

# 基于误差修正的多模型融合空气质量预测

李淑婷\*, 吕卫东#, 朵俞霖

兰州交通大学数理学院, 甘肃 兰州

收稿日期: 2023年8月12日; 录用日期: 2023年9月6日; 发布日期: 2023年9月13日

## 摘要

由于空气质量指数序列的复杂性和非线性, 创造性提出将误差修正与多模型融合相结合的预测方法。首先, 利用变分模态分解(VMD)将原始不平稳的空气质量指数序列分解为多个不同时间尺度的平稳固有模态分量; 其次, 使用改进的SA-BiGRU用于空气质量预测, 叠加各个子序列得到空气质量指数初始预测值, 实现了长距离时间模式的特征提取; 最后, 在初始预测模型的基础上建立误差修正模型, 通过SVR预测模型得到训练集的预测误差, 与初步预测结果用加法器合并, 增强模型的表达能力。与单一模型BP、LSTM以及混合模型VMD-LSTM、VMD-GRU、VMD-BiGRU、VMD-SA-BiGRU模型对比, 其预测的平均绝对误差分别降低了32.037%、24.581%、18.134%、11.448%、9.320%、5.802%。实验结果表明, VMD-SA-BiGRU-SVR模型在对空气质量指数进行预测时具有更高的精度, 预测性能更优异。

## 关键词

空气质量预测, 变分模态分解, 自注意力机制, 双向门控循环网络, 误差修正

# Multi Model Fusion Air Quality Prediction Based on Error Correction

Shuting Li\*, Weidong Lyu#, Yulin Duo

School of Mathematics and Physics, Lanzhou Jiaotong University, Lanzhou Gansu

Received: Aug. 12<sup>th</sup>, 2023; accepted: Sep. 6<sup>th</sup>, 2023; published: Sep. 13<sup>th</sup>, 2023

## Abstract

Due to the complexity and nonlinearity of the air quality index, a creative prediction method combining error correction with multi model fusion is proposed. Firstly, variational mode decomposition (VMD) is used to decompose the original unstable air quality index sequence into multiple

\*第一作者。

#通讯作者。

stationary natural mode components at different time scales. Secondly, the improved SA-BiGRU was used for air quality prediction, and the initial predicted values of the air quality index were obtained by overlaying various sub-sequences, achieving feature extraction for long-distance time patterns. Finally, an error correction model is established on the basis of the initial prediction model, and the prediction error of the training set is obtained through the SVR prediction model. The initial prediction results are combined with an adder to enhance the model's expression ability. Compared with the single model BP, LSTM and the mixed model VMD-LSTM, VMD-GRU, VMD-BiGRU and VMD-SA-BiGRU, the average absolute error of its prediction decreased by 32.037%, 24.581%, 18.134%, 11.448%, 9.320% and 5.802% respectively. The experimental results show that the VMD-SA-BiGRU-SVR model has higher accuracy and better prediction performance in predicting air quality index.

## Keywords

Air Quality Prediction, VMD, Self-Attention, BiGRU, Error Correction

Copyright © 2023 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

空气污染被认为是影响人口和环境的主要因素之一[1]。随着我国经济快速增长,工业化进程加快,发展的同时也造成了空气污染问题。目前,已有大量的文献对空气质量预测模型进行了研究,主要可以分为三类。第一类是经典的统计计量模型,如ARIMA模型,王建军等人构建了ARIMA(1,1,1)模型对苏州市空气质量指数进行预测[2]。但这类模型一般都要满足数据平稳性检验等假设检验,而空气质量由于受到多种因素的影响,得到的数据往往是非平稳、非线性的,传统的统计模型对于这类数据的处理会相应复杂一些。第二类是单一机器学习预测模型,随着计算机模式、人工智能技术的高速发展,机器学习被广泛应用到空气质量预测的任务中[3]。然而随着社会经济的发展,监测系统能够得到高维数据,其复杂性不断增加,且从数据量上来看更加庞大,从时间序列的特征来看,数据变得非平稳、非线性[4]。此时单一预测模型对空气质量进行预测时有自身的局限性,比如只能考虑时间或空间的相关性的一种,而且缺失的部分数据特征也可能会弱化预测效果。很多学者考虑用新的工具、方法对数据进行处理从而进行预测,这时就提出第三类组合预测的模型,组合模型利用了不同模型的优点,可以获得更好的预测精确度、更小的误差和更优的鲁棒性。而针对时间序列非线性、非平稳的特征,汪寿阳教授提出“先分解后集成”的核心思想,该方法是将传统的统计预测模型与人工智能技术相结合的方法论,首先将复杂系统进行多尺度分解,接着使用第一类或第二类预测方法估计出预测对象的趋势,最后对预测结果进行结合[5]。史学良等人提出使用EEMD对AQI序列分解预测的组合预测模型EEMD-LSTM,结果表明组合模型比单一模型的泛化能力更强、预测精度更高[6];梁涛等人提出一种基于自适应噪声的完整集成经验模态分解的组合预测模型对PM10浓度进行预测,其模型显示出了较高精度[7]。

经分析发现,空气质量预测模型对精度有较高的要求,但以往的文献中很少有将误差修正用于空气质量预测模型中。本文基于模型预测误差,进行融合算法改进。通过分析空气污染物和气象两方面影响因素,使用变分模态分解(VMD)将原始不平稳的空气质量指数序列进行分解为多个平稳的固有模态分量;并结合双向门控循环单元(Bi-GRU)网络的同时引入自注意力机制,提出基于自注意力机制的

VMD-SA-BiGRU 初始预测模型；再建立结合误差修正的支持向量回归(SVR)模型进行预测得到误差预测值，将初始 AQI 预测值与误差预测值用加法器合并得到最终预测结果，以此探究空气质量变化规律，为污染治理提供理论依据。

