

# 重庆中心城区臭氧浓度网格预报方法 及其应用研究<sup>①</sup>

吉莉<sup>1</sup>, 刘晓冉<sup>2</sup>, 李强<sup>1</sup>

1. 重庆市北碚区气象局, 重庆 400700; 2. 重庆市气象科学研究所, 重庆 401147

**摘要:** 为探索不同机器学习方法在臭氧浓度预测中的应用效果, 基于重庆市中心城区及其周边区域 2017—2019 年臭氧数据及气象数据, 利用 KNN 数据挖掘算法与 BP 神经网络算法, 建立了 2 种以重庆中心城区北碚区为中心的臭氧浓度预报模型, 并运用反距离加权插值法将臭氧浓度预报值网格化。研究显示: ①2017—2019 年间, 臭氧浓度除铜梁区和沙坪坝区呈逐年上升趋势外, 其他区域呈“单峰形”的变化, 最大值出现在 2018 年; 夏季臭氧浓度最高, 冬季最低; 臭氧浓度与气象要素关系最密切的是最高气温、平均气温及相对湿度。②2 种预测模型在 7 个区的训练中平均预报准确度超过了 74%, 其中 KNN 模型的准确率高于 BP 神经网络模型, 2 种模型与实测值的相关性超过了 0.7, 拟合度较高; 在验证模型准确率中, BP 神经网络模型在检验过程中的整体效果高于训练效果; 总的来说, 基于 BP 神经网络模型的预测结果总体优于 KNN 模型的。③将模型应用于臭氧浓度网格预报中, 从数值来看, 高值与实测值基本一致, 低值低于实测值; 从地理位置来看, 低值与实测值位置基本一致, 分布在重庆中心城区中部偏西地区, 高值预测区域差距较大, 实测高值区分布在中心城区的南部和东部。

**关键词:** 机器学习; 臭氧浓度; 网格预报

中图分类号: X823

文献标志码: A

文章编号: 1000-5471(2023)01-0063-10

## Application of Grid Prediction of Ozone Concentration in Central Urban Area of Chongqing

JI Li<sup>1</sup>, LIU Xiaoran<sup>2</sup>, LI Qiang<sup>1</sup>

1. Weather Bureau in Beibei District of Chongqing City, Chongqing 400700, China;

2. Chongqing Institute of Meteorological Sciences, Chongqing 401147, China

**Abstract:** In order to explore the application effect of different machine learning methods in ozone concentration prediction, based on the ozone data and meteorological data in the central urban area of Chongqing and its surrounding areas from 2017 to 2019, this paper used KNN data mining algorithm and BP neural network algorithm to establish an ozone concentration prediction model centered on Beibei District of the

① 收稿日期: 2022-06-16

基金项目: 重庆市气象局业务技术攻关项目(YWJSGG-202134)。

作者简介: 吉莉, 高级工程师, 主要从事气象服务研究的工作。

通信作者: 刘晓冉, 正高级工程师, 博士。

central urban area of Chongqing, and used the inverse distance weighted interpolation method to grid the ozone concentration prediction values. The conclusions were as follows: ① from 2017 to 2019, except that Tongliang and Shapingba showed an upward trend year by year, the trend of other district showed a “single peak”, and the maximum appeared in 2018. The concentration reached the highest in summer and the lowest in winter. The most closely related between ozone and meteorological elements are the maximum temperature, average temperature and relative humidity. ② The average prediction accuracy of the two prediction models in the training of seven areas exceeded 74%, of which the accuracy of KNN model was higher than that of BP neural network model. The correlation between the two models and the measured value exceeded 0.7, and the fitting degree was high. In verifying the accuracy of the model, the overall effect of BP neural network model in the process of testing was higher than that of training; In general, the prediction results based on BP neural network model were better than those of KNN model. ③ The model was applied to the grid prediction of ozone concentration. From the numerical point of view, the high value was basically consistent with the measured value, but the low value is lower than the measured value; From the perspective of geographical location, the location of low value was basically consistent with the measured value, which is distributed in the central by west area, the gap between high value prediction areas is large, and the measured high value area is in the South and East.

**Key words:** machine learning; ozone concentration; grid prediction

作为城市空气主要污染物之一的臭氧,是一种光化学反应的微量气体,它主要来自大气层中氮氧化物和碳氢化合物等.近地面的臭氧主要来源于氮氧化物和挥发性有机化合物,光照作用下的大气光化学反应、平流层输入、前体物浓度水平、气象条件和区域传输等都会对其产生影响<sup>[1]</sup>.高浓度臭氧不仅威胁人类健康,还能对植物、农作物等造成危害<sup>[2]</sup>.

