

# 基于 LSTM 网络与误差补偿的预测模型

王健<sup>1,2</sup>, 宋颖<sup>1,2</sup>, 吴涛<sup>1,2</sup>

(1. 安徽大学 数学科学学院, 安徽 合肥 230031;

2. 安徽大学 计算智能与信号处理教育部重点实验室, 安徽 合肥 230039)

**摘要:**随着时代发展,空气质量逐渐受到人们的重视,所以对未来空气污染物的变化预测显得尤为重要。首先,针对 $PM_{2.5}$ 的非线性变化以及变化所具有的周期性,选取完整年度数据进行训练和预测,使用对非线性序列数据拟合效果较好的 LSTM 网络作为初步预测模型,选择合适的滑动窗口,使用训练数据,建立了 LSTM 网络预测模型。由于 LSTM 网络预测结果中存在相邻年份误差分布相似,但整年分布不均匀的现象,使用 FCM 对训练数据及误差进行模糊聚类。通过聚类中心,对当前预测数据进行分类,并利用聚类结果,得到当前预测数据的误差补偿值,对 LSTM 网络的当前预测结果进行误差补偿,得到最终预测结果。最后,通过合肥 2017 年至 2021 年的空气污染数据对该方法进行了验证,结果表明,所建模型的效果优于其他对比模型,具有一定的可行性与有效性。

**关键词:**  $PM_{2.5}$  预测; 长短时记忆网络; 模糊聚类; 误差相似性; 误差补偿

中图分类号: TP183

文献标识码: A

文章编号: 1673-629X(2023)03-0133-06

doi: 10.3969/j.issn.1673-629X.2023.03.020

## Prediction Model Based on LSTM Network and Error Compensation

WANG Jian<sup>1,2</sup>, SONG Ying<sup>1,2</sup>, WU Tao<sup>1,2</sup>

(1. School of Mathematical Sciences, Anhui University, Hefei 230031, China;

2. Key Lab of Intelligent Computing and Signal Processing of Ministry of Education, Hefei 230039, China)

**Abstract:** With the development of the times, people pay more and more attention to air quality, so it is particularly important to predict the change of air pollution in the future. Firstly, aiming at the nonlinear change and periodicity of  $PM_{2.5}$ , select the complete annual data for training and prediction. LSTM network, which has excellent fitting effect on the nonlinear sequence data, is used as the preliminary prediction model. The LSTM network prediction model is established by selecting the appropriate sliding window and using the training data. For the phenomenon of similar error distribution in adjacent years but uneven distribution in the whole year in the prediction results of LSTM network, the training data and errors are fuzzy clustered by FCM. Through the clustering results, the current prediction data are classified, and the error compensation value of the current prediction data is obtained by using the clustering center. Using the error compensation value to modify the prediction result of LSTM network, the final prediction result is obtained. Finally, the proposed method is verified by the air pollutant data set of Hefei from 2017 to 2021. The results show that the model is better than other comparison models, which is feasible and effective.

**Key words:**  $PM_{2.5}$  prediction; LSTM; fuzzy clustering; error similarity; error compensation

## 0 引言

$PM_{2.5}$ 作为大气中的一种细小颗粒物,具有粒径小、面积大、毒性强的特点,严重威胁着人们的健康。近年来,国家不断加大各项环境治理措施, $PM_{2.5}$ 治理已取得很好的效果,雾霾天气已少有发生,但未来 $PM_{2.5}$ 的治理仍任重道远。因此,对未来 $PM_{2.5}$ 的变化进行预测与模拟尤为重要。

当前,空气质量预测模型主要有两种类型,一种是基于大气动力学的数值预报模型,即基于大气内部运动的物理规律,通过污染源、化学动力学等相关知识,建立起空气污染浓度预测的动态模型,对大气情况进行模拟预测<sup>[1-2]</sup>。

另一种则是基于统计学、机器学习等方法,利用已有的数据,从数据中挖掘其中所蕴含的数理规律,建立

收稿日期: 2022-05-04

修回日期: 2022-09-08

基金项目: 国家自然科学基金青年项目(618006001A); 安徽大学研究生创新项目(Z010111021)

