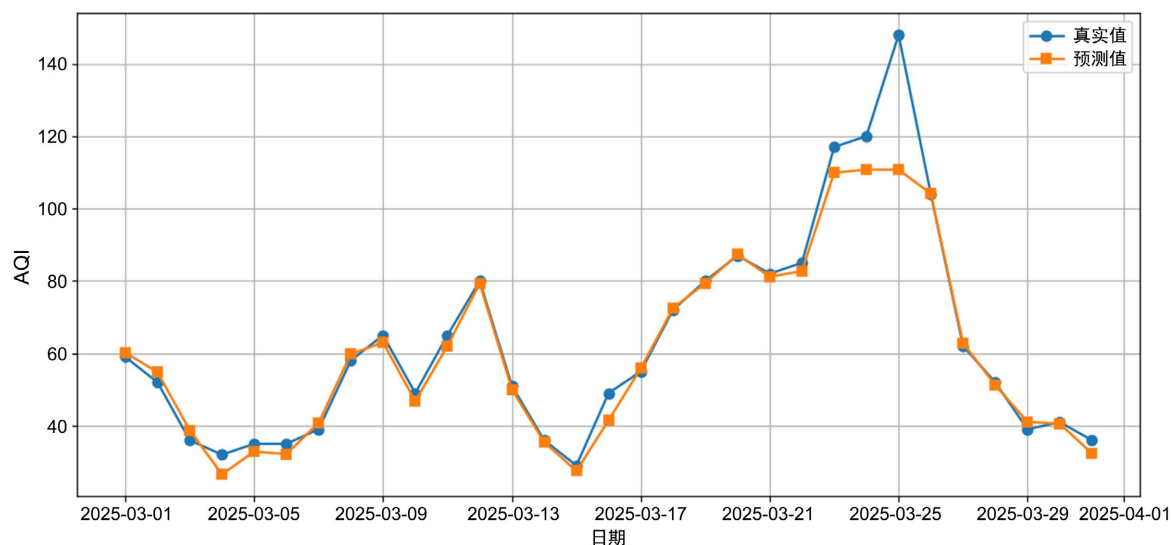


Figure 8. Prediction Results of the Enhanced STL-XGBoost Model

图 8. 增强版 STL-XGBoost 模型预测结果

4.2. 基于 BiLSTM 与 STL-XGBoost 的多阶段协同预测模型构建

1) 模型构建

在 STL-XGBoost 基础上进一步优化趋势项建模方式, 采用 BiLSTM 模型替代传统机器学习模型处理趋势部分, 利用其对时序依赖性和非线性特征的敏感性, 提高对趋势变化的捕捉能力。季节项与残差部分仍由 XGBoost 完成拟合, 从而实现多模型模块化协同建模。

2) 模型预测

图 9 显示的是基于 BiLSTM 与 STL-XGBoost 的多阶段协同预测模型在测试集上的效果, 可以看出模型在该时间段内真实值与预测值有较好的一致性, 能较好地捕捉 AQI 变化趋势。该模型的 MAE 值为 3.66, RMSE 值为 7.73, R^2 为 0.9275。

