

耦合注意力机制DNN的PM_{2.5}估算及时空特征分析

陈镔捷¹, 叶扬¹, 林溢², 游诗雪¹, 邓劲松^{1,3,4}, 杨武^{1,4}, 王珂¹

1. 浙江大学 环境与资源学院, 杭州 310058;

2. 浙江工业大学 环境学院, 杭州 310014;

3. 浙江生态文明研究院, 安吉 313300;

4. 浙江大学 智慧生态与绿色发展研发中心, 杭州 310058

摘要: PM_{2.5}作为指示环境质量的重要因子之一, 不仅影响着灰霾天气的发生, 还与公众健康息息相关, 近年来受到广泛的关注。尽管PM_{2.5}地面观测站点在不断地扩张, 其覆盖范围依旧有限, 难以反映全域PM_{2.5}浓度的时空异质性。本研究运用卫星遥感气溶胶光学厚度数据, 辅助因子除常规的气象因子等以外, 还加入了针对中国人民生产生活习惯的农历日因子, 提出一种耦合注意力机制的深度神经网络模型, 对长三角区域2015年—2020年PM_{2.5}浓度进行了逐日的高精度估算。模型交叉验证结果显示决定系数R²高达0.85, 斜率0.86, 与地面站点观测值有较高的一致性, 优于多元线性回归和随机森林模型。长三角区域PM_{2.5}浓度时空特征分析结果表明, PM_{2.5}浓度在空间上呈现北高南低的趋势; 季节特征以冬季浓度最高, 夏季浓度最低, 春秋过渡。此外, 长三角区域2015年—2020年整体PM_{2.5}浓度呈下降趋势, 其中以上海市最为明显, 下降速率为3.30 μg/(m³·a), 其次为江苏省(2.65 μg/(m³·a)); 浙江省与安徽省下降速率都小于2 μg/(m³·a), 但由于安徽省PM_{2.5}浓度远高于浙江省, 提升空间更大, 需要更多的关注。综上所述, 利用卫星数据结合本研究提出的方法能弥补地面观测站点的不足, 获得高精度全域PM_{2.5}浓度时空分布特征, 从而更科学地指导相关政策的规划与落地。

关键词: 遥感, 气溶胶光学厚度 (AOD), 深度学习, 注意力机制, 长三角区域, 空气质量

引用格式: 陈镔捷, 叶扬, 林溢, 游诗雪, 邓劲松, 杨武, 王珂. 2022. 耦合注意力机制DNN的PM_{2.5}估算及时空特征分析. 遥感学报, 26(5): 1027-1038

Chen B J, Ye Y, Lin Y, You S X, Deng J S, Yang W and Wang K. 2022. Spatiotemporal estimation of PM_{2.5} using attention-based deep neural network. National Remote Sensing Bulletin, 26(5): 1027-1038 [DOI: 10.11834/jrs.20221362]

1 引言

PM_{2.5}是指悬浮于空气中的空气动力学直径小于等于2.5 μm的细颗粒物干质量浓度。有大量研究表明, 由于PM_{2.5}能够进入到人体呼吸道深处, 同时极易附着大量有毒有害物质, 长期暴露在较高PM_{2.5}环境中, 会导致人类寿命的缩短, 以及各种呼吸疾病、心血管疾病的发生 (Dockery等, 1993)。在最新全球疾病负担 (Global Burden of Disease Study) 调查中, PM_{2.5}造成的死亡人数及其影响位列所有环境和职业风险因子中的首位 (GBD 2019 Risk Factors Collaborators, 2020)。此外, PM_{2.5}也是造成灰霾的主要原因, 是最直观反

映环境质量的要素之一。

自改革开放以来, 中国经济发展迅速, 快速工业化和城镇化进程中不可避免地对环境造成了污染。尽管自2013年以来政府出台了一系列相关政策来控制与监测空气质量, 地面监测站的覆盖范围也越来越大, 但是这些地面固定站点更多分布于城市中, 对于全方位监测空气质量还是过于稀疏。卫星遥感技术的蓬勃发展, 使得对空气站点缺失区域的空气质量监测成为可能, 为环境空气质量的大面积同步监测提供了全新的技术手段。利用卫星遥感获得的气溶胶光学厚度 (AOD) 产品, 构建AOD-PM_{2.5}的模型来估算全域PM_{2.5}, 可以有效弥补地面站点相对稀疏的缺陷, 获得更大范围时空连续的空气质

收稿日期: 2021-05-30; 预印本: 2022-01-09

基金项目: 国家重点研发计划(编号:2020YFC1807500), 浙江省重点研发计划(编号:2022C03078)

第一作者简介: 陈镔捷, 研究方向为环境与资源遥感。E-mail:bj_chen@zju.edu.cn

通信作者简介: 邓劲松, 研究方向为环境与资源遥感。E-mail:jsong_deng@zju.edu.cn

数据。目前,常用的AOD产品有:中分辨率成像光谱仪(MODIS)、先进的葵花成像仪(AHI)、可见光红外成像辐射仪(VIIRS)的气溶胶产品,空间分辨率在1 km左右;也有学者利用Landsat、高分一号等卫星获取高空间分辨率的AOD产品,但其重返周期较长,应用相对受限。而MODIS卫星自2000年起即有观测数据,同时能够提供每日的全球覆盖数据,因此其AOD产品被广泛应用于空气质量的估算研究中。MODIS卫星提供的AOD产品经过较长时间的发展,已经从最初的MOD04暗像元算法产品,发展到最新的多角度大气校正算法(MAIA) AOD产品,精度提升的同时空间分辨率也从10 km提升至1 km (Lyapustin等, 2018)。

准确构建AOD与 $PM_{2.5}$ 的关系模型是估算空气质量的关键。初期,学者尝试基于物理机理进行关系的构建,包括对AOD的湿度、标高及细模态粒子比例的订正等,建立具备严密逻辑的物理关系方程(Li等, 2016; Lin等, 2015; 张莹和李正强, 2013)。但由于物理机理过程复杂,模型中大量参数难以获得,只能依靠近似的计算,存在较大的不确定性,精度一般较低。随着地面 $PM_{2.5}$ 观测数据的日渐增多,越来越多的学者利用高效的统计模型,同时结合大量辅助因子来建立AOD与 $PM_{2.5}$ 的关系。多元线性回归模型最先被用于AOD- $PM_{2.5}$ 关系的构建,但由于AOD与 $PM_{2.5}$ 呈非线性关系,关系较为复杂,简单的线性回归模型无法很好地拟合这一关系(Liu等, 2007)。马宗伟(2015)、Just等(2015)、杨立娟等(2018)利用线性混合效应模型(LME)和广义加和模型(GAM),并结合气象要素、土地利用类型等参数建立更为复杂的非线性模型,估算得到的 $PM_{2.5}$ 浓度精度大幅提升。He和Huang(2018)利用地理加权回归(GWR)及其改进版本模型,开展了 $PM_{2.5}$ 浓度模拟研究,同样取得了优于简单空间插值与线性回归的结果。近年来,机器学习算法得到了快速发展,相比于传统的统计模型,机器学习模型有更强大的非线性拟合能力,因此也被广泛应用于大气环境质量研究中。Zang等(2019)利用广义回归神经网络结合AOD、气象要素、高程(DEM)、归一化植被指数(NDVI)等因子模型估算了全国 PM_{10} 的浓度。Wei等(2020)构建时空极端回归树模型估算得到全国长时序的 $PM_{2.5}$ 空间分布产品,同样用到了大量的辅助因子数据。