目前学界对于城市臭氧污染以及臭氧污染与气象条件关系的研究已有不少报道<sup>[3-10]</sup>,Shen 等<sup>[11]</sup>利用空气质量模式 CMAQ 和 CAMxO3 评估了珠三角地区臭氧预报的效果;Kathleen 等<sup>[12]</sup>利用空气质量模式 WRF-Chem 评估了欧洲地面臭氧的预报效果;杜勤博等<sup>[13]</sup>采用 BP 神经网络和逐步线性回归 2 种模型,对汕头市 2014—2017 年的 6 种污染物浓度进行预测,结果表明这两种预报模型对历史数据的拟合效果相似,总体来说,冬春季模型拟合度优于夏秋季;杜云松等<sup>[14]</sup>选用成都最高气温、日气温差和臭氧浓度,建立单因子预测方法,结果表明臭氧污染预测效果较好.可见,通过机器学习方法能够有效地预测臭氧污染.

成渝城市群是大气污染的重点防控区域之一<sup>[15]</sup>,因此,对重庆中心城区臭氧浓度进行精细化预报,对臭氧污染防治有重要意义.受环境监测点分布限制,以往研究多基于站点资料进行,重庆中心城区受地形影响,周边各站点的臭氧浓度存在一定差别,随着城市建设和社会经济的快速发展,单一的站点资料已不能完全满足臭氧污染治理需求,网格化插值技术在降水、气温、风等气象要素的处理中已显现较好的应用效果<sup>[16-18]</sup>,但用于臭氧预报的应用还较少.因此,本文以重庆中心城区北碚区为研究中心,利用臭氧浓度、常规气象要素观测资料,采用 KNN 数据挖掘算法与 BP 神经网络算法建立臭氧预报模型,运用反距离加权插值法将臭氧浓度预报值网格化,以期对臭氧污染的科学防治提供必要的技术支撑.

## 1 资料与方法

### 1.1 研究区域

重庆市中心城区由北碚、巴南、沙坪坝、渝北、南岸、渝中、江北、九龙坡和大渡口等 9 个区组成,其中设有气象机构的为北碚、巴南、沙坪坝、渝北 4 个区.本文选取重庆市中心城区北碚、巴南、沙坪坝、渝北 4 个区及周边区域璧山、合川、铜梁 3 个区为主要研究区域.

### 1.2 资料来源

本文采用的 2017—2019 年逐日  $\text{O}_3_{\text{8H}}$  环境监测数据为重庆市环境监测中心提供, 同期气象资料为重庆市气象信息与技术保障中心提供的国家气象站地面气象观测资料, 包括平均气温( $T$ )、最高气温( $T_{\text{max}}$ )、相对湿度( $RH$ )、降水量( $R$ )等常规气象观测数据。

### 1.3 研究方法

#### 1.3.1 KNN 数据挖掘算法

KNN(K-NearestNeighbor)数据挖掘算法是经典的分类与回归分析的邻近算法, 即  $K$  最近邻, 就是每个样本都可以用它最接近的  $K$  个邻近值来代表. 邻近算法是将数据集合中每一个记录进行分类的方法<sup>[19]</sup>, 在气象的研究中, 其主要原理是通过在历史数据库中查找与未来相似背景下的相似样本, 并依据排序结果客观赋予权重, 根据不同权重下的历史样本预报结果. KNN 算法在预报空气污染和霾等级分类中效果较好, 准确率达到了  $80\% \sim 90\%$ <sup>[20-21]</sup>. 本文的基本算法是假设预报日的平均气温、最高气温、相对湿度、降水量等气象要素为预测样本, 臭氧浓度实测值为标志量, 在训练样本集合中找到与预测样本最相似的  $K$  个近邻, 然后找出  $K$  个标志量, 最后按照投票多数原则, 选取最多的标志量为预测样本预测结果. 因此, KNN 算法也是一种相似算法.

#### 1.3.2 BP 神经网络算法

BP 神经网络算法是目前应用最广的预测方法, 不少学者将神经网络模型应用于环境气象等方面的研究, 取得了较好的成果<sup>[22-24]</sup>. 其基本思想是: 该算法工作信号正向传递和误差信号反向传递两个子过程, 学习规则和目标使用最速下降法, 通过反向传播不断调整网络的权值和阈值, 使全局误差系数最小, 学习本质是对连接权值的动态调整. 基本结构由输入层、隐层和输出层构成<sup>[25]</sup>.

#### 1.3.3 模型精度评价

对预测模型精度的评价主要采用相对误差值、绝对误差值和臭氧浓度预测准确率<sup>[26]</sup>进行比较. 其中相对误差值与绝对误差值越小, 说明预测精度越高; 若臭氧浓度预测准确率小于 0, 均按 0 处理.

$$\text{绝对误差值} = |\text{预测值} - \text{实测值}| \tag{1}$$

$$\text{相对误差值} = \text{绝对误差值} / \text{实测值} \tag{2}$$

$$\text{预测准确率} = (1 - \text{绝对误差值} \div \text{实测值}) \times 100\% \tag{3}$$

#### 1.3.4 反距离加权插值法

反距离加权插值法<sup>[27]</sup>是基于地理学第一定律相近相似的原理, 认为每个采样点对邻近区域的插值点有一定的影响, 且影响的大小随距离的增大而减小. 该插值法的特点是可以保证插值结果在采样点位与实际值一致.