作者简介: 王健(2000-),男,硕士研究生,从事机器学习与智能计算相关研究; 通讯作者: 吴涛(1970-),男,教授,博士,从事机器学习、智能计算与统计决策等相关研究。

预测模型,实现对未来情况的预测。常用的统计方法与机器学习方法有多元线性回归<sup>[3]</sup>、ARIMA<sup>[4]</sup>、支持向量回归<sup>[5]</sup>等等。而随着深度学习的快速发展,神经网络得到了各个行业的关注。由于神经网络在拟合非线性关系上的优越性,也被广泛应用于空气质量预测当中。常见的网络结构有BP神经网络<sup>[6]</sup>、卷积神经网络<sup>[7]</sup>、循环神经网络<sup>[8]</sup>及其相关变种LSTM<sup>[9-10]</sup>等,并通过神经网络,建立了相应雾霾预警系统<sup>[11]</sup>。在污染物预测建模过程中,通过历史污染物数据,对网络进行训练,建立相应的网络预测模型,取得了不错的效果。

同时,为了进一步提高预测的准确率,很多学者通过时空变化的规律进行分析建模,提高原有预测模型精度。例如:空间方面,李焱航<sup>[12]</sup>通过引入历史风场数据,将离散的PM<sub>2.5</sub>插值成网格图的形式,然后使用U-net网络对未来时刻进行预测,提高了对PM<sub>2.5</sub>的突变情况的预测能力;宋飞扬<sup>[13]</sup>通过KNN算法为所需预测的站点选取与其相关的空间相邻站点,得到其空间特征,再对目标站点建立LSTM预测模型,以达到对空间影响因素的利用。

时间方面,张怡文<sup>[14]</sup>通过对不同的季节分别进行建模的方式,通过季节性预测方法提高了预测准确率;蒋洪迅等人<sup>[15]</sup>通过序列分解的思路,分解为若干子序列,提取出不同时间尺度的周期序列,再利用LSTM网络对各子序列进行预测得到预测值,使得预测精度得到了提升;Wang等人<sup>[16]</sup>采用的基于区间值的二元分解与最优组合集成的学习方法,将日均值浓度转变为区间数,并将不同年份的同季节数据进行重构,采用二元经验分解对序列进行分解,再利用区间多层感知机进行预测,使得预测精度有所上升。

该文主要从时间方面入手,首先选择了对序列数据预测表现较好的长短时记忆网络(LSTM)作为初步预测模型,完成多因素影响下PM<sub>2.5</sub>的初步预测。然后利用PM<sub>2.5</sub>浓度变化所具有的季节性与周期性,探索相邻年份历史预测误差分布特点。通过模糊聚类的方法,分析比较当前数据与历史数据的相似性,对当前预测数据的预测误差进行估计,并做出相应的误差补偿,以提高当前预测结果的准确性。

## 1 理论与模型

### 1.1 长短时记忆网络(LSTM)

LSTM神经网络<sup>[17]</sup>大体结构与传统循环神经网络结构相同,主要为了解决长时间序列问题中的梯度爆炸和梯度消失的问题,在神经网络内部结构中引入了“记忆细胞(cell state)”这一概念。LSTM单元中存在三个门控单元,即遗忘门、输入门、输出门。这三种

门控单元均为非线性单元,其中遗忘门决定了从上一个状态中舍弃什么信息,即遗忘阶段;输入门决定了当前状态保存什么信息,即选择记忆阶段;而输出门决定了从记忆细胞中输出什么信息,即输出阶段。相关计算公式与网络结构(见图1)如下:

遗忘门:

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

输入门:

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

$$C_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \quad (3)$$

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

输出门:

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

$$h_t = o_t * \tanh(C_t) \quad (6)$$

其中, $W_f$ 、 $W_i$ 、 $W_c$ 、 $W_o$ 是与各个部分之间连接的权重,均为可训练参数; $h_{t-1}$ 为 $t-1$ 时刻的输出值; $x_t$ 为 $t$ 时刻的输入; $\sigma$ 为sigmoid函数; $i_t$ 为输入信息进入单元状态的程度; $\tilde{C}_t$ 为新的候选值; $o_t$ 为输出门的值; $h_t$ 为当前时刻输出。