尽管这些模型都取得了不错的效果,但由于大气环境的时空异质性较强,参数固定的机器学习模型难以捕获AOD与 $PM_{2.5}$ 随时空变化的关系(Fang等, 2016),一定程度上限制了机器学习模型的可推广性。注意力机制是一种广泛应用于自然语言处理,计算机视觉任务中的结构,其目的为使模型能够动态地将注意力放在感兴趣的区域,即动态识别更重要的因子,进而增强模型的鲁棒性与解释性(Vaswani等, 2017; 朱均安, 2020)。为解决参数固定的机器学习模型无法考虑时空异质性的问题,本研究提出一种结合注意力机制的深度神经网络模型(DNN)来构建AOD- $PM_{2.5}$ 的关系,模型中的注意力机制能够在时空动态中识别对于建模影响更为重要的因子,从而估算得到更可靠的 $PM_{2.5}$ 浓度。此外,本研究选取的辅助因子除了常规的DEM、NDVI和气象要素等因子外,还加入了针对于中国人民生产生活习惯的农历日因子。最后利用这些数据构建得到AOD- $PM_{2.5}$ 的关系模型,估算长三角区域2015年—2020年 $PM_{2.5}$ 浓度的时空分布,探索长三角区域空气质量的时空变化特征。

2 研究区和数据

2.1 研究区概况

研究区为长三角地区包括“三省一市”,即江苏省、浙江省、安徽省和上海市,其作为一带一路和长江经济带的交汇地带,是中国经济最为发达的区域之一,同时也是空气污染相对严重的区域之一。随着城市区域的不断扩张和人口的不断聚集,长三角区域高密度人口更易受到空气污染的影响,因此也对空气污染的监测提出了更高的要求。利用卫星遥感手段对长三角区域空气质量进行全域估算,能够弥补地面站点的缺陷,从而更科学地指导长三角地区空气污染防治政策的制定,分析空气污染对人体健康的影响。研究区地理空间范围与 $PM_{2.5}$ 监测站点分布如图1所示。

2.2 AOD数据

本研究采用MODIS卫星最新的MAIA气溶胶产品,相比于此前版本的MOD04气溶胶产品,MAIA AOD具有更高的空间分辨率(1 km),更高的数据质量与更广的覆盖范围,能更好地用于 $PM_{2.5}$ 的估算。覆盖长三角区域自2015-01-01—

2020-12-31的每日MAIAC数据从NASA网站下载获得(<https://ladsweb.modaps.eosdis.nasa.gov/>[2021-05-30]),通过重投影和裁剪操作,提取获得研究区范围内的AOD值。由于已经有大量研究验证了MAIAC AOD在长三角区域的精度,本研究不再对其精度进行验证(Xiao等,2017)。

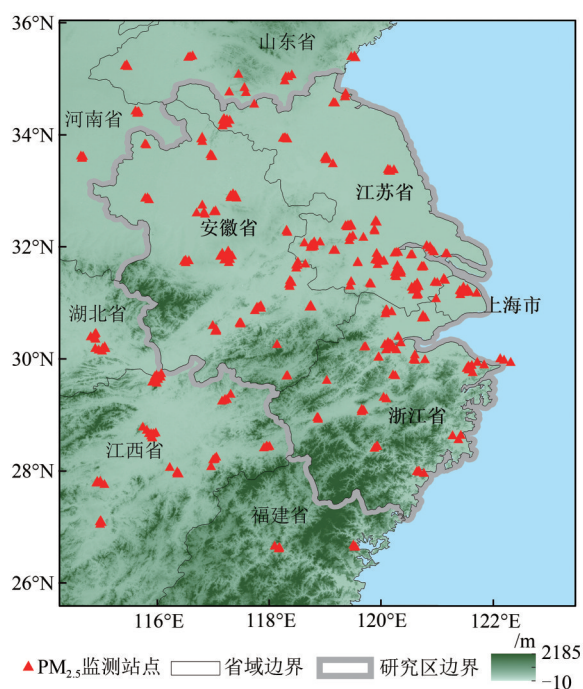


图1 研究区范围与PM_{2.5}监测站点分布

Fig.1 Study area and distribution of PM_{2.5} monitoring sites

2.3 地面PM_{2.5}监测数据

2015-01-01—2020-12-30每日的逐小时地面PM_{2.5}监测数据来源于国家生态环境部的“全国城市空气质量实时发布平台”(<http://www.cnemc.cn/>[2021-05-30]),选取研究区范围及周边共计352个国控空气质量监测站点作为本研究的地面观测PM_{2.5}浓度数据来源(图1)。

2.4 其他辅助数据

气象条件对PM_{2.5}的形成与扩散有重要的影响,本研究所用气象数据来源于欧洲中期天气预报中心ECMWF(<https://cds.climate.copernicus.eu/>,[2021-05-30])最新发布的第五代全球气候再分析数据集ERA5,从中提取2015年—2020年每日的地表2 m处气温(T2M)、地表10 m处东西向风速(WU)、南北向风速(WV)、相对湿度(RH)、行星边界层高度(PBLH)、地表气压(SP)作为构建AOD与PM_{2.5}关系模型的气象参数,空间分辨率为

0.125°。此外,数字高程模型(DEM)从地理空间数据云(<http://www.gscloud.cn/>[2021-05-30])上获得,空间分辨率30 m。2015年—2020年每月的1 km分辨率NDVI同AOD数据从NASA网站下载获得。

2.5 数据预处理

由于AOD受云等干扰存在大量缺失现象,导致无法获得全域覆盖的结果,为解决这一问题,本研究利用ECMWF的再分析数据中的550 nm AOD与现有每日MAIAC AOD建立线性关系模型,进而插补得到每日完整的MAIAC AOD数据,再进行后续AOD-PM_{2.5}建模。将AOD插补结果与AERONET地面站点进行验证,结果显示插补前二者相关系数(R)为0.85,均方根误差(RMSE)为0.17,插补后 R 为0.76, RMSE为0.28,尽管精度略有下降,但仍可以满足后续使用要求。由于气象数据、DEM、AOD与NDVI空间分辨率各不相同,而模型的构建需要这些数据一一对应,因此将所有栅格数据统一利用双线性插值法重采样至与MAIAC AOD相同的1 km分辨率。地面监测的PM_{2.5}数据可能存在由于设备故障导致的数据缺失,为保证数据精度,剔除每日观测数小于12个小时的数据,再对剩余的数据进行日均值的计算。由于1个1 km栅格内可能存在多个地面站点,计算这些站点的均值作为该栅格最终地面实测PM_{2.5}日均浓度。将所有预处理完的数据,进行时空匹配,即根据每日PM_{2.5}站点数据,提取其所在栅格的包括气象数据,DEM,AOD与NDVI的对应数据结果,构建模型所需数据集。最后对数据集进行Z-score标准化处理,使得每个因子的均值为0,标准差为1。

3 研究方法

本研究基于DNN加入注意力机制构建了自适应深度神经网络模型SADNN(Self-Adaptive DNN)对AOD-PM_{2.5}的关系进行建模。DNN属于机器学习模型的一种,由一个输入层,多个隐藏层和一个输出层构成,每一层又由不同个数的神经元构成,层内神经元互不连接,层间神经元通过全连接与激活函数的组合,对数据进行非线性的变换(LeCun等,2015)。DNN的这种结构可以使模型学习到数据的多层次抽象特征,拥有强大的非线性拟合能力,从而表达复杂的模式,其基本原理可表示为

$$X_i^{j+1} = g\left(\sum_{i=0}^{s_j} \theta_{ii} X_i^j\right) \quad (1)$$

$$t = 0, 1, 2, \dots, s_{j+1}; j = 0, 1, 2, \dots, n$$

式中, X_i^j 表示第 j 层的第 i 个神经元, s_j 表示第 j 层神经元的总数, θ_{ii} 为神经元 X_i^j 和 X_i^{j+1} 之间的权重, $g(\cdot)$ 代表激活函数, n 为DNN总的层数。