#### 1.3.5 数据预处理

通常使用的机器学习算法将数据样本分为训练集与测试集, 通过训练集数据来建立模型, 测试数据则用于检验模型的泛化能力<sup>[28]</sup>. 因此在确定建立模型前, 为消除指标之间的量纲影响, 需对数据进行归一化处理. 本文采用对数据分段建模方式进行拟合, 确定 2017—2018 年 731 个有效臭氧浓度数据作为训练集数据, 2019 年逐日数据作为检验数据, 再结合平均气温、最高气温、相对湿度、降水量等 4 个气象因子, 以臭氧浓度为输入目标, 基于 KNN 数据挖掘算法和 BP 神经网络算法, 分别构建 7 个区的预测模型.

## 2 臭氧浓度与气象因子分析

### 2.1 臭氧浓度变化特征

绘制 2017—2019 年研究区臭氧浓度趋势图(图 1). 图 1a 可知, 除铜梁区和沙坪坝区呈逐年上升趋势, 北碚区及其他区最大值均出现在 2018 年, 3 年趋势呈“单峰形”; 从图 1b 可知, 研究区臭氧浓度季节变化均为夏季浓度最高, 冬季最低, 其中沙坪坝区夏季浓度最高, 铜梁区夏季浓度最低, 北碚区夏季臭氧浓度

在整个研究区中为第三高，但冬季浓度最低。

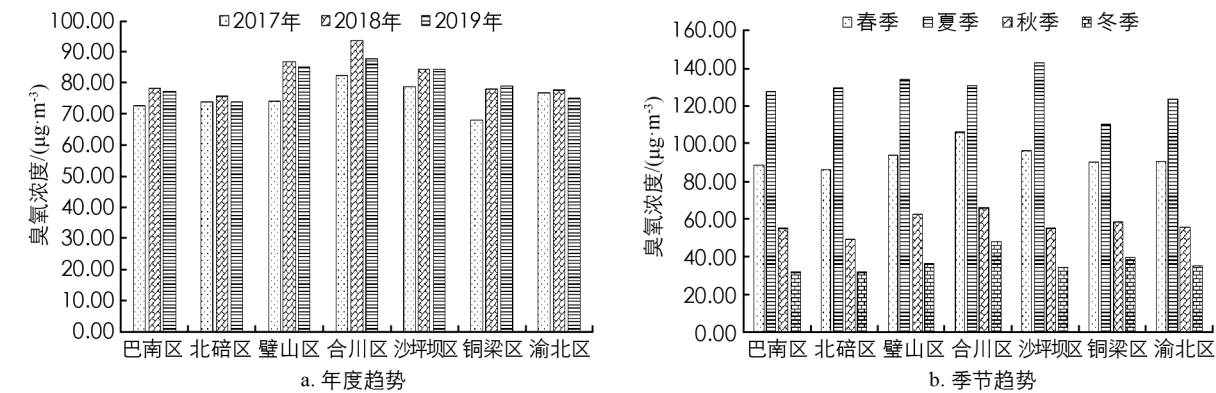


图 1 2017—2019 年研究区臭氧浓度趋势图

图 2 为研究区月平均臭氧浓度趋势图。由图 2 可知，2017—2019 年 7 个区臭氧浓度的月平均浓度呈单峰型，其中 8 月最高，12 月最低，即从 1 月起，每月的臭氧浓度都呈增加趋势，8 月达到峰值后开始逐月降低。

2.2 臭氧浓度与气象要素相关性

2017—2019 年研究区臭氧日浓度与 4 种气象要素的相关性如表 1 所示。从表 1 可看出，7 个区臭氧浓度与气象要素密切相关，对比分析发现，除降水量与臭氧浓度呈弱相关性外，其余 3 个气象要素均与臭氧浓度呈较强相关性，其值均大于 0.6，其中关系最密切的是最高气温，其次是平均气温，说明臭氧浓度的变化受温度的影响较大。臭氧浓度与相对湿度呈负相关性，这是因为空气水汽中含有较多的 H，OH 等自由基，可以分解臭氧中的氧分子，从而使臭氧浓度降低<sup>[24]</sup>。臭氧浓度与降水量相关性较低，最大相关系数也接近 0.1，说明研究区降水与光化学反应效率关系不明显。总体来说，臭氧浓度与气象要素关联度由强到弱依次为最高气温、平均气温、相对湿度、降水量。

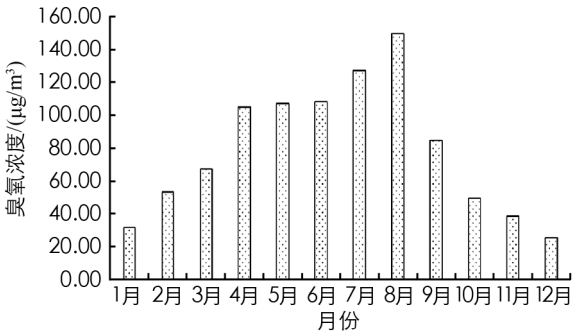


图 2 研究区臭氧浓度逐月趋势图

表 1 2017—2019 年研究区臭氧日浓度与气象要素相关性

区域	$T$	$T_{\max}$	$RH$	$R$
巴南区	0.753**	0.805**	-0.678**	-0.077*
北碚区	0.762**	0.843**	-0.681**	-0.107**
璧山区	0.716**	0.775**	-0.603**	-0.058
合川区	0.740**	0.797**	-0.715**	-0.102**
沙坪坝区	0.775**	0.833**	-0.680**	-0.051
铜梁区	0.613**	0.666**	-0.544**	-0.057
渝北区	0.751**	0.810**	-0.677**	-0.051