通过式(1)计算,得到 $f_t \in (0,1)$ ,来决定历史信息 $C_{t-1}$ 通过的多少,即完成遗忘的功能。然后通过式(2)计算输入值 $x_t$ 进入单元的程度值 $i_t$ ,通过式(3)生成新的候选值 $\tilde{C}_t$ ,作为当前层产生的候选值添加到细胞单元中,进而通过式(4)计算出细胞单元中需要更新的新的信息 $C_t$ 。最后,进入输出门,由式(5)计算出当前时刻输出门的值,通过式(6)计算出最后的输出值 $h_t$ 。

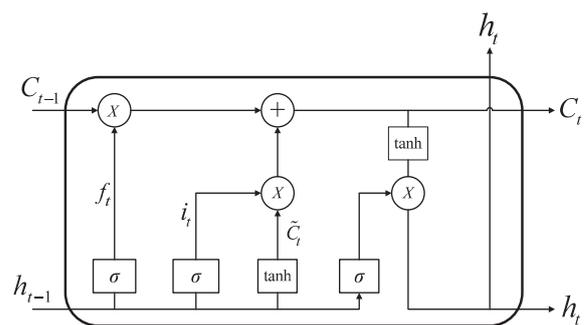


图1 LSTM网络结构

### 1.2 模糊C均值聚类(FCM)

传统的聚类分析是一种硬划分,即划分事物即是非此即彼的特性,而模糊聚类则采用的是一种软划分的方式,通过赋予每个样本点隶属度的方式,来衡量每个样本对每个类的影响,可以更好地衡量数据对所在类的影响。现假定有一个数据集,若把这些数据划分为 $c$ 类,那么对应的就有 $c$ 个聚类中心 $C_i (i = 1, 2, \dots, N)$ ,每个样本属于某一类的隶属度为 $u_{ij}$ ,定义一个FCM目标函数(7)及约束条件(8)。

$$J = \sum_{i=1}^c \sum_{j=1}^n u_{ij}^m \|x_j - c_i\|^2 \quad (7)$$

$$\sum_{i=1}^c u_{ij} = 1, \quad j = 1, 2, \dots, n \quad (8)$$

式中,  $u_{ij} \in (0, 1)$ , 表示第  $j$  个数据对第  $i$  个类的隶属度。通过不断迭代, 使式(7)取得最小值, 从而确定聚类中心  $C_i$  和每个样本点对应聚类中心的隶属度  $u_{ij}$  :

$$C_i = \frac{\sum_{j=1}^n u_{ij}^m x_j}{\sum_{j=1}^n u_{ij}^m} \quad (9)$$

$$u_{ij} = \frac{1}{\sum_{k=1}^c \left(\frac{d_{ij}}{d_{kj}}\right)^{\frac{2}{m-1}}} \quad (10)$$

### 1.3 基于 LSTM 网络和 FCM 误差补偿的 PM<sub>2.5</sub> 预测模型

下面对该文所使用的算法进行介绍, 具体流程如下:

步骤 1: 将收集的数据集  $T$  的数据进行 max-min 归一化。对完整年度将数据集进行分割, 分割成三个部分, 用于神经网络训练的历史训练数据  $T_1$ ; 用于模糊聚类实验的测试集  $T_2$ ; 用于误差补偿后验证效果的验证集  $T_3$ 。并设置输入 LSTM 网络的数据窗口宽度为  $w$ 。

步骤 2: 使用训练集  $T_1$  完成对 LSTM 网络的训练, 并把测试集  $T_2$  输入已训练完成的神经网络中, 得到预测结果  $\hat{y} = \{\hat{y}_1, \hat{y}_2, \dots, \hat{y}_n\}$ , 并计算出相应的误差  $e_i = \hat{y}_i - y_i$ , 得到对应的误差数据  $e = \{e_1, e_2, \dots, e_n\}$ 。