DNN模型的训练通过后向传播更新神经元间的权重实现,早期的激活函数(如Sigmoid)和后向传播算法(如随机梯度下降)容易出现梯度消失或弥散的情况,导致深度模型难以训练,因此只能训练浅层模型,如人工神经网络(ANN)。将整流线性单元(ReLU)作为激活函数,利用均方根传播(RMSProp)算法进行后向传播的计算可以有效地缓解这一现象,同时加快训练速度,使得DNN的训练成为可能(Reddy等, 2018; Nair和Hinton, 2010)。因此本研究激活函数选用ReLU,后向传播算法选择RMSProp。

注意力机制最早被提出应用于自然语言处理任务上,它能够自动识别上下文的依赖关系,进而提升模型的准确性,同时使得用户能够了解模型识别了哪些关键词语来理解整个句子,增强了模型的可解释性(Vaswani等, 2017)。在语音情感识别(Speech emotion recognition)研究中,注意力机制能够在一大段语音中识别目标情感的关键区域,提取关键情感特征(Li等, 2021)。随后注意力机制也被广泛迁移至计算机视觉任务中,注意力机制在计算机视觉中可分为空间注意力和通道注意力,空间注意力能够寻找图片中对于结果最重要的部位,通道注意力机制的本质在于构建不同通道特征之间的重要性的模型,对输入进行特征分配,选择对任务更有用的特征(Fu等, 2019; 朱均安, 2020)。此外,在遥感应用上, Ye等(2019)结合了空间注意力和通道注意力对遥感影像中的建筑物进行了高精度提取; Mou和Zhu(2020)提出光谱注意力模块,对超光谱数据的各个波段进行校正,对重要的波段赋予较大权重,不重要的波段赋予较小权重,以获得更好的地物分类精度。注意力机制的原理为首先通过全连接层识别因子间的相互关系,然后通过一定的变化函数如Sigmoid函数计算得到每个因子的重要性,最后将原始因子与重要性相乘即得到校正后的特征, Sigmoid表示如下:

$$g(z) = \frac{1}{1 + e^{-z}} \quad (2)$$

式中, $g(z)$ 为变量 z 的Sigmoid函数, e 代表自然常数。

本研究将注意力机制融入到DNN中,构建SADNN模型,其目的就在于识别对于 $PM_{2.5}$ 的估算更重要的特征,进而更好地构建AOD- $PM_{2.5}$ 关系。输入的因子首先经过注意力机制的校正,对原始因子中重要的部分赋予一个较高的权重,不重要的部分赋予一个较低的权重,随后输入因子与其对应的权重相乘得到校正后的因子,再输入到后续网络中。这样模型即可以自适应地时空动态校正各输入因子,从而捕获AOD- $PM_{2.5}$ 关系的时空异质性。模型结构采用启发式搜索确定,将隐藏层的层数设为1,并以1为步长进行迭代训练验证,隐藏层神经元从16个开始,每次增加1倍。最终综合考虑模型精度与效率,确定为包含4个隐藏层,每层1024个神经元的结构,整个SADNN模型结构如图2。

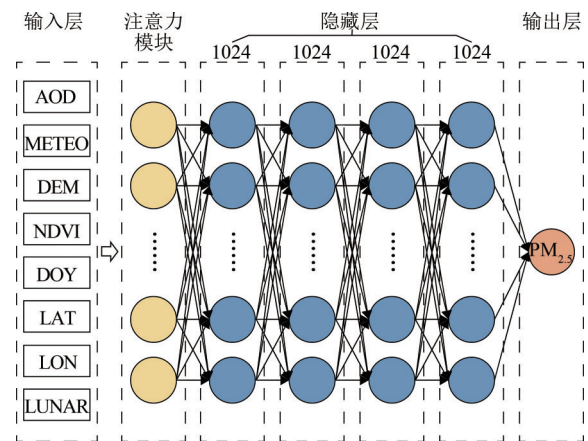


图2 SADNN模型结构

Fig.2 The structure of SADNN model

模型的输入因子除了AOD、DEM、NDVI和上述提到的气象因子外,还加入了积日(DOY)、经度(LON)和纬度(LAT),此外还加入了针对于中国人民生产生活习惯的农历日(LUNAR)作为一个因子,来更好地体现人为活动对 $PM_{2.5}$ 浓度的影响。农历作为中国传统历法,根据月相变化进行编排,既能反映季节变化,又指导着物候和农事。此外,中国许多法定节假日根据农历安排,影响着人们的生活习惯,例如大年三十与正月初一,中国普遍存在集体燃放烟花爆竹的习俗,会对大气环境造成严重的影响。模型可表示为

$$\text{PM}_{2.5_s} = f(\text{AOD}_{st}, \text{DEM}_s, \text{NDVI}_{st}, \text{METEO}_{st}, \text{LON}_s, \text{LAT}_s, \text{DOY}_t, \text{LUNAR}_t) \quad (3)$$

式中， $\text{PM}_{2.5_s}$ 表示栅格 s 在时间 t 时的日均 $\text{PM}_{2.5}$ 浓度， METEO 表示气象因子，包括前文提到的T2M、WV、WD、RH、PBLH和SP。

此外本研究利用上述相同因子，构建多元线性回归模型与随机森林模型进行对比分析。多元线性回归模型是构建AOD- $\text{PM}_{2.5}$ 关系的最基础模型(Liu等, 2007)，可表示为

$$\text{PM}_{2.5_s} = \beta_1 \text{AOD}_{st} + \beta_2 \text{DEM}_s + \beta_3 \text{NDVI}_{st} + \beta_4 \text{METEO}_{st} + \beta_5 \text{LON}_s + \beta_6 \text{LAT}_s + \beta_7 \text{DOY}_t + \beta_8 \text{LUNAR}_t + \alpha \quad (4)$$

式中， α 为截距， $\beta_1, \beta_2, \dots, \beta_8$ 为每个因子的斜率。

随机森林是近年最为流行的机器学习模型之一，它通过集合多个决策树的结果，以有放回抽样的方式进行训练，能够快速准确地学习数据模式，进行回归或分类，被广大学者用于 $\text{PM}_{2.5}$ 浓度的估算中(Hu等, 2017)。其作为非参数模型，公式表达同式(3)。

模型精度的定量评价选取决定系数(R^2)、均方根相对误差(RMSRE)、均方根误差(RMSE)和平均绝对误差(MAE)作为指标，公式如下

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_{i,\text{Meas}} - y_{i,\text{Esti}})^2}{\sum_{i=1}^n (y_{i,\text{Meas}} - \overline{y_{i,\text{Meas}}})^2} \quad (5)$$

$$\text{RMSRE} = \sqrt{\frac{1}{n} \sum_{i=1}^n \left(\frac{y_{i,\text{Meas}} - y_{i,\text{Esti}}}{y_{i,\text{Meas}}} \right)^2} \quad (6)$$

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_{i,\text{Meas}} - y_{i,\text{Esti}})^2} \quad (7)$$

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |y_{i,\text{Meas}} - y_{i,\text{Esti}}| \quad (8)$$

式中， n 代表数据个数， $y_{i,\text{Meas}}$ 代表第 i 条实测 $\text{PM}_{2.5}$ 值， $y_{i,\text{Esti}}$ 代表第 i 条估算的 $\text{PM}_{2.5}$ 值， $\overline{y_{i,\text{Meas}}}$ 代表所有实测 $\text{PM}_{2.5}$ 的平均值。