注：\*\*表示在 0.01 级别(双尾)，相关性显著；\*表示在 0.05 级别(双尾)，相关性显著。

3 预测模型建立及应用

3.1 KNN 数据挖掘算法模型构建

3.1.1 训练结果

本文利用 SPSS Modeler 软件、采用 KNN 数据挖掘算法对 2017—2018 年研究区的臭氧浓度进行建模

预测, 预测值与实测值对比结果见表 2。由表 2 可知, KNN 模型在训练模型中准确率均在 75% 以上, 7 个区模型的平均准确率为 77.3%, 其中合川区模型的准确率最高, 达 79.7%, 巴南区模型的准确率最低, 只有 75.3%, 北碚区模型的准确率为 78.3%, 高于平均准确率; 从决定系数来看, 北碚区决定系数为 0.892, 高于其余 6 个区, 铜梁区决定系数最低只有 0.701, 其余区决定系数在 0.82~0.87 之间, 模拟效果较好。7 个区模型平均相对误差为 0.29, 其中合川区误差最小(0.23), 璧山区误差最大(0.34); 7 个模型平均均方根误差为 22.5, 其中铜梁区误差最大(24.47), 北碚区误差最小(18.93)。综上可见, 基于地面气象要素建立的 KNN 模型可以较好地模拟预测污染物的浓度。

表 2 KNN 训练效果统计检验

区域	准确率/%	决定系数	相对误差	均方根误差
巴南区	75.30	0.833	0.34	23.02
北碚区	78.30	0.892	0.28	18.93
璧山区	75.80	0.804	0.34	26.47
合川区	79.70	0.823	0.23	21.11
沙坪坝区	77.00	0.875	0.32	21.67
铜梁区	77.90	0.701	0.29	24.47
渝北区	76.90	0.841	0.28	21.87

图 3 为 2017—2018 年北碚区夏季臭氧浓度值与实测值趋势对比图, 从图中可看出, 臭氧浓度模拟值与实测值趋势基本一致, 但是有部分模拟值低于实测值。



图 3 2017—2018 年北碚区夏季臭氧浓度日均值的 KNN 模拟值与实测值趋势图

3.1.2 检验结果

对 2019 年研究区的臭氧浓度进行预测检验, 将预测值与实测值对比得到表 3, 由此可知, KNN 模型在检验模型中准确率略低于训练模型, 7 个区 KNN 模型的平均准确率为 74%, 比训练模型准确率低 3.3%, 其中北碚区的 KNN 模型准确率最高, 为 78.60%, 较训练模型高 0.3%; 铜梁区的 KNN 模型准确率最低, 只有 70.40%, 较训练模型低了 7.5%; 从决定系数来看, 北碚区决定系数在 0.9 左右, 铜梁区决定系数最低, 其余区的决定系数在 0.7~0.8 之间, 模拟效果较好。7 个区的模型平均相对误差为 0.33, 其中北碚区误差最小, 为 0.23, 渝北误差最大, 为 0.40; 7 个区 KNN 模型均方根误差与训练模型结果基本一致, 铜梁区误差最大, 北碚区误差最小。



表 3 KNN 检验效果统计检验

区域	准确率/%	决定系数	相对误差	均方根误差
巴南区	71.80	0.786	0.42	26.70
北碚区	78.60	0.905	0.23	16.89
璧山区	74.30	0.769	0.29	29.30
合川区	76.50	0.787	0.28	25.91
沙坪坝区	73.50	0.803	0.34	27.74
铜梁区	70.40	0.688	0.35	33.74
渝北区	72.90	0.756	0.40	25.69

图 4 为 2019 年北碚区夏季臭氧浓度的模拟值与实测值趋势图,从图中可看出,该区臭氧浓度模拟值与实测趋势基本一致,两者之间的误差波动不大,平均相对误差 0.18,其中最大误差 0.9.



图 4 2019 年北碚区夏季臭氧浓度日均值的 KNN 模拟值与实测值趋势图

3.2 BP 神经网络模型构建

3.2.1 训练结果

利用 BP 神经网络算法对 2017—2018 年研究区的臭氧浓度进行预测,将预测值与实测值对比得到表 4. 由表 4 可知, BP 神经网络算法在训练模型中准确率均在 74.57% 以上,7 个区的平均准确率为 74.6%,其中合川区的准确率最高,为 77.50%,璧山区的模型准确率最低,只有 71.60%,北碚区的准确率为 74.30%,略低于平均准确率;从决定系数来看,北碚区决定系数高于其余 6 个区的模型,铜梁区最低,只有 0.711,其余区的决定系数在 0.758~0.840 之间,模拟效果较好.7 个区的模型平均相对误差为 0.34,其中合川区误差最小为 0.26,巴南区误差最大为 0.39;7 个区的模型平均均方根误差为 22.50,其中璧山区误差最大(29.52),北碚区误差最小(21.59). 综上可见,基于地面气象要素建立的 BP 神经网络算法可以较好地模拟预测臭氧浓度.