## 2. 模型与方法

### 2.1. VMD 分解原理

变分模态分解(Variational Modal Decomposition, VMD)是一种信号分解技术，它最主要的特征是自适应、完全非递归，将实际值信号即非平稳信号进行频域分析并将其分解为多个得到若干独立模态分量的过程就是变分问题的构造与求解过程[8] [9]。

假设每个本征模态函数(Intrinsic mode function, IMF)是具有中心频率  $\omega_k$  的有限带宽函数，变分问题是寻找将实值序列  $X_n$  分解为  $k$  个 IMF 分量  $u_k(t)$ ，使得每个模态的估计带宽之和最小，约束条件为每个 IMF 分量之和与原始输入信号  $s(t)$  相等。具体模型的构造与求解如下：

模型构造：

- 1) 通过希尔伯特变换，计算得到每个 IMF 分量  $u_k(t)$  的解析信号及单边谱：

$$\left( \delta(t) + \frac{j}{\pi t} \right) * u_k(t) \quad (1)$$

式中， $\delta(t)$  为冲激函数。

- 2) 将每个模态的解析信号与算子  $e^{-j\omega_k t}$  相乘，进而将  $u_k(t)$  的谱频调至相应的基频带上：

$$\left[ \left( \delta(t) + \frac{j}{\pi t} \right) * u_k(t) \right] e^{-j\omega_k t} \quad (2)$$

- 3) 通过高斯平滑处理(梯度的  $L^2$  范数)对第二步信号解调得：

$$\left\| \partial_t \left[ \left( \delta(t) + \frac{j}{\pi t} \right) * u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 \quad (3)$$

- 4) 估计各个 IMF 的带宽使得其之和最小，建立式(3)的约束变分模型：

$$\begin{aligned} \min_{\{u_k, \omega_k\}} & \left\{ \sum_k \left\| \partial_t \left[ \left( \delta(t) + \frac{j}{\pi t} \right) * u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 \right\} \\ \text{s.t.} & \sum_k u_k = s(t) \end{aligned} \quad (4)$$

模型求解：

- 1) 为求得约束变分问题的最优解，通过引入拉格朗日乘子  $\tau(t)$  和二阶惩罚因子  $\alpha$ ，将该问题转化为非约束变分问题。

$$\begin{aligned} L(\{u_k\}, \{\omega_k\}, \tau) &= \alpha \sum_k \left\| \partial_t \left[ \left( \delta(t) + \frac{j}{\pi t} \right) * u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 \\ &+ \left\| s(t) - \sum_k u_k(t) \right\|_2^2 + \left\langle \tau(t), s(t) - \sum_k u_k(t) \right\rangle \end{aligned} \quad (5)$$

其中，拉格朗日乘子  $\tau(t)$  的作用是保证约束条件的严格性可以确保约束条件的严格性；二阶惩罚因子  $\alpha$  的作用是确保在存在高斯噪声环境下信号重构的准确性。

2) 利用交替方向乘子法(Alternating Direction Method Of Multipliers, AMMD)迭代更新各分量与中心频率, 得到无约束变分模型的鞍点为模型的最优解, 分量表示为:

$$\hat{u}_k^{n+1}(\omega) = \frac{\hat{s}(\omega) - \sum_{i \neq k} \hat{u}_i(\omega) + \hat{\tau}(\omega)/2}{1 + 2\alpha(\omega - \omega_k)^2} \quad (6)$$

式中,  $\omega$  表示频率,  $\hat{u}_k^{n+1}$ 、 $\hat{s}(\omega)$ 、 $\hat{\tau}(\omega)$  分别为对应函数的傅里叶变换。

3) 直到满足迭代停止条件, 得到  $k$  个分量为变分模态分解如下所示:

$$\frac{\sum_k \|\hat{u}_k^{n+1} - \hat{u}_k^n\|_2^2}{\|\hat{u}_k^n\|_2^2} < \varepsilon \quad (7)$$

式中,  $\hat{u}_k^n$  为第  $n$  次迭代分量,  $\hat{u}_k^{n+1}$  为第  $n+1$  次迭代分量,  $\varepsilon$  为误差。

## 2.2. 双向门控循环单元网络 BiGRU

门控循环单元(Gated Recurrent Unit, GRU)网络是一种比长短期记忆网络(LSTM)更加简单的循环神经网络, 在 LSTM 的基础上作了结构简化的改进, 在解决长期依赖的问题外, 还有结构简单、参数少和运行速度快的优点[10] [11] [12]。虽然 GRU 内部也是使用门来控制信息的传递, 但 GRU 网络去除了细胞状态的传递, 其历史信息的传递通过重置门(Reset gate)和更新门(Update gate)控制。重置门学习上一时刻与当前时刻信息组合权重, 更新门学习上一时刻信息的保留权重, 取值在 0 到 1 之间, 当值越接近 1, 表明更趋向于对上一时刻传入信息的留存, 反之则表示对上一时刻信息更倾向舍弃。当更新门设置为 0, 重置门设置为 1, 则表示上一个时刻的信息完全传递到当前时刻。GRU 网络结构如图 1 所示。

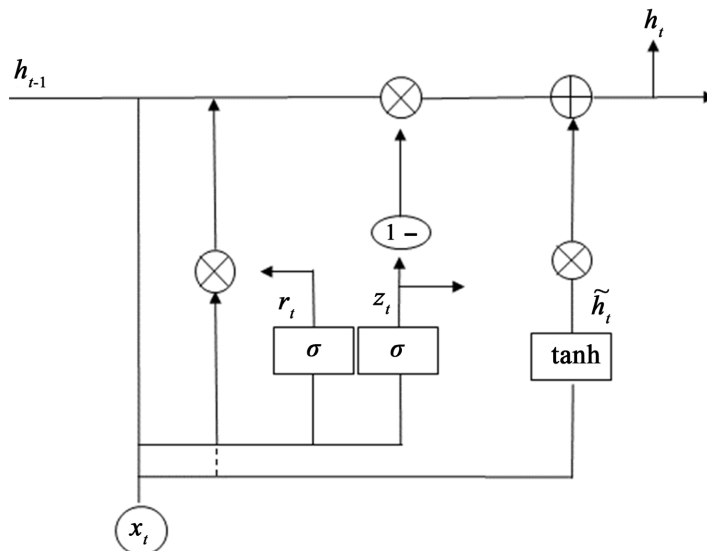


Figure 1. Schematic diagram of GRU network structure

图 1. GRU 网络结构示意图

重置门和更新门的计算公式为:

$$r_t = \sigma(W_r x_t + U_r h_{t-1} + b_r) \quad (8)$$

$$z_t = \sigma(W_z x_t + U_z h_{t-1} + b_z) \quad (9)$$

其中  $r_t$  表示重置门, 用来决定之前的信息有哪些需要保留,  $z_t$  表示更新门, 它在隐藏层状态更新时会决定哪些信息会被传递到下个时刻。

GRU 网络是按照时间顺序正向的方式提取信息特征[13]。当发生异常状态后, 数据会出现一种恢复趋势的现象, 如果反向利用这种趋势将会更加敏感的捕捉异常状态。双向门控循环单元网络 BiGRU 是由前向 GRU 与后向 GRU 组合而成, 其可以对时间序列正向和反向的特征信息进行同时提取。

BiGRU 网络的具体结构如图 2 所示, BiGRU 具有两个独立的隐藏层, 其中前向隐藏层学习历史信息, 逆向隐藏层学习未来信息, 最后整合历史与未来信息作为输出结果。

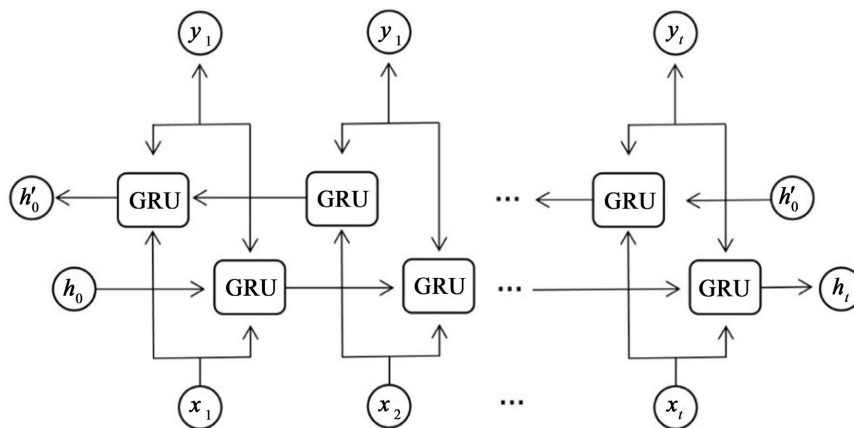


Figure 2. Schematic diagram of BiGRU network structure  
图 2. BiGRU 网络结构示意图

### 2.3. Self-Attention 机制

注意力机制是受到了人类视觉注意力机制的启发, 人类的注意力机制是在长期进化过程中形成特有的大脑信号处理机制, 即全局扫描物体整体状况判断哪部分需要重点关注并快速定位核心部分投入更多的注意力获取细。而自注意力机制首次提出是主要解决在实际训练中, 神经网络接受大小不一且有相互关系的向量时, 无法使用向量之间的相互关系而导致模型训练效果不佳的问题[14]。自注意力的结构是由查找(query)、键(key)和值(value)组成, 且三者相等, 它能把输入序列的不同信息联系起来, 更擅长捕捉重要数据[15]。该机制主要分为三步, 首先计算每个 query 和 key 之间的相似度, 以获得相应的 value; 然后使用 Softmax 函数对权重进行归一化, 获得权重系数; 最后将权重系数和相应值加权求和, 以获得最终的注意力值。其计算公式如下:

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (10)$$

其中,  $d_k$  为 key 的维度。

### 2.4. SVR

SVR 算法是在支持向量机 SVM 分类的基础上, 引入核函数和损失函数, 通过非线性映射函数即核函数将输入数据映射至高维特征空间, 找到最优拟合超平面, 使所有训练样本与该平面的总偏差最小[16][17]。其回归函数表达式如下:

$$f(x) = \omega^T \varphi(x) + b \quad (11)$$

其中,  $\omega$  为权值系数,  $\varphi(x)$  为非线性映射函数,  $b$  为偏移量。

再通过引入松弛变量  $\xi_i$ ,  $\xi_i^*$ , 使用最小化正则分线准则, 将原来的回归问题转化成无约束的二次规划问题, 即

$$\begin{cases} \min \left\{ \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^n (\xi_i, \xi_i^*) \right\} \\ \text{s.t.} \begin{cases} f(x_i) - y_i \leq \varepsilon + \xi_i, i=1, 2, \dots, n \\ y_i - f(x_i) \leq \varepsilon + \xi_i^* \\ \xi_i \geq 0 \\ \xi_i^* \geq 0 \end{cases} \end{cases} \quad (12)$$

其中,  $C$  为惩罚因子。

### 3. VMD-SA-BiGRU-SVR 模型建立

对于空气质量的随机性强、波动性大的特点, 本文采用变分模态分解以及加入自注意力机制的双向门控循环单元网络的组合算法对空气质量指数进行预测。而空气质量的影响因素是多方面的, 我们选取了空气污染物浓度数据和气象数据作为主要影响因素对 AQI 值进行预测。历史的 AQI 值中也蕴含着重要的价值信息, 所以为了降低其序列的非平稳性, 使用 VMD 对历史 AQI 数据进行分解, 得到多组 IMF 分量。将数据集以 8:2 的比例将数据集划分为训练数据集和测试数据集。将训练样本输入 SA-BiGRU 的模型中, 其中 BiGRU 中添加的自注意力机制可以对高维数据参数学习中出现的梯度爆炸与梯度消失问题进行稀释, 将各个子序列和其他影响因素的预测结果叠加求和得到空气质量指数初始预测值。为了进一步提升该模型的预测性能, 本文建立了误差修正模型, 通过 SVR 预测模型得到训练集的预测误差, 将原始数据的输入与 VMD-SA-BiGRU 模型输出的初始预测值之差作为误差模型的输入进行训练, 得到误差预测模型, 再用此误差预测模型对测试集数据进行预测得到误差的预测值, 与初步预测结果用加法器合并得到最终预测结果。通过对比实验, 验证了此误差修正模型能大幅提升模型的预测精度。该方法流程如图 3 所示。