模型的验证采用5折交叉验证：将整个数据集随机平均分为5份，轮流选取其中1份作为验证集，剩余4份作为训练集进行模型的训练，最后将5次结果进行整合得到最终的模型精度。通过不同模型比较，选取验证精度相对最优的模型，作为最终确定的模型。再利用完整的数据集，对确定的模型进行训练，应用于后续长三角区域 $\text{PM}_{2.5}$ 的估算，得到研究区域的 $\text{PM}_{2.5}$ 时空分布。

4 结果和分析

4.1 模型精度验证

将2015-01-01—2020-12-31研究区范围内所有数据清理整合，最终得到有效匹配数据共341768条用于模型的训练、验证与比较，结果如图3所示。本研究提出的SADNN模型交叉验证结果 R^2 高达0.85，RMSE为 $12.77 \mu\text{g}/\text{m}^3$ ，RMSRE为0.35，MAE仅 $8.22 \mu\text{g}/\text{m}^3$ 。作为流行的机器学习模型，随机森林也达到了一个不错的验证结果， R^2 为0.82，RMSRE为0.36，但其拟合线的斜率(0.74)相比于SADNN的0.86要低不少，说明随机森林模型更易低估 $\text{PM}_{2.5}$ 高值，其估算能力相对较弱。传统的多元线性回归模型表现最差，不能很好地对AOD- $\text{PM}_{2.5}$ 关系进行建模， R^2 仅为0.30，RMSRE与RMSE分别高达0.78与 $27.25 \mu\text{g}/\text{m}^3$ 。由此可见，本研究提出的SADNN模型利用注意力机制与深层网络模型结构，通过对大规模数据模式的学习，对于估算 $\text{PM}_{2.5}$ 的浓度相较随机森林和多元线性回归模型更为精准，误差更小，有更高的应用与推广潜力。

为进一步说明注意力机制与农历日因子的有效性，同时验证SADNN模型对历史逐日 $\text{PM}_{2.5}$ 浓度估算的能力，选取2016-02-06—2016-02-08(腊月廿八、除夕与春节)3d的不同模型估算结果进行对比分析，模型包括SADNN、不含农历日因子的SADNN(SADNN-noLunar)、不含注意力的SADNN(SADNN-noAtt)(图4)。与地面站点的对比验证显示，腊月廿八、除夕与春节SADNN的 R^2 分别为0.78, 0.73, 0.68; SADNN-noLunar的 R^2 分别为0.72, 0.73, 0.60; SADNN-noAtt的 R^2 分别为0.75, 0.73, 0.64。可以看到，3日中除了除夕3个模型站点验证结果不相上下外，其余二日SADNN模型的结果要明显更优，加入农历日因子后二日 R^2 有0.06和0.08的提升，加入注意力后二日 R^2 提升了0.03和0.04。除夕深夜与春节凌晨中国有燃放烟花爆竹的习俗，极易导致春节白天空气质量的显著下降，加入农历日因子的模型更好地捕捉了这一特征。从空间分布趋势来看，SADNN模型在腊月廿八成功估算了浙江北部的高值区域，另外两个模型估算的高值区域明显偏小。除夕当日江苏省空气污染严重，地面监测站大部分都超过

80 $\mu\text{g}/\text{m}^3$, 尽管3个模型与地面站点验证结果表现不相上下, SADNN对江苏省高值区域的估算要明显优于另外两个模型。春节当日, SADNN-noAtt在安徽西南山区有错误的高估, 而SADNN-noLunar对研究区北部的高值区域有明显的低估。综上所述

述, 农历日因子和注意力的加入能够对模型起到一定程度的优化作用, 尤其是在 $\text{PM}_{2.5}$ 的高值区域和空间细节方面, 同时模型在单日 $\text{PM}_{2.5}$ 的空间估算上与地面站点有较强的一致性, 能够较为精准地估算研究区范围每日 $\text{PM}_{2.5}$ 的浓度。

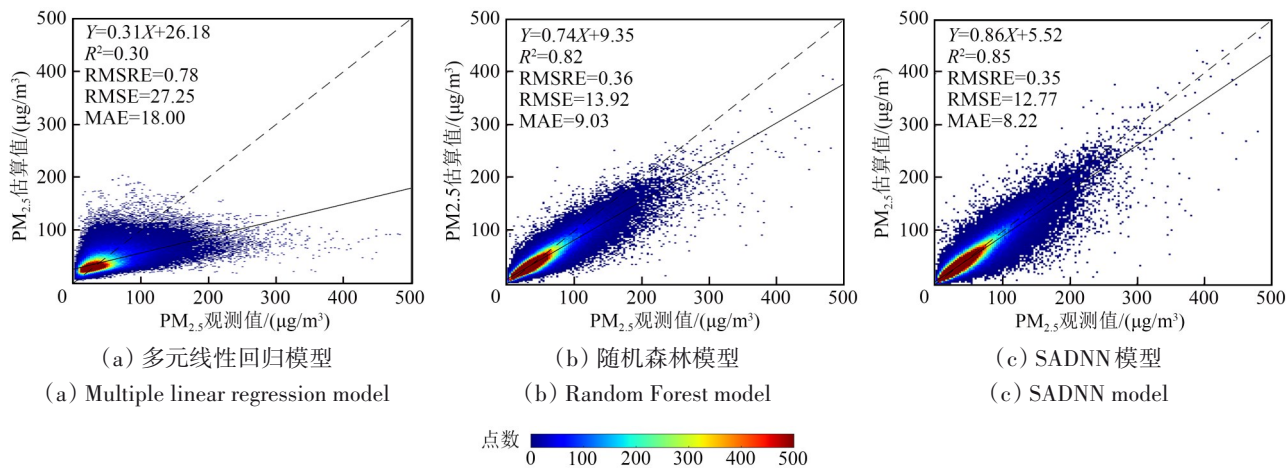
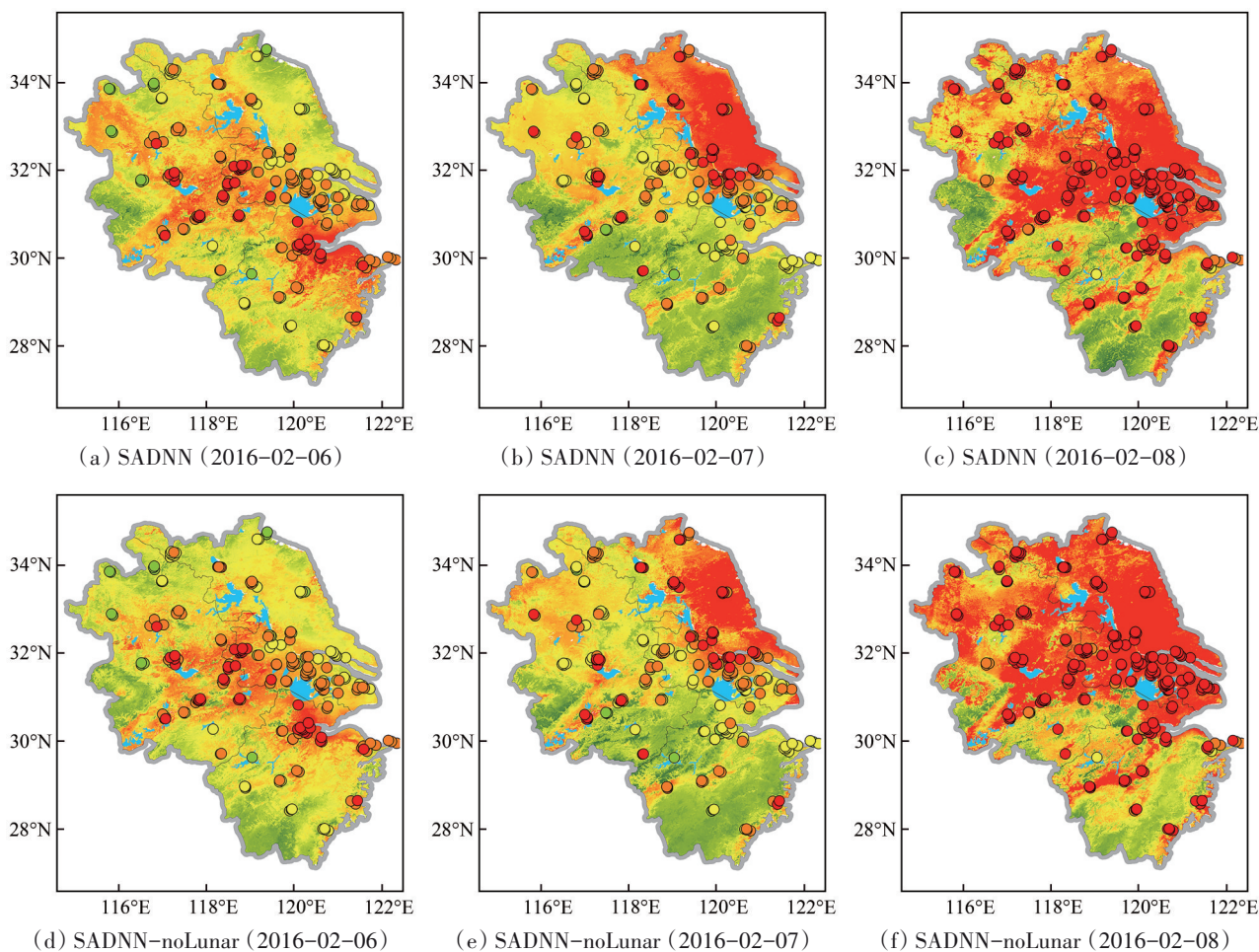