步骤 3: 将测试集  $T_2$  与误差数据  $e$  组成新的数据集  $(T_2, e)$ , 并对数据集  $(T_2, e)$  进行 FCM 模糊聚类。确定聚类中心  $L_1(i) = \{c_1, c_2, \dots, c_k\}$ , 其中  $c_i = (x_{i1}, x_{i2}, \dots, x_{im}, e_i)$ 。

步骤 4: 将验证集  $T_3$  输入已训练好的 LSTM 网络中进行预测, 得到初步预测结果  $\hat{y} = \{\hat{y}_1, \hat{y}_2, \dots, \hat{y}_n\}$ 。将  $T_2$  经过聚类所确定的聚类中心  $c_i^* = (x_{i1}, x_{i2}, \dots, x_{im})$ , 作为验证集  $T_3$  的分类中心, 对验证集  $T_3$  进行分类。其中, 距离度量使用欧氏距离, 分类数为  $k$ 。

步骤 5: 对分类后验证集  $T_3$  中属于不同类的预测值  $\hat{y}_i$ , 进行相应的误差补偿, 误差补偿公式如下:

$$\tilde{y}_i = \hat{y}_i - e_{i,j}$$

其中,  $\hat{y}_i$  为时刻  $i$  的预测值,  $e_{i,j}$  为时刻  $i$  的预测数据所在类的分类中心  $c_j = (x_{j1}, x_{j2}, \dots, x_{jm}, e_j)$  中的,  $\tilde{y}_i$  为  $i$  时刻误差修正后的预测值。

具体算法流程如图 2 所示。

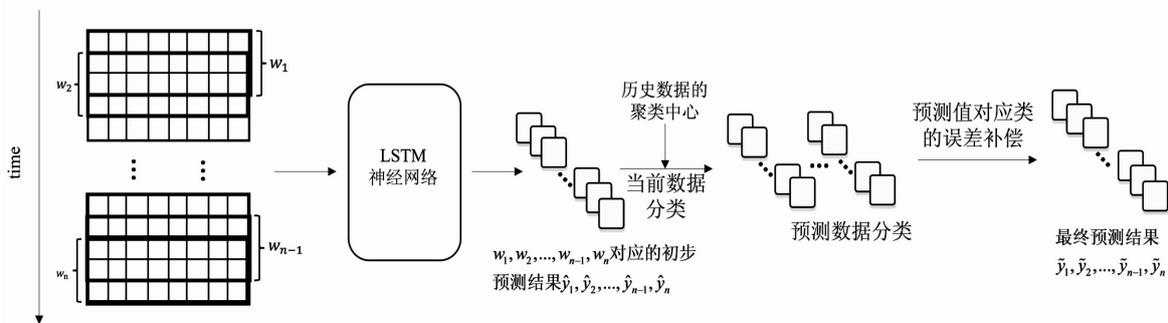


图 2 基于 LSTM 网络和 FCM 误差补偿的 PM<sub>2.5</sub> 预测模型

## 2 PM<sub>2.5</sub> 浓度预测的实例分析

### 2.1 数据的初步分析

该文所使用的数据集为 2017 年 1 月 1 日至 2021 年 12 月 31 日安徽省合肥市的 PM<sub>2.5</sub>、PM<sub>10</sub>、SO<sub>2</sub>、CO、NO<sub>2</sub>、O<sub>3</sub> 这六种空气污染物日均值浓度数据 (其中 CO 数据单位 mg/m<sup>3</sup>, 其余数据单位均为 μg/m<sup>3</sup>), 共收集数据 1 825 条。为充分利用 PM<sub>2.5</sub> 变化的周期性, 分割数据集时, 将前三年共计 1 095 条数据作为神经网络的历史训练集  $T_1$ , 用于训练神经网络; 将数据集的第四年, 即第 1 096 条至 1 461 条共计 365 条作为测试集  $T_2$ , 用于历史预测误差的模糊聚类实验; 数据集的第五年, 即第 1 462 条数据到 1 825 条数据  $T_3$  作为经过误