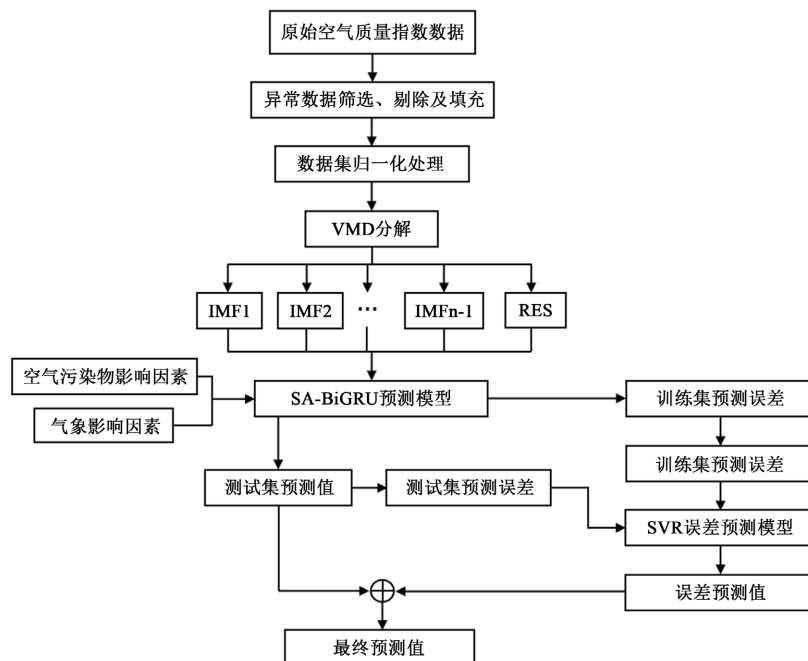


Figure 3. Schematic diagram of the combined model process

图 3. 组合模型流程示意图

## 4. 实例分析

### 4.1. 模型数据来源

本文选取西安市 2021 年 1 月~2022 年 12 月的实际空气质量数据, 其中包括 PM<sub>2.5</sub>、PM<sub>10</sub>、SO<sub>2</sub>、CO、NO<sub>2</sub> 和 O<sub>3</sub> 六方面的污染物浓度, 除 O<sub>3</sub> 外计算各站点 24 h 平均浓度值的算术平均值作为日均浓度, O<sub>3</sub> 采用日最大 8 h 平均浓度值的算术平均值作为日均浓度, 数据来源于中国环境检测总站, 部分数据如表 1 所示; 另一部分是对应地区的气象数据, 其中包括大气温度、大气压、相对湿度、平均风速、最低气温、最高气温、水平能见度和降水量, 部分数据如表 2 所示。

**Table 1.** Xi'an air quality index (Partial)

**表 1.** 西安市空气质量指数(部分)

年份	AQI 指数	PM2.5	PM10	SO <sub>2</sub>	CO	NO <sub>2</sub>	O <sub>3</sub>
2021-01-01	96	71	107	11	1.07	65	13
2021-01-02	166	126	183	19	1.72	79	10
2021-01-03	164	124	179	18	1.6	85	11
2021-01-04	208	158	230	18	1.91	85	6
2021-01-05	144	109	155	15	1.28	63	15

**Table 2.** Meteorological data of Xi'an City (Partial)

**表 2.** 西安市气象数据(部分)

年份	大气温度(°C)	大气压(kPa)	相对湿度(%)	平均风速(m/s)	能见度(km)
2021-01-01	-1.675	735.275	51.5	1.375	10.375
2021-01-02	0.1625	733.7625	46.5	1.875	8.5
2021-01-03	1.675	734.8375	51.625	2	5.9875
2021-01-04	0.4125	735.5875	57.5	1.25	4.875
2021-01-05	1.7875	738.9875	42.125	2	16.7125

原始的 AQI 值数据存在异常值, 如果直接分解其时间序列并进行预测, 则会导致预测误差较大。因此在分解预测之前, 先去除原始数据的异常值。2021 年 1 月~2022 年 12 月某地每日空气质量指数原始 AQI 变化趋势如图 4 所示, 一共 731 个观测值。从图中可以看出, 空气质量指数具有非线性、非平稳的复杂特性。

### 4.2. 数据标准化

由于数据种类和规范单位的不同, 如果直接进行数据分析有可能会造成数据间的差异导致结果误差变大, 不利于后续的研究, 所以本文对原始数据进行 Min-Max 标准化, 消除不同指标量纲和取值范围差异的影响, 使各组数据处于同一量纲, 其计算公式如下:

$$x_{\text{new}} = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad (13)$$

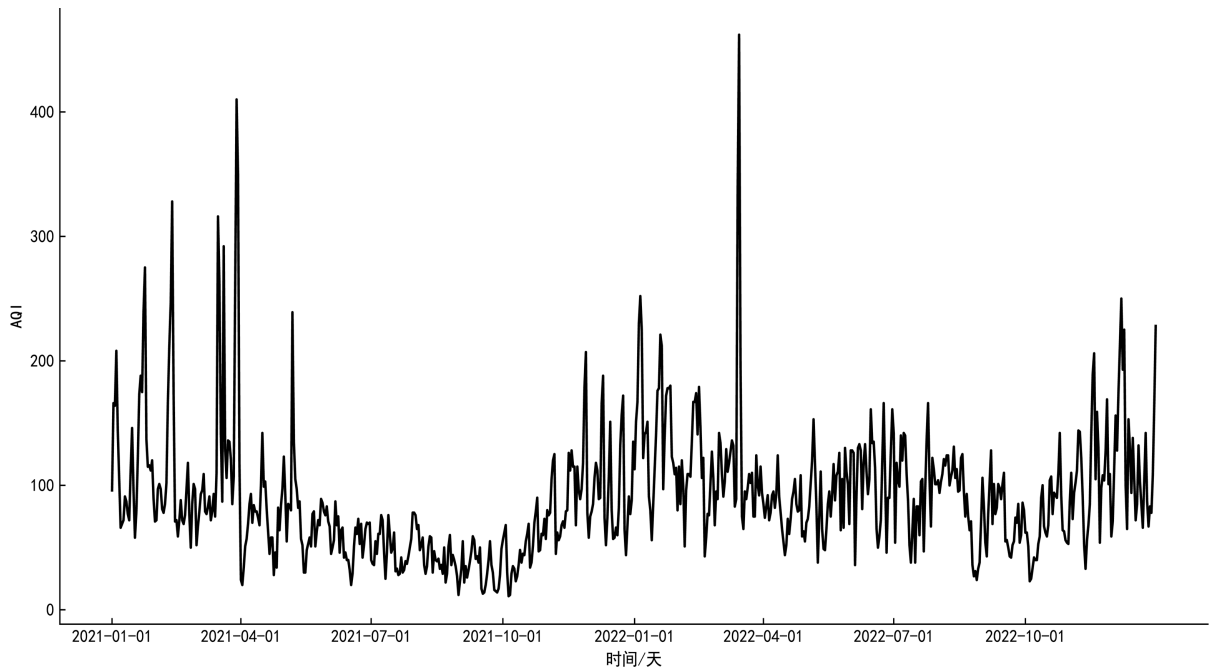


Figure 4. Time series diagram of air quality index  
图 4. 空气质量指数的时间序列图

### 4.3. 评价标准

为评估不同空气质量预测模型的优劣，本文选用平均绝对值百分比误差(Mean Absolute Percentage Error, MAPE)和均方根误差(Root Mean Square Error, RMSE)作为模型的评价指标，计算公式如下：

$$\text{MAPE} = \frac{1}{n} \sum_{i=1}^n \left| \frac{\hat{y}_i - y_i}{y_i} \right| \times 100\% \quad (14)$$

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2} \quad (15)$$

式中， $y_i$  和  $\hat{y}_i$  分别表示第  $i$  次输出的实际值和模型预测值。其中，MAPE 和 RMSE 的值越小，说明模型预测效果越好，反之，模型预测效果相对较差。

### 4.4. 基于 VMD-SA-BiGRU 模型的预测

为了能够充分提取数据中潜在的关联，本文采用 VMD 算法处理空气质量指数数据。首先利用中心频率法确定分解模态数  $K=4$ ；惩罚因子  $\alpha=5000$ ；收敛误差  $\varepsilon=10^{-7}$ 。将原始数据按低频到高频分解为单一频率的 IMF 分量，展现不同尺度的数据信息。VMD 分解算法能够将空气质量指数分解为不同特征成分的 IMF 分量，呈现出的 IMF 分量特征清晰。图 5 为分解的序列图像，分解后各模态的本征函数如图 6 所示，从图中可以看出，IMF1 本征函数的变化趋势与原序列大致相同且变化程度比较缓慢，IMF2 本征函数变化也比较缓慢，剩余分量则为高频分量，变化较快。一般认为，噪声主要集中在高频 IMF 分量上，低频 IMF 分量受噪声影响较小，但删除噪声会大大降低预测的准确性，因此保留了所有的 IMF 成分。经过 VMD 分解后的子序列复杂度明显低于原始序列，有助于模型拟合。