图3 不同模型交叉验证结果

Fig.3 Cross-validation results for different models



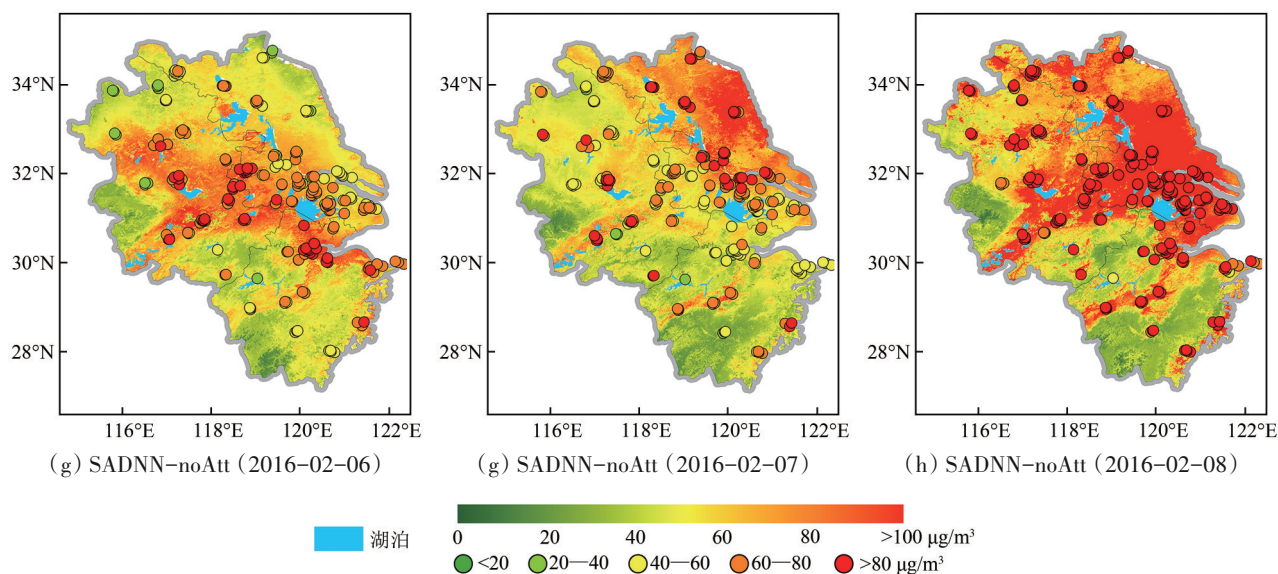


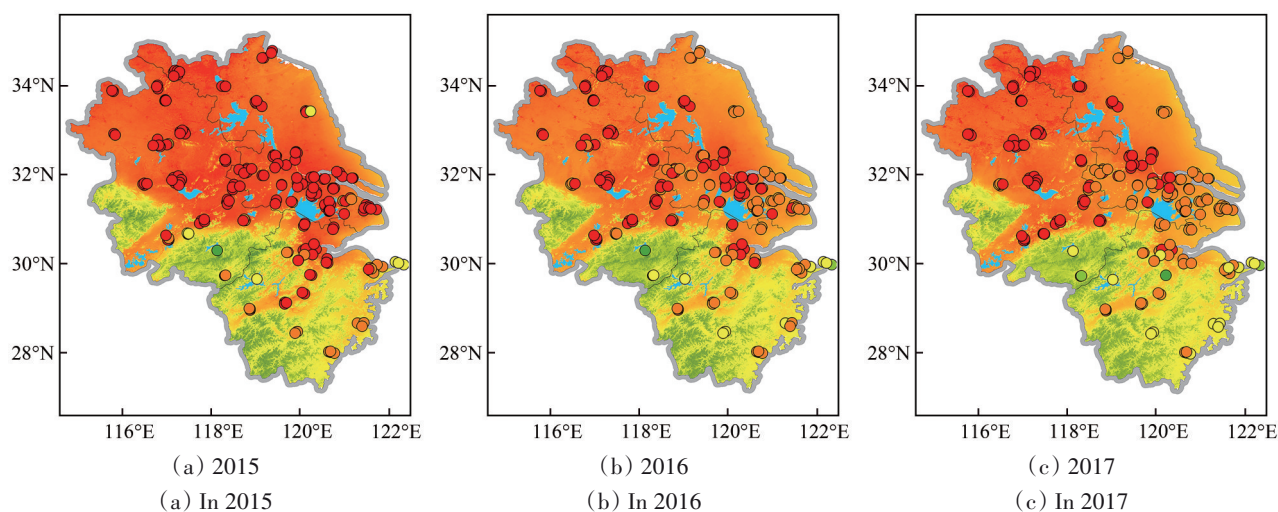
图 4 不同模型对 2016 年典型日期的 PM_{2.5} 估算

Fig. 4 PM_{2.5} estimation for different models at typical days in 2016

4.2 年际 PM_{2.5} 浓度时空分布特征

基于 SADNN 模型估算得到的每日 PM_{2.5} 浓度数据，计算长三角区域 2015 年—2020 年每年的 PM_{2.5} 浓度均值时空分布，结果如图 5 所示。估算的 PM_{2.5} 浓度空间分布上与地面观测站点保持一致，整体呈现北高南低的趋势，安徽省与江苏省空气污染更为严重。2015 年—2020 年估算得到的长三角区域 PM_{2.5} 均值分别为：44.88 μg/m³、40.91 μg/m³、41.27 μg/m³、39.34 μg/m³、37.64 μg/m³、34.07 μg/m³，除 2017 年较 2016 年有小幅反弹外，整体空气质量在逐年改善。2015 年，研究区内除黄山区域和浙江南部山区 PM_{2.5} 浓度低于国家二级标准所规定的