差补偿后的验证集, 验证该方法的效果与可行性。

通过绘制 PM<sub>2.5</sub> 的浓度时序变化图 (见图 3), 可以看出 PM<sub>2.5</sub> 的变化具有周期性, 存在“秋冬高, 春夏低”的现象, 并且相邻两年之间 PM<sub>2.5</sub> 浓度的分布具有相似性。此外还可以看出五年的 PM<sub>2.5</sub> 浓度变化总体有一定的下降趋势, 并且单日 PM<sub>2.5</sub> 含量激增的情况也大大减少, 猜测应该与国家环境相关治理有关, 城市整体空气质量正在逐年提高。

对该文所使用数据集的五种与 PM<sub>2.5</sub> 相关的污染因子进行 Pearson 相关因子检验, 检验结果表明, PM<sub>2.5</sub> 与 PM<sub>10</sub>、CO 有较强的正相关性, 与 NO<sub>2</sub>、SO<sub>2</sub> 存在相关性, 而与 O<sub>3</sub> 呈现一定的负相关性, 说明了所选 PM<sub>2.5</sub> 影响因子的合理性。具体分析结果如表 1 所示。

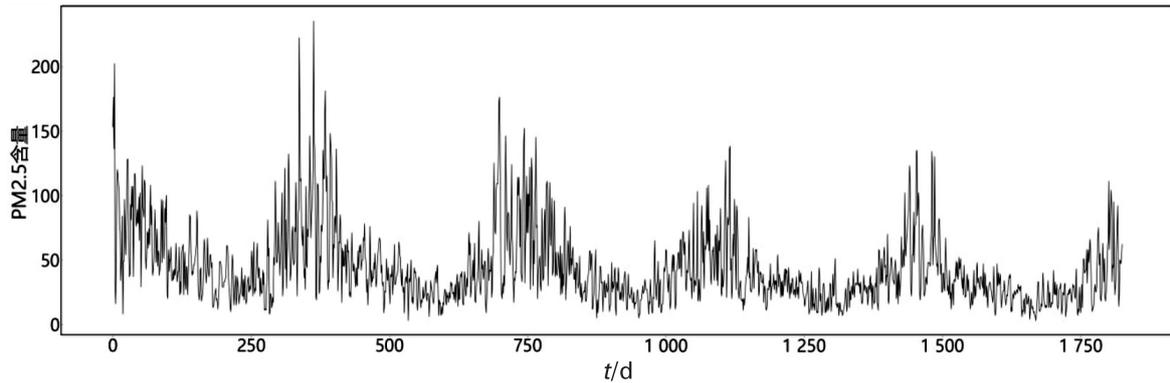


图 3 PM<sub>2.5</sub>浓度时序变化

表 1 Pearson 相关因子分析

因素	PM <sub>10</sub>	SO <sub>2</sub>	NO <sub>2</sub>	CO	O <sub>3</sub>
相关系数	0.73	0.59	0.47	0.83	-0.19

## 2.2 数据预处理

### (1) 缺失数据填补。

由于收集的数据存在少量缺失值,且污染物的变化存在一定的趋势性,而线性插值可以对趋势性的数据进行很好的补充,故采用线性插值的方式对缺失的数据值进行合理填补。

### (2) 数据归一化。

为了提高神经网络在训练过程中的收敛速度和模型准确率,以及各污染物由于量纲不同对数据聚类所产生的影响,而 max-min 归一化处理可以有效解决这些问题,故对数据进行 max-min 归一化处理,使得数据范围为 [0,1]。

### (3) 数据滑动窗口化处理。

在使用 LSTM 神经网络进行预测时,需要将时序型数据转化为监督型数据,建立一个输入 X、输出 Y 的映射关系,而神经网络则是学习 X 到 Y 的模式,故需将数据处理为适合网络的输入形式。所以采用滑动窗口的方式建立这种映射关系,具体处理方式如下:

$$\text{设原始数据集为 } T = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \text{ 其中 } x_i = (x_{i1}, x_{i2}, x_{ij}, \dots, x_{im}), x_{ij} \text{ 表示第 } i \text{ 个样本的第 } j \text{ 个特征, 设窗口大小为 } w, \text{ 则需要建立如下的映射关系:}$$