表 4 BP 训练效果统计检验

区域	准确率/%	决定系数	相对误差	均方根误差
巴南区	72.80	0.776	0.39	26.69
北碚区	74.30	0.860	0.33	21.59
璧山区	71.60	0.758	0.38	29.52
合川区	77.50	0.800	0.26	23.10
沙坪坝区	75.10	0.840	0.36	24.48
铜梁区	75.80	0.711	0.31	24.17
渝北区	74.90	0.803	0.32	23.77

图 5 为 2017—2018 年北碚区夏季臭氧浓度的 BP 模拟值与实测值趋势图,从图中可看出,夏季臭氧浓度模拟值与实测值趋势基本一致,平均相对误差 0.18,其中最大误差 1.02.

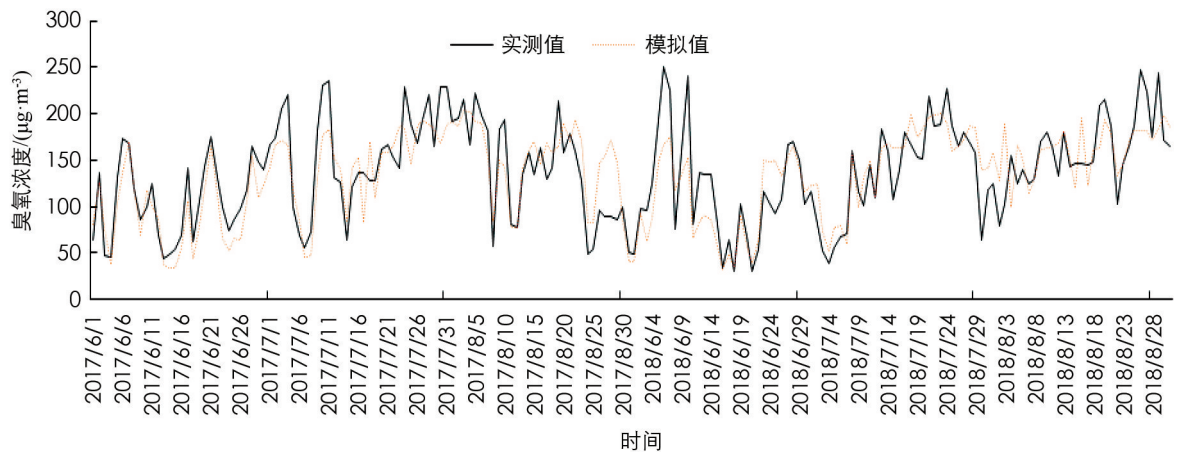


图 5 2017—2018 年北碚区夏季臭氧浓度日均值的 BP 模拟值与实测值趋势图

3.2.2 检验结果

对研究区 2019 年的臭氧浓度进行预测检验, 将预测值与实测值对比得到表 5. 由表 5 可知, BP 神经网络算法构建的模型在检验模型中准确率略高于训练模型, 7 个区的模型平均准确率为 75.68%, 比训练模型准确率高 1.1%, 其中合川区的模型准确率最高为 78.50%, 较训练模型高 1% 的准确率; 巴南区的模型准确率最低只有 73.90%, 较训练模型高了 1.1%, 北碚区的检验模型准确率也较训练模型高了 2.2%; 从决定系数来看, 北碚区系数在 0.86 以上, 高于其余 6 个区, 璧山区决定系数最低, 其余区的决定系数在 0.815~0.851 之间. 7 个区的模型平均相对误差为 0.29, 其中合川区误差最小(0.25), 巴南区误差最大(0.38); 7 个区的模型均方根误差与 KNN 模型结果一致, 铜梁区误差最大, 北碚区误差最小.

表 5 BP 检验效果统计检验

区域	准确率/%	决定系数	相对误差	均方根误差
巴南区	73.90	0.822	0.38	24.58
北碚区	76.50	0.892	0.26	18.19
璧山区	75.00	0.577	0.27	27.11
合川区	78.50	0.824	0.25	23.86
沙坪坝区	76.40	0.851	0.29	24.78
铜梁区	74.70	0.823	0.27	29.57
渝北区	74.80	0.815	0.37	22.18

图 6 为 2019 年北碚区夏季臭氧浓度的 BP 模拟值与实测值趋势图, 从图中可看出臭氧浓度模拟值与实测值趋势基本一致, 两者之间的误差波动在 8 月, 平均相对误差 0.18, 其中最大误差 0.9.