35 μg/m³ 外，其余内陆地区都远超这一标准，空气质量不容乐观。2016 年 PM_{2.5} 浓度整体下降约 4 μg/m³，其中以杭嘉湖及其周边区域最为明显。2017 年年均 PM_{2.5} 浓度的小幅反弹主要由于安徽省空气质量变差导致了整个研究区平均浓度的上升（王学梅，2020）。随着 2018 年《打赢蓝天保卫战三年行动计划》的发布，长三角地区作为全国重点区域之一，大力推进产业绿色发展，优化能源结构，严格管控措施，2018 年—2020 年空气质量改善明显，2020 年长三角区域 PM_{2.5} 浓度年均值已低于国家二级标准，治理成效显著（http://www.mee.gov.cn/zcwj/gwywj/201807/t20180704_446068.shtml [2021-05-30]）。



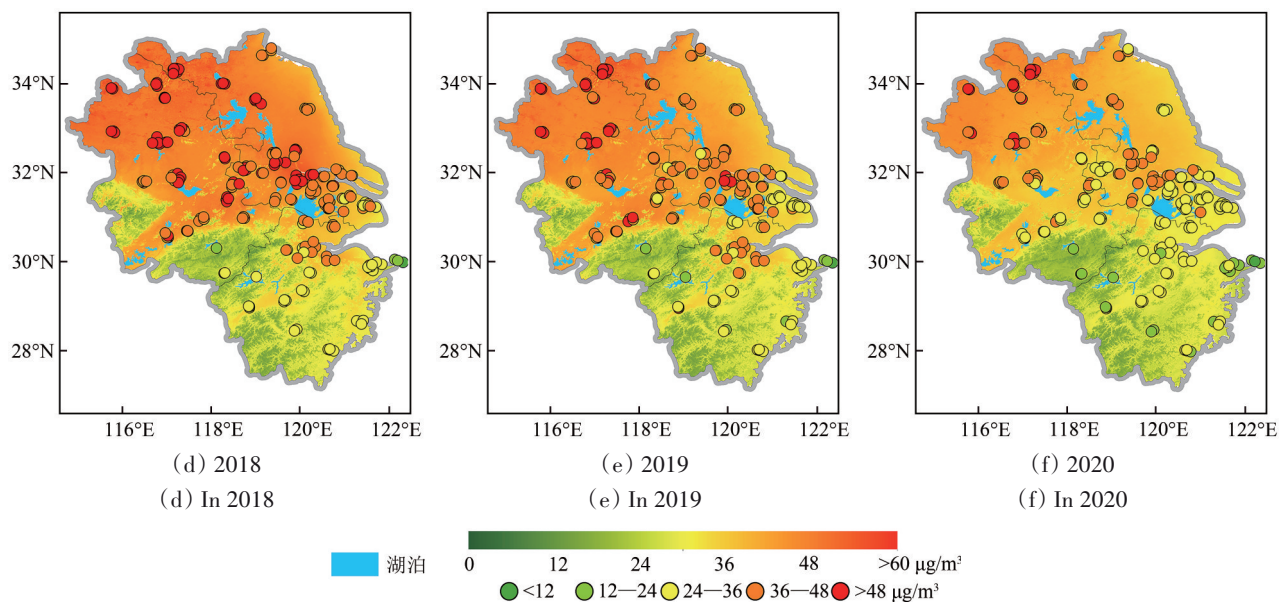


图5 长三角区域PM_{2.5}浓度年际分布

Fig.5 Annual PM_{2.5} concentration in the Yangtze River Delta.

4.3 季节PM_{2.5}浓度时空分布特征

为分析长三角区域PM_{2.5}浓度在不同季节的时空分布特征，本研究采用四季分类法将3、4、5月作为春季，6、7、8月作为夏季，9、10、11月作为秋季，12、1、2月作为冬季，对上述估算得到的2015年—2020年每日PM_{2.5}数据进行整合，得到四季的PM_{2.5}浓度均值（图6），与地面站点也同样具有较高的一致性。长三角区域季节PM_{2.5}浓度整体为冬季最高（58.35 µg/m³），春季次之（40.18 µg/m³），其后为秋季（36.32 µg/m³），夏季浓度最低（25.37 µg/m³）。地面站点季节趋势相同，

PM_{2.5}浓度均值分别为：冬季最高（68.58 µg/m³），春季次之（46.42 µg/m³），其后为秋季（41.44 µg/m³），夏季浓度最低（31.23 µg/m³）。冬季由于华北地区污染物传输、不利的气象条件和供暖因素，从而导致空气质量变差（马宗伟，2015；Xie等，2015）。长三角区域夏季对流天气强，雨水丰富，污染物沉降和扩散条件较好，因此夏季PM_{2.5}浓度在四季中最低（Kim等，2007）。空间分布上每个季节都类似，均为北高南低，安徽省和江苏省污染最为严重，上海市次之，浙江省空气质量相对最优，与上一节结果相同。