Figure 5. VMD decomposition  
图 5. VMD 分解

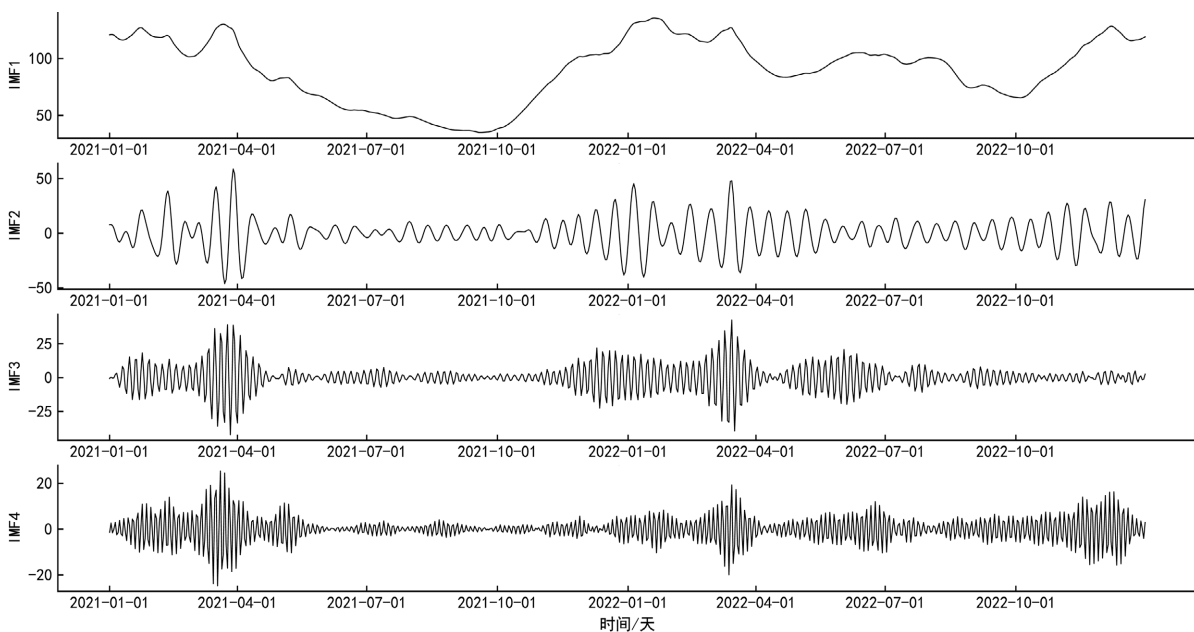


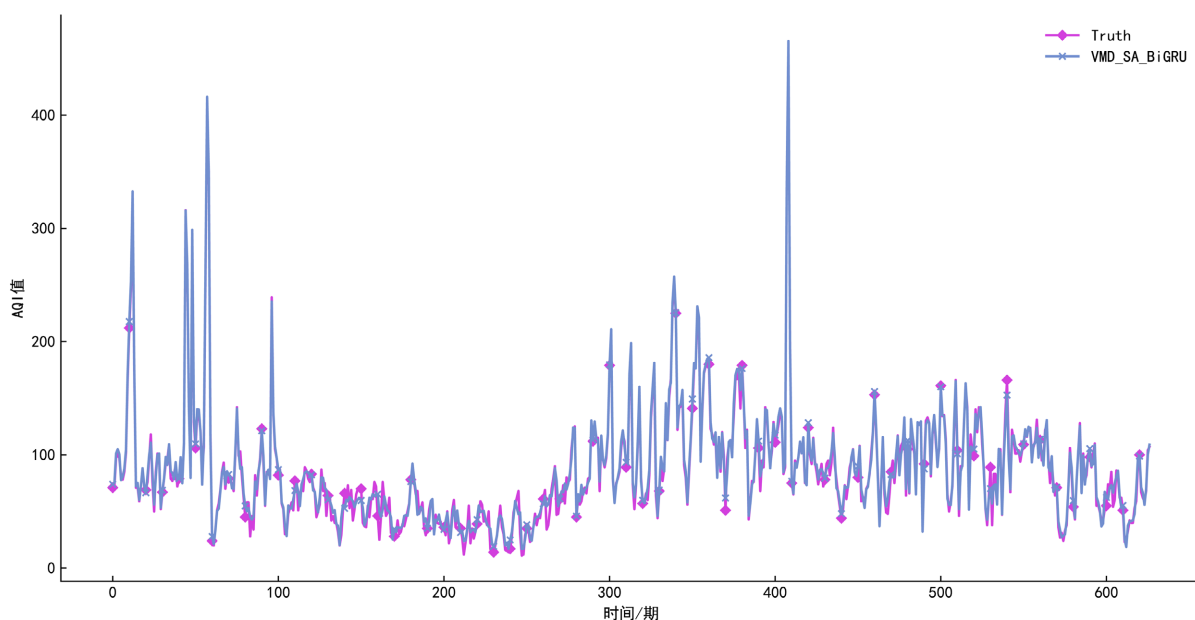
Figure 6. IMF function  
图 6. IMF 函数

将通过信号分解的原始空气质量指数数据与污染物、气象影响因素整合后输入 SA-BiGRU 模型，华东窗口选择 30 期历史信息对训练数据训练、更新参数后，在测试集数据中进行验证。本文建立的 BiGRU 模型网络结构如表 3 所示。

在迭代后输出预测结果如图 7，该图为 VMD-SA-BiGRU 组合模型预测结果与真实值的对比图，从图中我们可以看到空气质量指数预测值与真实值之间的结果十分接近，在除空气质量极端值外的其它时刻该组合模型在数据集上有很好的效果。

**Table 3.** SA-BiGRU model structure  
**表 3.** SA-BiGRU 模型结构

网络层序数	网络结构	参数
第一层	Self-Attention	20
第二层	输入层	20
第三层	BiGRU 层	4
第四层	Dropout 层	0.5
第五层	输出层	1



**Figure 7.** AQI prediction chart  
**图 7.** AQI 预测图

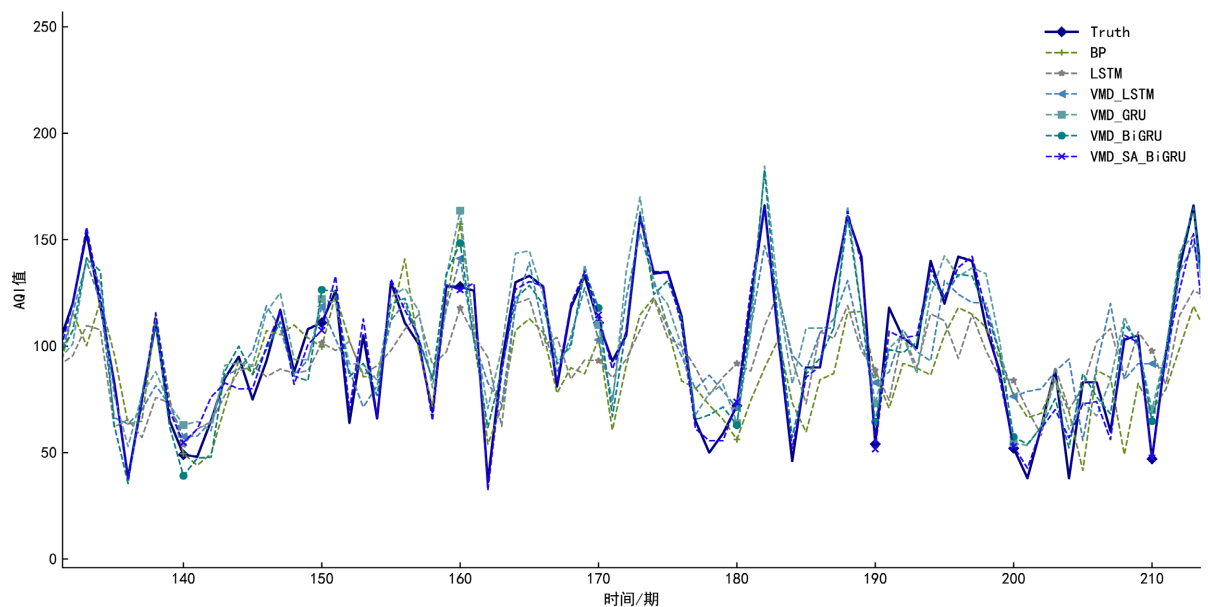
#### 4.5. 对照实验

为了验证本文方法的优越性，我们对比了不同模型下对该空气质量数据预测，其中有 BP、LSTM、VMD-LSTM、VMD-GRU、VMD-BiGRU 两种单一模型和三种不同组合方式的预测模型。