其中,  $y_{p,q}$  为待预测第  $p$  天的 PM<sub>2.5</sub> 真实值。

$$\begin{cases} (x_1, \dots, x_w)^T \rightarrow y_{w+1,q} \\ (x_2, \dots, x_{w+1})^T \rightarrow y_{w+2,q} \\ \vdots \\ (x_{n-w}, \dots, x_{n-1})^T \rightarrow y_{n,q} \end{cases}$$

其中,  $y_{p,q}$  为待预测第  $p$  天的 PM<sub>2.5</sub> 真实值。

## 2.3 模型参数设置

该文基于深度学习 keras 框架,搭建 LSTM 神经网络

作为初步预测模型。在神经网络进行训练之前,需要对神经网络的超参数进行设置,所使用的 LSTM 网络超参数具体设置如下:两层 LSTM,节点数分别为 128 和 64;一层全连接层,节点数为 32;训练所使用的损失函数为 MSE(均方误差),优化器采用 Adam 优化器, batchsize 设置为 64。为了防止网络过拟合,在每个 LSTM 层后,设置丢弃率 dropout 为 0.2。

在滑动窗口处理的过程中,不同的窗口大小会对实验结果造成影响。经多次实验,确定了最佳窗口大小,窗口大小为 3。

## 2.4 模型评估指标

评估指标选用均方根误差(RMSE)、平均绝对误差(MAE),具体公式如下:

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

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

其中,  $y_i$  为  $i$  时刻的 PM<sub>2.5</sub> 的真实值,  $\hat{y}_i$  为  $i$  时刻的 PM<sub>2.5</sub> 的预测值。

## 2.5 结果分析

### 2.5.1 LSTM 初步预测结果分析

使用 LSTM 网络对测试集数据进行预测,对一年中不同的时间段的误差进行误差统计,并绘制相邻年间的误差时序变化图,具体结果如表 2 所示。

表 2 一年不同时间段误差统计

时间	RMSE	MAE
1月-3月	20.522	14.191
4月-6月	10.526	8.247 8
7月-9月	8.921	7.255
10月-12月	20.722	15.614

通过表 2,可知一年间不同时间段的误差分布不均匀,表现为年初和年末误差较大,年中误差较小,这个特点也可以在图 4 中得到体现。由图 4 可知,相邻年间误差分布具有相似性,且 PM<sub>2.5</sub> 变化具有季节性,故推断,相似数据的误差分布是相似的。