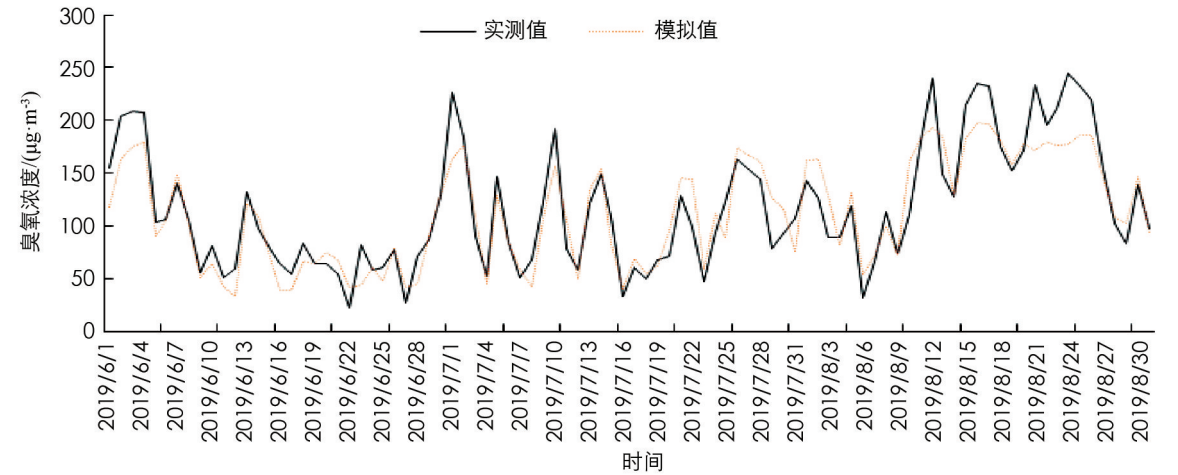


图 6 2019 年北碚区夏季臭氧浓度日均值的 BP 模拟值与实测值趋势图

3.3 模型对比

本文分别基于 KNN 数据挖掘算法与 BP 神经网络算法, 构建了 2017—2019 年臭氧浓度预测模型, 并对 2 种预测模型的预报效果进行对比检验, 筛选最优预测模型. 结果表明, 2 种预测模型在 7 个区的训练中平均预报准确度超过了 74%, 其中 KNN 模型的准确率高于 BP 神经网络模型, 2 种模型与实测值的相关性超过了 0.7, 拟合度较高. 在验证模型准确率中, KNN 预测模型趋势与 BP 神经网络模型比较, BP 神经网络模型在检验过程中的效果高于训练效果, 平均误差也低于训练值; 同时从北碚区预测结果来看, 基于 BP 神经网络模型的预测结果总体优于 KNN 模型. 总体来说, 基于 BP 神经网络模型更优于基于 KNN 数据挖掘算法的预测模型.

3.4 臭氧浓度网格化插值应用对比检验

根据上文分析, 以 BP 神经网络模型为预测模型, 选取研究区 2020 年 6 月 1 日平均气温、最高气温、相对湿度、降水量等 4 个气象因子为预测要素, 预测 6 月 1 日研究区臭氧浓度, 并采用 ArcGIS 空间分析模块中的反距离加权插值法计算、绘制未来 24 h 分辨率为 5 km 的研究区臭氧浓度网格预报图(图 7), 将预测值与实测值进行对比检验. 根据图 7a 可知, 北碚区未来 24 小时臭氧浓度在  $145\sim153\text{ }\mu\text{g}/\text{m}^3$  之间, 低值区出现在北碚城区, 高值区为研究区北部, 即与合川区交界的一带. 与北碚区 6 月 1 日实测值(图 7b)对比, 其余区的实测值在  $92\sim155\text{ }\mu\text{g}/\text{m}^3$  范围, 预测值高值与实测值基本一致, 低值差距较大. 从地理位置来看, 实测值的低值与预测的低值位置基本一致, 位于研究区中部偏西, 而实测高值预测区域差距较大, 位于研究区南部和东部. 造成预测误差的原因主要是因为臭氧是二次生成的光化学污染物, 生成和积累的时间主要集中在白天, 同时北碚区水土和蔡家这两个监测点位于发展区, 存有大量的在建工地, 易产生大量的 VOCs(挥发性有机物)不利于污染物扩散.