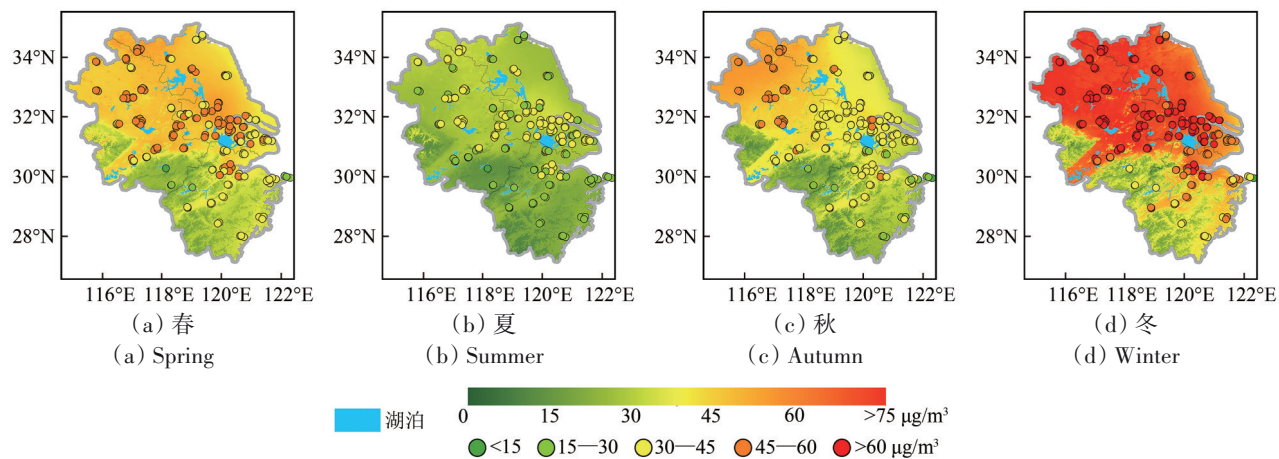


图6 长三角区域PM_{2.5}浓度季节分布

Fig.6 Seasonal PM_{2.5} concentration in the Yangtze River Delta

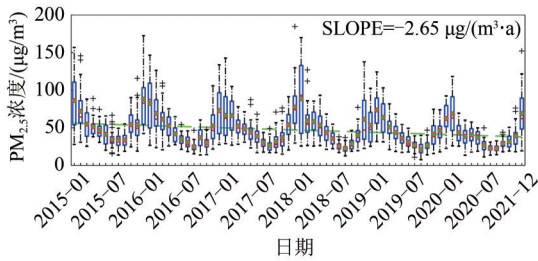
4.4 各省市PM_{2.5}浓度分析

为进一步分析长三角内部各省市的PM_{2.5}污染

情况，绘制各省市每月PM_{2.5}浓度箱线图并计算其变化趋势如图7。所有省市PM_{2.5}浓度总体趋势都在

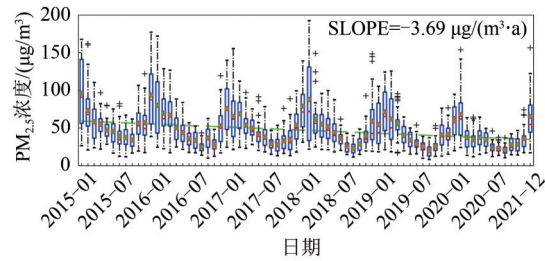
好转, 但估算结果的下降速率与站点结果有所差异, 同时估算结果的月际波动幅度较站点结果也更小。出现这种差异的主要原因是由于地面观测站点大都分布于PM_{2.5}浓度较高的城市区域, 其总体浓度较全域均值都偏高, 而本研究估算的PM_{2.5}浓度还包括山区等地, 反映的是全域整体情况。此外, PM_{2.5}浓度的月度变化趋势呈现显著的周期性, 并且各省市的趋势都相似: 冬季月份高, 夏季月份低, 春秋过渡。箱线图也直观的指示了除夏季的6、7、8月外, 各省市其余月份都有超过75%的日子PM_{2.5}日均浓度无法达到国家一级标准(日均35 μg/m³), 而冬季月份(12、1、2月)除浙江省外有超半数的天数超过国家二级标准(日均75 μg/m³)。本研究估算PM_{2.5}的结果显示, 四省市中以上海市下降趋势最为明显(下降速率=-3.30 μg/(m³·a))。上海市在前4 a出现PM_{2.5}浓度异常高值的情况最为频繁, 最高的几天接近200 μg/m³, 近2 a有所好转, 出现极端污染的情况明显减少, 极值也下降了近一半。地面站点趋势与本研究估算结果类似, 下降速率为4.06 μg/(m³·a)。地面站点

结果中下降速率最快的为浙江省, 达4.09 μg/(m³·a), 然而本研究结果显示浙江省下降趋势最为平稳, 下降速率仅为1.79 μg/(m³·a)。其原因为浙江省大部分为山区, 6年来山区PM_{2.5}浓度一直保持在较低水平, 而地面观测站点不包含这些区域, 导致其结果浓度偏高。同时站点结果与本研究估算结果的差异也说明了浙江省在城市地区PM_{2.5}的控制方面做到了较高水平, 值得其余地区学习借鉴。江苏省地处平原, 全域PM_{2.5}浓度空间异质性较低, 因此其估算结果与地面站点结果差异相对较小, 下降速率分别为2.65 μg/(m³·a)和3.69 μg/(m³·a)。江苏省在2018年以前月间PM_{2.5}浓度波动变化较大, 夏冬差距明显; 2018年以后冬季重污染天气减少, 空气质量改善, 月间变化变缓。安徽省由于黄山等山区的存在, 其总体下降速率的差异同浙江省类似, 估算结果为1.72 μg/(m³·a), 相对较低, 地面站点结果为3.37 μg/(m³·a), 同样为长三角区域最缓。但安徽省PM_{2.5}浓度较浙江省高, 月季变化幅度也较大, 因此仍有较大改善空间, 是后续更需要政府关注的重点区域。



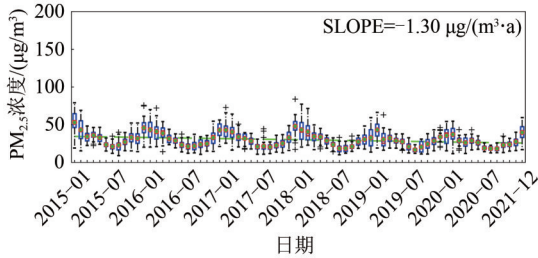
(a) 江苏省(估算的PM_{2.5})

(a) Jiangsu Province (estimated PM_{2.5})



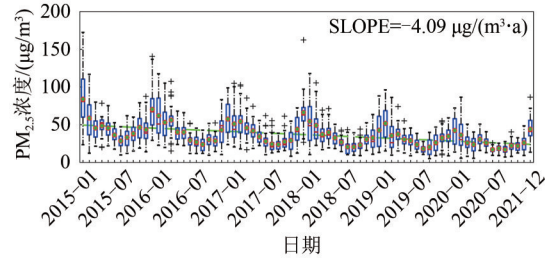
(b) 江苏省(站点监测的PM_{2.5})

(b) Jiangsu Province (in-situ PM_{2.5})



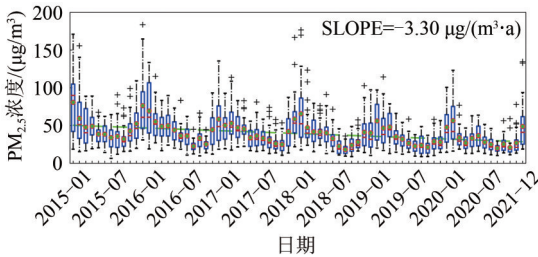
(c) 浙江省(估算的PM_{2.5})

(c) Zhejiang Province (estimated PM_{2.5})



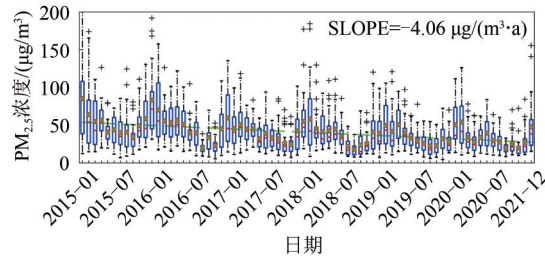
(d) 浙江省(站点监测的PM_{2.5})

(d) Zhejiang Province (in-situ PM_{2.5})



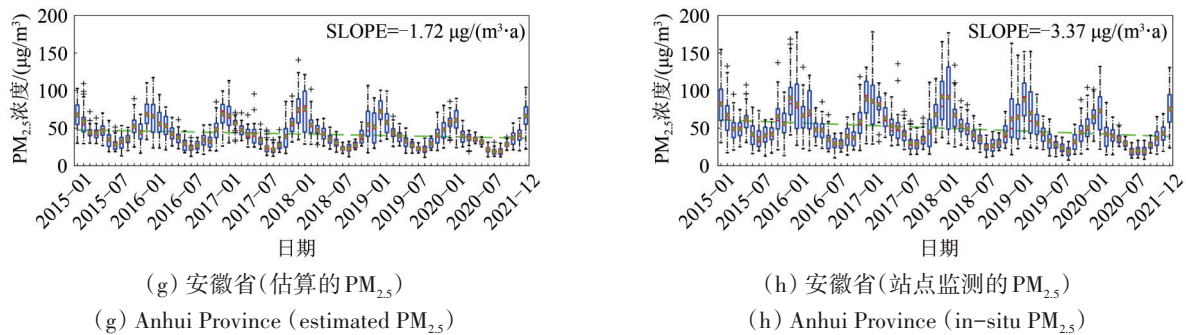
(e) 上海市(估算的PM_{2.5})

(e) Shanghai Province (estimated PM_{2.5})



(f) 上海市(站点监测的PM_{2.5})

(f) Shanghai Province (in-situ PM_{2.5})

图7 长三角区域各省市2015年—2020年PM_{2.5}浓度趋势Fig.7 Provincial trend of PM_{2.5} concentration from 2015 to 2020 in the Yangtze River Delta

5 结论

(1) 针对现有机器学习模型不能很好地捕获AOD-PM_{2.5}关系的时空异质性问题, 本研究构建了一种结合注意力机制的自适应深度神经网络模型(SADNN), 在拥有复杂非线性建模能力的同时, 对输入因子动态校正, 以识别其时空异质性。模型交叉验证结果指示所提出的模型相比多元线性回归和随机森林拥有更高的估算性能, R^2 达0.85, RMSRE为0.35, 能更准确地估算逐日PM_{2.5}浓度。