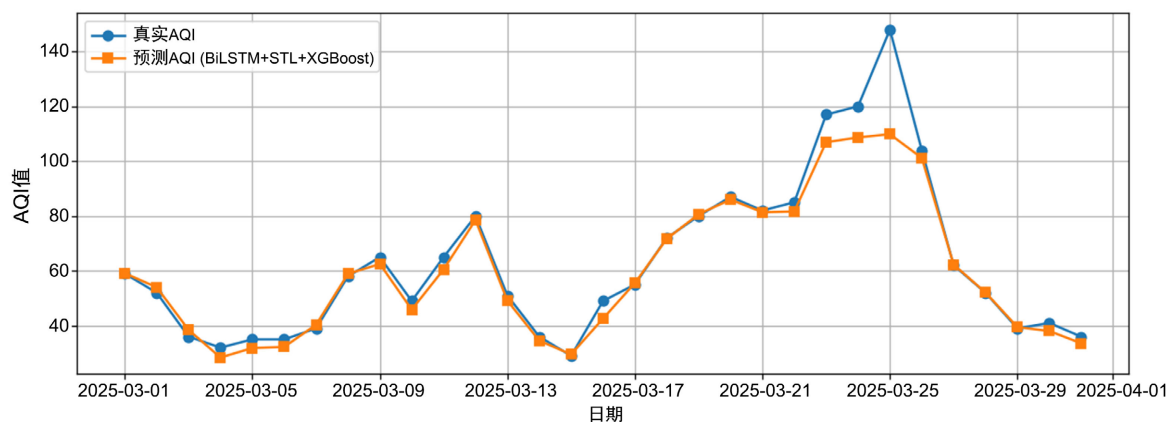


Figure 9. Prediction Results of the BiLSTM-STL-XGBoost Model

图 9. BiLSTM-STL-XGBoost 模型预测结果

4.3. 基于 BP 与 BiLSTM 的层次化 Stacking 融合预测模型设计

1) 模型构建

为进一步提升 AQI 预测模型的鲁棒性与综合性能,本文引入 Stacking 集成学习思想,构建了一种融合 BP 神经网络与 BiLSTM 的多模型协同优化框架。该方法通过集成多个具有异质结构的预测模型,从静态与动态两个角度充分挖掘数据中的时序特征与变量关系,有效提升了预测精度。

2) 模型预测

图 10 显示的是基于 B 神经网络与 BiLSTM 的层次化 Stacking 融合预测模型在测试集上的效果,可以看出模型在该时间段内真实值与预测值有较好的一致性,能较好地捕捉 AQI 变化趋势。该模型的 MAE 值为 2.13, RMSE 值为 2.78, R^2 为 0.9906。

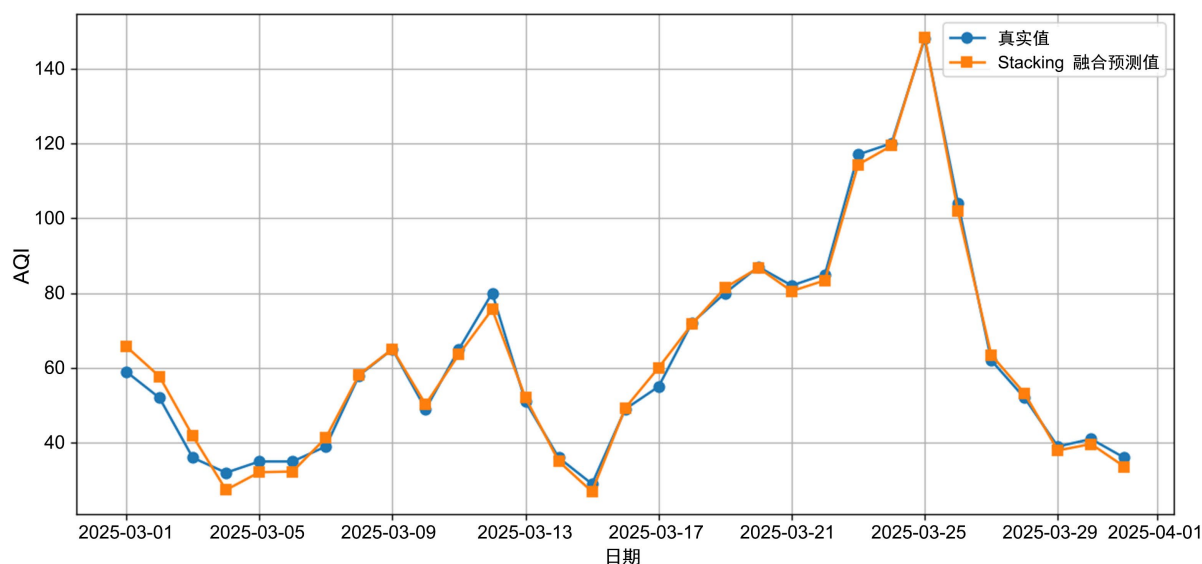


Figure 10. Prediction results of the BP-BiLSTM-Stacking model

图 10. BP-BiLSTM-Stacking 预测结果

4.4. 组合模型预测结果对比

表 3 为本文研究的四个组合模型在测试集上的关键评估指标(MAE, RMSE, R^2)。基于 BP 与 BiLSTM 的层次化 Stacking 融合模型以 MAE = 2.13、RMSE = 2.78 和 R^2 = 0.9906 的综合优势居于首位,表明其通过异质模型集成实现了对 AQI 时序特征的深度挖掘。增强版 STL-XGBoost 模型的效果相较于 STL-XGBoost 模型有明显的提升,并且其通过结构化时序特征与外部变量融合,较原始 XGBoost 的 RMSE 降低 3.0% (7.62→7.39),验证了时序信息精细化建模的必要性。BiLSTM-STL-XGBoost 模型虽未能超越 BiLSTM 模型,但为多阶段协同建模提供了可行思路。