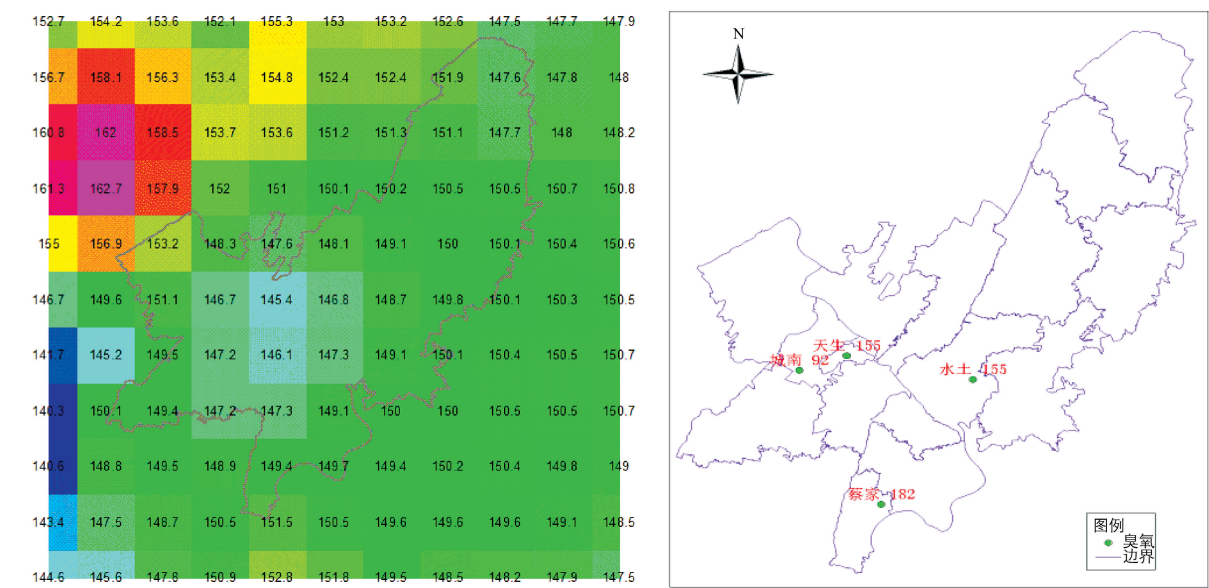


图 7 研究区与北碚区臭氧浓度网格预报对比图

4 结论

为探索不同机器学习方法在臭氧浓度预测中的应用效果, 本文利用 KNN 数据挖掘算法与 BP 神经网络算法, 挑选与臭氧密切相关的平均气温、最高气温、相对湿度、降水量等 4 个气象因子, 建立 2017—2019 年以重庆中心城区北碚区为中心的臭氧浓度预报模型, 并运用反距离加权插值法将臭氧浓度预报值网格化, 得到以下结论:



1) 2017—2019年间,除铜梁区和沙坪坝区呈逐年的上升趋势外,北碚区及其他区的趋势呈“单峰形”,最大值出现在2018年.北碚区及其周边区域的臭氧浓度季节变化是一致的,夏季浓度最高,冬季浓度最低,且逐月平均臭氧浓度呈单峰型,8月最高,12月最低.臭氧浓度与气象要素密切相关,除了降水量与臭氧浓度呈弱的负相关性外,其余3个气象要素均与臭氧浓度相关性较强,均大于0.6.

2) 基于KNN数据挖掘算法与BP神经网络算法,构建了2017—2019年臭氧浓度预测模型,并对2种预测模型的预报效果进行对比检验,筛选最优预测模型.结果表明,2种预测模型在7个区的训练中平均预报准确度超过了74%,其中KNN模型的准确率高于BP神经网络模型,2种模型与实测值的相关性超过了0.7,拟合度较高.在验证模型准确率中,KNN模型预测趋势与BP神经网络模型比较,BP神经网络模型在检验过程中的效果高于训练效果,平均误差也低于训练时;同时从北碚区预测结果来看,基于BP神经网络模型的预测结果总体优于KNN模型.

3) 采用ArcGIS空间分析模块中的反距离加权插值法计算北碚区臭氧浓度网格预报,发现北碚区域内未来24小时臭氧浓度预测在 $145\sim 153\text{ }\mu\text{g}/\text{m}^3$ 之间,高值与实测值基本一致,低值低于实测值.从地理位置来看,低值与实测值位置基本一致,在中部偏西,高值预测区域差距较大,实测高值区在南部和东部.

综上所述,在运用2种机器学习方法对臭氧浓度进行模拟预测中,BP神经网络算法在臭氧浓度的预测中有更好的表现;其次臭氧作为二次污染物,受气象和前体物等因素的影响的同时,也受污染物区域传输以及传输过程中前体物的光化学反应的重要影响<sup>[29]</sup>,而预测模型主要基于气象条件的影响,因此预测值较实测值有所偏差.