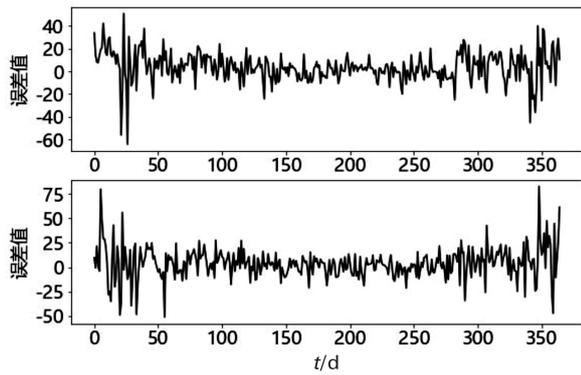


图4 相邻年份预测误差变化

2.5.2 预测结果与误差

由 2.5.1 节的结论,对测试集及对应误差组成的数据集进行 FCM 模糊聚类。利用聚类中心,对验证集

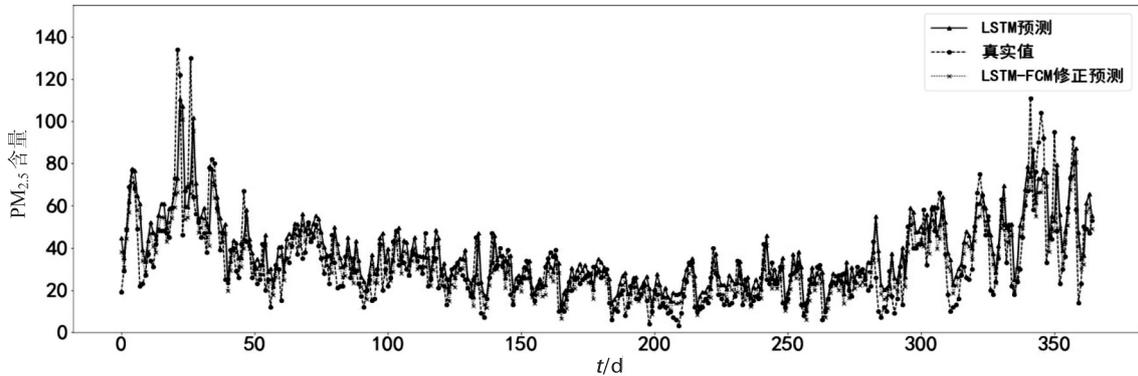


图5 全年 PM<sub>2.5</sub> 预测结果

3 结束语

使用合肥 2017 年 1 月 1 日至 2021 年 12 月 31 日的空气质量六因子的日均值数据进行了 PM<sub>2.5</sub> 的浓度预测。实验过程中,首先对少量缺失数据进行了线性插值填补,并进行归一化处理。归一化后,对数据集进行滑动窗口化处理,处理成适合 LSTM 神经网络输入的滑动窗口数据,并完成 LSTM 网络的训练以及初步预测。通过实验,发现 PM<sub>2.5</sub> 预测误差存在相邻年份误差分布相似,但整年误差分布不均匀,呈现出年初、年末误差较大,而年中误差较小。

根据此特点,采用了 FCM 对测试集数据进行聚类。根据聚类结果对验证集分类,并根据聚类中心做出相应的误差补偿。最终实验结果表明,均方根误差 (RMSE) 与绝对平均误差 (MAE) 均得到了降低,总体取得了不错的效果。

但仍存在很多不足。在研究初期,仅对污染因子进行了相关性分析,验证了所选特征之间所具有相关性,并未对诸如风速、风压等气象因子进行分析和考虑。

在模糊聚类方面,通过模糊聚类确定当前数据与历史预测数据相似的部分,利用对应类的聚类中心作

数据进行分类。根据分类结果,对属于不同类的预测数据进行相应的误差补偿,得到最终预测结果。并与其他实验方法进行对比,结果如表 3 所示。

表 3 实验预测误差结果

模型名称	RMSE	MAE
SVR	16.915	13.051
BP	15.627	11.758
GRU	13.417	10.077
LSTM	13.537	9.976
LSTM-FCM	12.398	8.696

由表 3 可知,经过误差修正后的预测结果相较于其他对比的实验方法,效果最佳,从而验证了该方法的有效性。相关预测结果如图 5 所示。

为误差补偿值,可以对 LSTM 初步预测结果进行很好的补偿。但由于误差存在正负,聚类并不能很好区分误差的正负,可以在接下来的研究中考虑如何确定误差的方向,以进一步提高预测精度。

时间因素方面,本次实验采用的为城市的日均值数据,而天气情况具有实时性,若将数据更换为以小时为单位的数据,则模型的预测精度应该会更好,并更具有实际意义。在后续的研究中,可以针对以上问题进行一定的改进,取得更好的预测效果。

参考文献:

[1] LELIEVELD J, EVANS J S, FNAIS M, et al. The contribution of outdoor air pollution sources to premature mortality on a global scale[J]. Nature, 2015, 525 (7569) :367-371.

[2] SHANG Z, DENG T, HE J, et al. A novel model for hourly PM<sub>2.5</sub> concentration prediction based on CART and EELM [J]. Science of the Total Environment, 2019, 651: 3043-3052.

[3] DIMITRIOU K, KASSOMENOS P. A study on the reconstitution of daily PM<sub>10</sub> and PM<sub>2.5</sub> levels in Paris with a multivariate linear regression model[J]. Atmospheric Environment, 2014, 98:648-654.

[4] 孟凡强. ARIMA 模型在空气污染指数预测中的应用[J].