Table 3. Comparison of prediction results of ensemble models

表 3. 组合模型预测结果对比表

模型名称	MAE	RMSE	R^2
STL-XGBoost	6.71	8.21	0.9182
增强版 STL-XGBoost	3.42	7.39	0.9337
BiLSTM-STL-XGBoost	3.66	7.73	0.9275
BP-BiLSTM-Stacking	2.13	2.78	0.9906

5. 基于支持向量机模型的空气质量等级分析

为了对 SVM 模型效果进行评估，给出模型优化前的混淆矩阵如图 11：

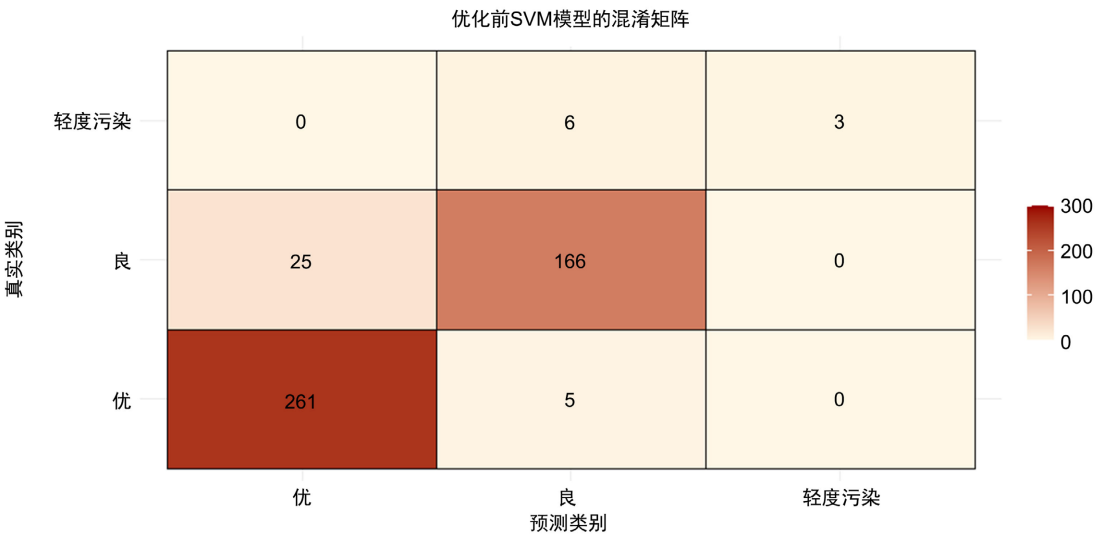


Figure 11. Confusion matrix of the SVM model before optimization

图 11. 优化前 SVM 模型的混淆矩阵

可以看到优化前模型的整体准确度为 92.27%，其中，真实类别为优的被正确预测的概率为 98%，真实类别为良的被正确预测的概率为 87%，真实类别为轻度污染的被正确预测的概率仅有 33%，对多数类（“优”，“良”）的预测准确率高，但对少数类（“轻度污染”）预测准确度很低，说明模型难以识别这类样本，故对模型进行优化。