## 参考文献:

- [1] 王旭东,尹沙沙,杨健,等.郑州市臭氧污染变化特征、气象影响及输送源分析[J].环境科学,2021,42(2):604-615.
- [2] MOURA B B, ALVES E S, MARABESI M A, et al. Ozone Affects Leaf Physiology and Causes Injury to Foliage of Native Tree Species from the Tropical Atlantic Forest of Southern Brazil [J]. Science of the Total Environment, 2018(610-611): 912-925.
- [3] 徐国强,朱乾根,李晓燕.南北极区和青藏高原臭氧变化与中国降水和温度的联系[J].气象,2004,30(1):8-12.
- [4] 纪飞,秦瑜.对流层臭氧研究进展[J].气象科技,1998,26(4):17-24.
- [5] 肖辉,沈志来,黄美元,等.我国西南地区地面和低层大气臭氧的观测分析[J].大气科学,1993,17(5):621-628.
- [6] RAFIQ L, TAJBAR S, MANZOOR S. Long Term Temporal Trends and Spatial Distribution of Total Ozone over Pakistan [J]. The Egyptian Journal of Remote Sensing and Space Science, 2017, 20(2): 295-301.
- [7] MICHAEL R, OLENICK C R, MONAGHAN A, et al. Application of Geostatistical Approaches to Predict the Spatio-Temporal Distribution of Summer Ozone in Houston, Texas [J]. Journal of Exposure Science & Environmental Epidemiology, 2019, 29(6): 806-820.
- [8] 刘建,吴兑,范绍佳,等.前体物与气象因子对珠江三角洲臭氧污染的影响[J].中国环境科学,2017,37(3):813-820.
- [9] 吴锴,康平,王占山,等.成都市臭氧污染特征及气象成因研究[J].环境科学学报,2017,37(11):4241-4252.
- [10] 王磊,刘端阳,韩桂荣,等.南京地区近地面臭氧浓度与气象条件关系研究[J].环境科学学报,2018,38(4):1285-1296.
- [11] SHEN J, WANG X S, LI J F, et al. Evaluation and Intercomparison of Ozone Simulations by Models-3/CMAQ and CAMx over the Pearl River Delta [J]. Science China Chemistry, 2011, 54(11): 1789-1800.
- [12] MAR K A, OJHA N, POZZER A, et al. Ozone Air Quality Simulations with WRF-Chem (V3. 5. 1) over Europe: Model Evaluation and Chemical Mechanism Comparison [J]. Geoscientific Model Development, 2016, 9(10): 3699-3728.
- [13] 杜勤博,陈欢欢,李玥莹,等. BP神经网络模型在 $\text{O}_3$ 质量浓度预报中的应用[J].广东气象,2019,41(3):29-32.
- [14] 杜云松,罗彬,陈建文,等.气温在成都地区臭氧预报的运用研究[J].环境科学与技术,2017,40(S1):329-334.
- [15] 梅梅,徐大海,朱蓉,等.减排措施与气象因子对2013—2019年中国大陆地区 $\text{PM}_{2.5}$ 浓度变化的贡献[J].环境科学学报,2021,41(7):2519-2529.

- [16] 高志伟, 杨坤琳, 潘坤, 等. 基于自然邻点插值法的洱海湖面降水特征分析 [J]. 气象科技, 2021, 49(5): 799-805.
- [17] 张祖莲, 张山清, 毛炜峰, 等. 基于平均滤波算法的北疆春季日最低气温格点预报产品订正分析 [J]. 沙漠与绿洲气象, 2021, 15(5): 16-23.
- [18] 陈豫英, 陈楠, 杨银, 等. 贺兰山东麓春季最低气温和霜冻精细化格点预报产品的检验 [J]. 自然灾害学报, 2019, 28(1): 198-206.
- [19] 王增民, 王开珏. 基于熵权的 K 最临近算法改进 [J]. 计算机工程与应用, 2009, 45(30): 129-131, 160.
- [20] 尹晓梅, 李梓铭, 熊亚军, 等. 2014—2017 北京市气象条件和人为排放变化对空气质量改善的贡献评估 [J]. 环境科学, 2019, 40(3): 1011-1023.
- [21] 熊亚军, 廖晓农, 李梓铭, 等. KNN 数据挖掘算法在北京地区霾等级预报中的应用 [J]. 气象, 2015, 41(1): 98-104.
- [22] 吴亚平, 张琦, 王炳赞, 等. 四川雅安三种主要大气污染物浓度与气象条件的关系及其预测研究 [J]. 高原气象, 2020, 39(4): 889-898.
- [23] 苏高利, 邓芳萍. 论基于 MATLAB 语言的 BP 神经网络的改进算法 [J]. 科技通报, 2003, 19(2): 130-135.
- [24] 韩余, 周国兵, 陈道劲, 等. 重庆市臭氧污染及其气象因子预报方法对比研究 [J]. 气象与环境学报, 2020, 36(4): 59-66.
- [25] 李萍, 曾令可, 税安泽, 等. 基于 MATLAB 的 BP 神经网络预测系统的设计 [J]. 计算机应用与软件, 2008, 25(4): 149-150, 184.
- [26] 吕蓓红, 曹三忠, 张勇. 环境空气质量预报准确率研究 [J]. 信阳师范学院学报(自然科学版), 2006, 19(4): 499-503.
- [27] 李翔宇, 李翌, 包艳英, 等. 基于反距离加权插值法评价海域水质类别空间分布 [J]. 中国环境监测, 2019, 35(6): 70-77.
- [28] 史达伟, 李超, 史逸民, 等. 基于机器学习的大雾天气背景下特强浓雾本地化诊断研究 [J]. 灾害学, 2018, 33(2): 193-199.
- [29] 杨显玉, 易家俊, 吕雅琼, 等. 成都市及周边地区严重臭氧污染过程成因分析 [J]. 中国环境科学, 2020, 40(5): 2000-2009.

责任编辑 胡杨 崔玉洁