为了更加直观地比较各个模型的优劣，从表 4 即各个模型的 MAPE、RMSE 两个指标的实验结果和图 8 即不同的组合模型预测效果图这两方面可以看出，组合模型 VMD-LSTM、VMD-GRU、VMD-BiGRU 和 VMD-SA-BiGRU 普遍要比单一模型 BP 和 LSTM 表现要好；模型 LSTM 和经过 VMD 分解后的 LSTM 的组合模型的预测结果对比可以看出加入 VMD 分解算法后的模型精度有明显的提升，其中 LSTM 模型参数与 VMD-LSTM 组合模型参数一致，以此验证本文提出的 VMD 算法对数据去噪的有效性；以 VMD 分解为基础的 VMD-LSTM、VMD-GRU 和 VMD-BiGRU 三种组合模型的预测结果可以看出，针对空气质量数据而言 BiGRU 的预测精度是优于 LSTM 和 GRU 的；融入自注意力机制后的 BiGRU 的预测模型是对序列特征的赋予关注度，相比于 VMD-BiGRU 模型进一步提高了模型的预测精度，同时 VMD-SA-BiGRU 模型的波动幅度与真实值最为接近，虽然某些点的预测效果略低于其他模型，但总体预测精度和稳定性较高，拟合度较好。

**Table 4.** Comparison of prediction error indicators for different models  
**表 4.** 不同模型预测误差指标对比

模型	RMSE	MAPE
BP	25.180	35.948
LSTM	21.004	28.492
VMD-LSTM	15.115	22.045
VMD-GRU	11.930	15.359
VMD-BiGRU	10.258	13.231
VMD-SA-BiGRU	8.023	9.713



**Figure 8.** Partial comparison of prediction results of different models  
**图 8.** 不同模型的预测结果部分对比图

#### 4.6. 误差修正

在 VMD-SA-BiGRU 模型中,  $t$  时刻的空气质量序列值为  $Y_t$ , 该模型输出预测结果为  $\hat{L}_t$  出预测, 则最后残差结果可以表示为:

$$e_t = Y_t - \hat{L}_t \quad (16)$$

将 VMD-SA-BiGRU 模型的残差值输入 SVR 模型得到残差预测值, 修正 VMD-SA-BiGRU 模型预测结果。在 SVR 模型中对得到的残差序列结果进行建模, 并得到残差值的预测值  $\hat{N}_t$ 。最终根据 VMD-SA-BiGRU 模型的预测值叠加 SVR 模型输出的残差预测值, 计算出组合模型的预测值  $\hat{Y}_t$ :

$$\hat{Y}_t = \hat{N}_t + \hat{L}_t \quad (17)$$

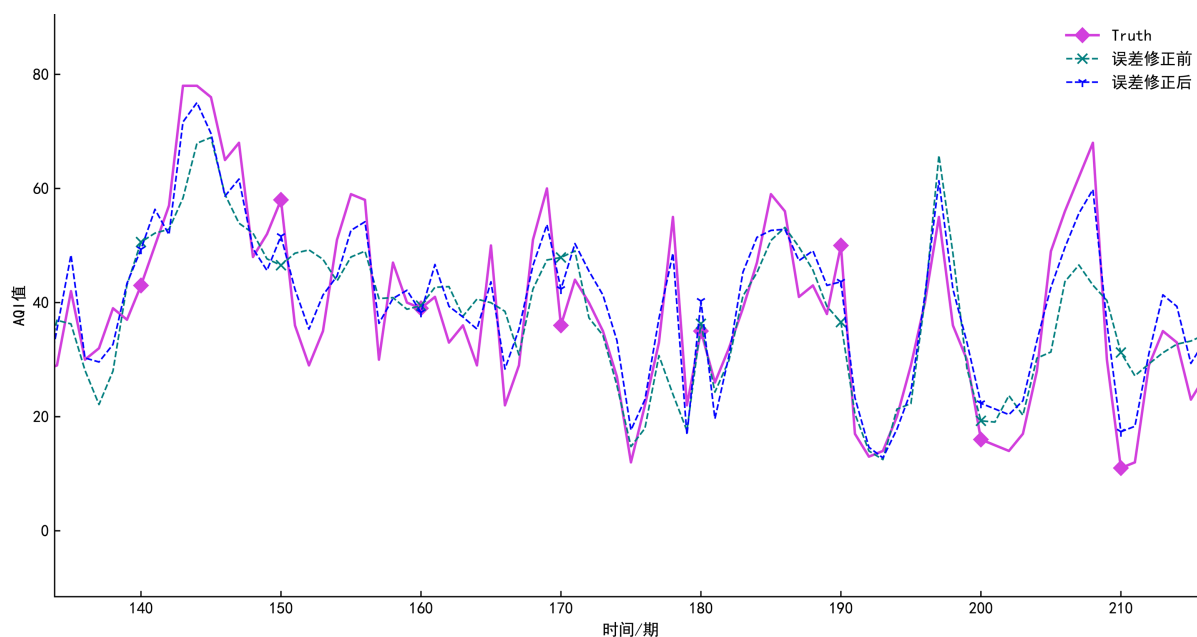
建立误差修正模型进行预测得到误差预测值, 将选择上述实验中最优的模型为基础预测模型输出的初始预测值与误差预测值用加法器合并得到最终结果。为验证加入误差修正后模型的有效性, 我们选择