- 统计与决策,2009(7):33-35.
- [5] MOGOLLÓN-SOTELO C, CASALLAS A. A support vector machine model to forecast ground-level  $PM_{2.5}$  in a highly populated city with a complex terrain[J]. *Air Quality, Atmosphere & Health*,2021,14(3):399-409.
- [6] 王敏,邹滨,郭宇,等.基于BP人工神经网络的 $PM_{2.5}$ 浓度空间预测[J]. *环境污染与防治*,2013,35(9):63-66.
- [7] 吴春霖,李琦,侯俊雄,等.卷积神经网络的 $PM_{2.5}$ 预报模型[J]. *测绘科学*,2018,43(8):68-75.
- [8] 范竣翔,李琦,朱亚杰,等.基于RNN的空气污染时空预报模型研究[J]. *测绘科学*,2017,42(7):76-83.
- [9] 白盛楠,申晓留.基于LSTM循环神经网络的 $PM_{2.5}$ 预测[J]. *计算机应用与软件*,2019,36(1):67-70.
- [10] SENG D,ZHANG Q,ZHANG X, et al. Spatiotemporal prediction of air quality based on LSTM neural network[J]. *Alexandria Engineering Journal*,2021,60(2):2021-2032.
- [11] 黄威,王星捷,阳清青.基于神经网络的雾霾预警系统研究与实现[J]. *计算机技术与发展*,2019,29(10):26-30.
- [12] 李焱航,翟卫欣,颜寒祺,等.基于U-net神经网络模型的 $PM_{2.5}$ 逐小时浓度值预测模型[J]. *北京大学学报:自然科学版*,2020,56(5):796-804.
- [13] 宋飞扬,铁治欣,黄泽华,等.基于KNN-LSTM的 $PM_{2.5}$ 浓度预测模型[J]. *计算机系统应用*,2020,29(7):193-198.
- [14] 张怡文,郭傲东,吴海龙,等.基于PCA-BP神经网络的 $PM_{2.5}$ 季节性预测方法研究[J]. *南京林业大学学报:自然科学版*,2020,44(5):231-238.
- [15] 蒋洪迅,闫超超,张立峰.基于时序分解和神经网络的 $PM_{2.5}$ 浓度预测研究——以沈阳市为例[J]. *系统科学与数学*,2021,41(12):3446-3460.
- [16] WANG Z C, CHEN L R, ZHU J M, et al. Double decomposition and optimal combination ensemble learning approach for interval-valued AQI forecasting using streaming data[J]. *Environmental Science and Pollution Research*,2020,27(30):37802-37817.
- [17] HOCHREITER S, SCHMIDHUBER J. Long short-term memory[J]. *Neural Computation*. 1997,9(8):1735-1780.
- 
- (上接第126页)
- (1):012083.
- [19] YAQUB M, ASIF H, KIM S, et al. Modeling of a full-scale sewage treatment plant to predict the nutrient removal efficiency using a long short-term memory (LSTM) neural network[J]. *Journal of Water Process Engineering*,2020,37:101388.
- [20] MI J X, SUN Y R, LU J, et al. Robust supervised sparse representation for face recognition[J]. *Cognitive Systems Research*,2020,62:10-22.
- [21] MIRJALILI S, LEWIS A. The whale optimization algorithm[J]. *Advances in Engineering Software*,2016,95:51-67.
- [22] TAO H, PIERRE P, SHU F, et al. Probabilistic energy forecasting: global energy forecasting competition 2014 and beyond[J]. *International Journal of Forecasting*,2016,32(3):454-461.
- [23] WU H, XU J, WANG J, et al. Autoformer: decomposition transformers with auto-correlation for long-term series forecasting [C]//Advances in neural information processing systems. [s. l.]: Curran Associates Inc,2021:22419-22430.
- [24] LIM B, ARK S, LOEFF N, et al. Temporal fusion transformers for interpretable multi-horizon time series forecasting[J]. *International Journal of Forecasting*,2021,37(4):1748-1764.
- [25] KE G, MENG Q, FINLEY T, et al. LightGBM: a highly efficient gradient boosting decision tree [C]//Proceedings of the 31st international conference on neural information processing systems. Red Hook: Curran Associates Inc,2017:3149-3157.
- [26] ADNAN R M, KHOSRAVINIA P, KARIMI B, et al. Prediction of hydraulics performance in drain envelopes using Kmeans based multivariate adaptive regression spline[J]. *Applied Soft Computing*,2021,100:107008.