得到优化后模型的混淆矩阵如图 12：

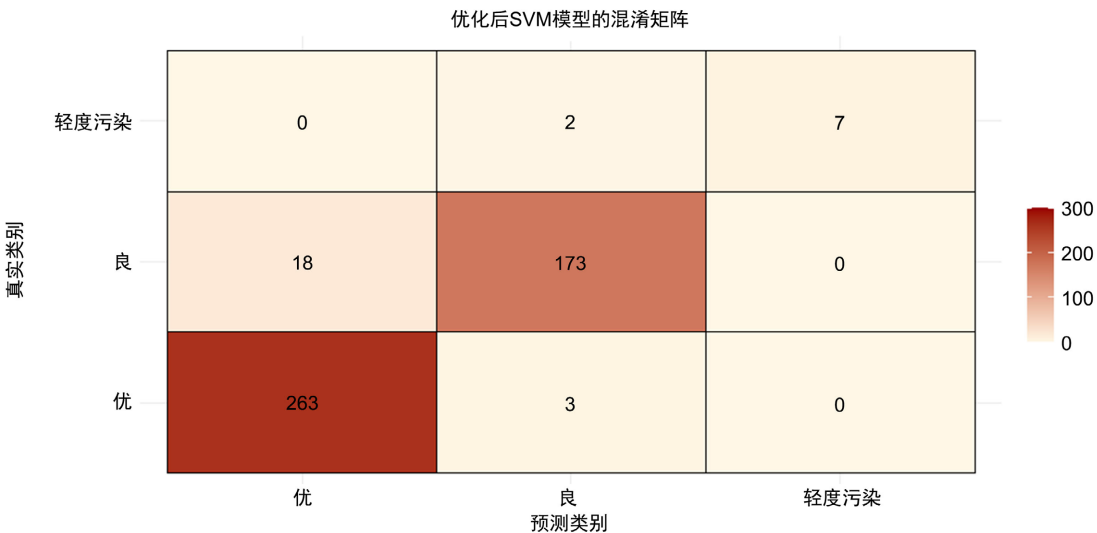


Figure 12. Confusion matrix of the SVM model after optimization

图 12. 优化后 SVM 模型的混淆矩阵

可以看到优化前模型的整体准确度达 95.06%，比优化前有显著上升，其中，真实类别为优的被正确

预测的概率为 99%，准确度几乎完美，真实类别为良的被正确预测的概率为 91%，也有所提高，而真实类别为轻度污染的被正确预测的概率从 33%显著提高到了 78%，说明优化后的 SVM 模型在 AQI 分类任务上表现优异，特别是通过参数调优显著提升了对少数类的识别能力[9]。

6. 结论与建议

6.1. 结论

(1) 福州市空气质量现状与特征：2021 年至 2024 年，福州市空气质量总体处于优良水平，但 AQI 存在波动，轻度污染偶有发生，表明空气质量仍有提升潜力。

(2) 单一模型预测性能：在选取的 XGBoost、BP 神经网络、BiLSTM 和 Prophet 四种单一模型中，BP 神经网络在福州市 AQI 预测任务上表现最佳，具有最高的拟合优度和较低的预测误差，而 Prophet 模型表现相对一般。

(3) 多模型协同优化效果：基于 BP 神经网络和 BiLSTM 的层次化 Stacking 融合模型综合性能最优，而增强版 STL-XGBoost 也通过融合时序分解信息或残差建模，相较于基础模型有明显改进。

(4) SVM 分类性能：优化后的 SVM 模型在空气质量等级分类任务上表现优异，整体准确率达 95.06%，且显著改善了对“轻度污染”这一少数类别的识别准确率，证明了其在处理类别不平衡数据和进行空气质量状态评估方面的有效性。

6.2. 建议

(1) BP 神经网络与 BP-BiLSTM-Stacking 模型在 AQI 预测中的表现最好，福州市相关部门可考虑将其应用到日常空气质量预报业务中，为市民健康防护和政府决策提供更精准的信息。

(2) 优化后的 SVM 模型能有效区别不同空气质量等级，可将其用于构建空气质量预警系统，及时发布健康提示和污染管控建议，特别是在预测到可能出现轻度及以上污染时。

(3) 结合相关性分析结果和高精度预测模型，应同时重点关注对 AQI 影响较大的污染物(如 PM_{2.5}、PM₁₀、O₃)，制定更具针对性的减排措施和应急预案，助力“双碳”目标实现。

参考文献

- [1] 李子熠, 张天宇, 李鸿强. 基于 XGBoost 和 LSTM 组合模型的 PM_{2.5} 浓度预测[J]. 河北建筑工程学院学报, 2023, 41(4): 219-223.
- [2] 王子航, 王子睿, 柳卓行, 杨娟. 基于 LSTM 多步预测模型的空气质量预测与预警[J]. 应用数学进展, 2023, 12(12): 5057-5071.
- [3] Smola, A.J. and Schölkopf, B. (2004) A Tutorial on Support Vector Regression. *Statistics and Computing*, **14**, 199-222. <https://doi.org/10.1023/b:stco.0000035301.49549.88>
- [4] Chen, T. and Guestrin, C. (2016) XGBoost: A Scalable Tree Boosting System. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, San Francisco, 13-17 August 2016, 785-794. <https://doi.org/10.1145/2939672.2939785>
- [5] 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>
- [6] Wolpert, D.H. (1992) Stacked Generalization. *Neural Networks*, **5**, 241-259. [https://doi.org/10.1016/s0893-6080\(05\)80023-1](https://doi.org/10.1016/s0893-6080(05)80023-1)
- [7] 纪峰, 陈荔, 李占利. 基于 STL 文件的模型及应用[J]. 长安大学学报(自然科学版), 2006, 26(1): 104-107.
- [8] 常恬君, 过仲阳, 徐丽丽. 基于 Prophet-随机森林优化模型的空气质量指数规模预测[J]. 环境污染与防治, 2019, 41(7): 758-761, 766.
- [9] 李英英, 纪昌杰. 基于信息熵加权去噪的半监督 SVM 分类器[J]. 电脑知识与技术, 2013, 9(25): 5705-5707, 5715.