VMD-SA-BiGRU 与所提的误差修正模型进行对比分析, 模型预测误差指标如表 5 所示, 模型修正前后对比如图 9 所示。从表 5 可以得到, 经过误差修正后的 VMD-SA-BiGRU 模型, 其预测的平均绝对误差低至 3.911%, 相比于未修正的模型降低了 5.802%。

**Table 5.** Comparison of model evaluation indicators before and after error correction

**表 5.** 误差修正前后的模型评价指标对比

模型	RMSE	MAPE
VMD-SA-BiGRU	8.023	9.713
VMD-SA-BiGRU-SVR	4.559	3.911



**Figure 9.** Comparison of predictions before and after error correction

**图 9.** 误差修正前后预测对比图

## 5. 结论

本文提出一种基于 VMD-SA-BiGRU-SVR 的空气质量预测模型, 提高了空气质量指数 AQI 预测精度, 并通过实验得到以下结论: 1) 对于随机性强、波动性大的空气质量指数 AQI 时间序列数据, 引入变分模态分解算法对该数据进行分解, 提取 AQI 特征并降低噪声影响, 提高预测精度。然后结合气象因素与空气污染物因素构成多维特征时间序列对其进行预测; 2) 融入自注意力机制模块, 对多维特征时间序列添加权重, 初步提取时间序列长距离依赖关系, 促进 BiGRU 网络对历史信息的充分提取, 同时减少网络参数, 提高模型的预测精度的同时提高预测速度; 3) 建立了误差修正模型。通过 VMD-SA-BiGRU 组合模型得到训练集的预测误差, 将其与原始模型的输入分别作为误差模型的输出与输入进行训练, 得到误差预测模型, 再用此误差预测模型对测试集数据进行预测得到误差的预测值, 与初步预测结果用加法器合并得到最终预测结果。通过对比实验, 验证了此误差修正模型大幅提升了模型的预测精度。这种组合预测模型为空气质量的预测提供了一种新的方法, 也为大气污染的有效治理和政策制定提供了精准的数据和科学的理论依据支撑。

## 基金项目

国家自然科学基金(11961039)。

## 参考文献

- [1] 宋国君, 李虹霖. 基于 PM<sub>2.5</sub> 的空气污染防治政策框架设计[J]. 中国人口·资源与环境, 2023, 33(2): 1-10.
- [2] 王建树, 王瑛, 赵敏娟, 等. ARIMA 模型在苏州市空气质量指数预测中的应用[J]. 公共卫生与预防医学, 2019, 30(2): 18-20.
- [3] 朱润苏. 基于机器学习的空气质量预测算法研究[D]: [硕士学位论文]. 无锡: 江南大学, 2021.
- [4] 付恩, 张益农, 杨帆, 等. 基于频率分解 Transformer 的时间序列长时预测模型[J]. 制造业自动化, 2022, 44(11): 177-181.
- [5] Wang, S.Y., Yu, L. and Lai, L.K. (2005) Crude Oil Price Forecasting with TEI@I Methodology. *Journal of Systems Science and Complexity*, **18**, 145-166.
- [6] 史学良, 李梁, 赵清华. 基于改进 LSTM 网络的空气质量指数预测[J]. 统计与决策, 2021, 37(16): 57-60.
- [7] 梁涛, 谢高峰, 米大斌, 等. 基于 CEEMDAN-SE 和 LSTM 神经网络的 PM<sub>10</sub> 浓度预测[J]. 环境工程, 2020, 38(2): 107-113.
- [8] Dragomiretskiy, K. and Zosso, D. (2013) Variational Mode Decomposition. *IEEE Transactions on Signal Processing*, **62**, 531-544. <https://doi.org/10.1109/TSP.2013.2288675>
- [9] 柯虎, 张新生, 陈章政. 基于二次分解 BAS-LSTM 的陕西省碳排放预测研究[J]. 经营与管理, 2023, 5: 1-14.
- [10] Hao, L., Ling, H.Z., Cheng, T., et al. (2023) Short-Term Load Forecasting Model Based on Gated Recurrent Unit and Multi-Head Attention. *The Journal of China Universities of Posts and Telecommunications*, **5**, 1-7.
- [11] Liu, X. and Guo, H.Y. (2022) Air Quality Indicators and AQI Prediction Coupling Long-Short Term Memory (LSTM) and Sparrow Search Algorithm (SSA): A Case Study of Shanghai. *Atmospheric Pollution Research*, **13**, Article ID: 101551. <https://doi.org/10.1016/j.apr.2022.101551>
- [12] Duan, J.K., Chang, M.H., Chen, X.Y., et al. (2022) A Combined Short-Term Wind Speed Forecasting Model Based on CNN-RNN and Linear Regression Optimization Considering Error. *Renewable Energy*, **200**, 788-808. <https://doi.org/10.1016/j.renene.2022.09.114>
- [13] Wang, D. and Mao, K. (2018) Multimodal Object Classification Using Bidirectional Gated Recurrent Unit Networks. 2018 *IEEE 3rd International Conference on Data Science in Cyberspace (DSC)*, Guangzhou, 18-21 June 2018, 685-690. <https://doi.org/10.1109/DSC.2018.00109>
- [14] Lin, Z., Feng, M., Santos, C.N., et al. (2017) A Structured Self-Attentive Sentence Embedding.
- [15] Vaswani, A., Shazeer, N., Parmar, N., et al. (2017) Attention Is All You Need. In: *Proceedings of the 31st International Conference on Neural Information Processing Systems*, Curran Associates Inc., Long Beach, 6000-6010.
- [16] Ustunu, B., Melssen, W.J. and Buydens, L.M.C. (2007) Visualisation and Interpretation of Support Vector Regression Models. *Analytica Chimica Acta*, **595**, 299-309. <https://doi.org/10.1016/j.aca.2007.03.023>
- [17] Lee, H., Kim, D. and Gu, J.H. (2023) Prediction of Food Factory Energy Consumption Using MLP and SVR Algorithms. *Energies*, **16**, Article No. 1550. <https://doi.org/10.3390/en16031550>