(2) 在输入因子的选取方面, 除了常规的AOD、气象因子、DEM、NDVI、DOY和经纬度外, 加入了针对于中国人民生产生活习惯的农历日因子。实验结果表明农历日因子的加入能够一定程度上校正AOD-PM_{2.5}的关系, 使得估算结果在空间分布上更为精确。

(3) 利用该模型进一步估算长三角区域2015年—2020年的PM_{2.5}浓度发现: PM_{2.5}浓度在空间分布上整体呈现北高南低的趋势, 安徽省与江苏省空气污染最为严重, 浙江省浓度最低。时间尺度上长三角区域2015年—2020年整体PM_{2.5}浓度除2017年外, 每年都在下降; 季节差异同样十分显著, 呈冬季>春季>秋季>夏季的规律。分省市分析结果表明: 上海市PM_{2.5}浓度下降最快, 下降速率为3.30 $\mu\text{g}/(\text{m}^3\cdot\text{a})$, 江苏省次之(2.65 $\mu\text{g}/(\text{m}^3\cdot\text{a})$), 安徽省与浙江省变化速度相近, 均不到2 $\mu\text{g}/(\text{m}^3\cdot\text{a})$, 但安徽省整体浓度远高于浙江省, 改善空间较大, 需要更多的关注。

(4) 利用本研究提出的模型能够较好地估算全域逐日PM_{2.5}浓度, 估算结果与地面观测有较好的一致性, 同时由于地面站点只分布于城市等PM_{2.5}浓度相对较高区域, 本研究结果可以弥补地

面观测在空间分布上的不足。然而现阶段模型的辅助因子以气象为主, 缺少具有时空特征的人为活动排放清单这类因子。未来期望再加入这类因子到模型中, 同时累积更多的地面观测数据进行模型的训练, 得到更为精确的PM_{2.5}浓度时空分布, 为流行病学研究和中国生态文明建设提供更为可靠的数据与理论支撑。

志 谢 本研究所用MODIS系列卫星数据由NASA提供, ERA5气象数据与AOD再分析数据由欧洲中期天气预报中心提供, PM_{2.5}数据由国家生态环境部提供, 地面AOD监测数据由AERONET提供, 在此表示衷心的感谢!

参考文献(References)

- Dockery D W, Pope C A, Xu X P, Spengler J D, Ware J H, Fay M E, Ferris B G Jr and Speizer F E. 1993. An association between air pollution and mortality in six U.S. cities. *New England Journal of Medicine*, 329(24): 1753-1759 [DOI: 10.1056/nejm19931209312093292401]
- Fang X, Zou B, Liu X P, Sternberg T and Zhai L. 2016. Satellite-based ground PM_{2.5} estimation using timely structure adaptive modeling. *Remote Sensing of Environment*, 186: 152-163 [DOI: 10.1016/j.rse.2016.08.027]
- Fu Y Y, Ye Z R, Deng J S, Zheng X Y, Huang Y B, Yang W, Wang Y H and Wang K. 2019. Finer resolution mapping of marine aquaculture areas using WorldView-2 imagery and a hierarchical cascade convolutional neural network. *Remote Sensing*, 11(14): 1678 [DOI: 10.3390/rs11141678]
- GBD 2019 Risk Factors Collaborators. 2020. Global burden of 87 risk factors in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. *The Lancet*, 396(10258): 1223-1249 [DOI: 10.1016/S0140-6736(20)30752-2]
- He Q Q and Huang B. 2018. Satellite-based high-resolution PM_{2.5} estimation over the Beijing-Tianjin-Hebei region of China using an improved geographically and temporally weighted regression

- model. *Environmental Pollution*, 236: 1027-1037 [DOI: 10.1016/j.envpol.2018.01.053]
- Hu X F, Belle J H, Meng X, Wildani A, Waller L A, Strickland M J and Liu Y. 2017. Estimating PM_{2.5} concentrations in the conterminous United States using the random forest approach. *Environmental Science and Technology*, 51(12): 6936-6944 [DOI: 10.1021/acs.est.7b01210]
- Just A C, Wright R O, Schwartz J, Coull B A, Baccarelli A A, Tellez-Rojo M M, Moody E, Wang Y J, Lyapustin A and Kloog I. 2015. Using high-resolution satellite aerosol optical depth to estimate daily PM_{2.5} geographical distribution in Mexico city. *Environmental Science and Technology*, 49(14): 8576-8584 [DOI: 10.1021/acs.est.5b00859]
- Kim S W, Yoon S C, Kim J and Kim S Y. 2007. Seasonal and monthly variations of columnar aerosol optical properties over East Asia determined from multi-year MODIS, LIDAR, and AERONET Sun/sky radiometer measurements. *Atmospheric Environment*, 41(8): 1634-1651 [DOI: 10.1016/j.atmosenv.2006.10.044]
- LeCun Y, Bengio Y and Hinton G. 2015. Deep learning. *Nature*, 521(7553): 436-444 [DOI: 10.1038/nature14539]
- Li S Z, Xing X F, Fan W Q, Cai B L, Fordson P and Xu X M. 2021. Spatiotemporal and frequential cascaded attention networks for speech emotion recognition. *Neurocomputing*, 448: 238-248 [DOI: 10.1016/j.neucom.2021.02.094]
- Li Z Q, Zhang Y, Shao J, Li B S, Hong J, Liu D, Li D H, Wei P, Li W, Li L, Zhang F X, Guo J, Deng Q, Wang B X, Cui C L, Zhang W C, Wang Z Z, Lv Y, Xu H, Chen X F, Li L and Qie L. 2016. Remote sensing of atmospheric particulate mass of dry PM_{2.5} near the ground: method validation using ground-based measurements. *Remote Sensing of Environment*, 173: 59-68 [DOI: 10.1016/j.rse.2015.11.019]
- Lin C Q, Li Y, Yuan Z B, Lau A K H, Li C C and Fung J C H. 2015. Using satellite remote sensing data to estimate the high-resolution distribution of ground-level PM_{2.5}. *Remote Sensing of Environment*, 156: 117-128 [DOI: 10.1016/j.rse.2014.09.015]
- Liu Y, Franklin M, Kahn R and Koutrakis P. 2007. Using aerosol optical thickness to predict ground-level PM_{2.5} concentrations in the St. Louis area: a comparison between MISR and MODIS. *Remote Sensing of Environment*, 107(1/2): 33-44 [DOI: 10.1016/j.rse.2006.05.022]
- Lyapustin A, Wang Y J, Korkin S and Huang D. 2018. MODIS collection 6 MAIAC algorithm. *Atmospheric Measurement Techniques*, 11(10): 5741-5765 [DOI: 10.5194/amt-11-5741-2018]
- Ma Z W. 2015. Study on Spatiotemporal Distributions of PM_{2.5} in China Using Satellite Remote Sensing. Nanjing: Nanjing University (马宗伟. 2015. 基于卫星遥感的我国PM_{2.5}时空分布研究. 南京: 南京大学)
- Mou L C and Zhu X X. 2020. Learning to pay attention on spectral domain: a spectral attention module-based convolutional network for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 58(1): 110-122 [DOI: 10.1109/TGRS.2019.2933609]
- Nair V and Hinton G E. 2010. Rectified linear units improve restricted boltzmann machines//*Proceedings of the 27th International Conference on International Conference on Machine Learning*. Haifa: Omnipress: 807-814
- Reddy R V K, Rao B S and Raju K P. 2018. Handwritten hindi digits recognition using convolutional neural network with RMSprop optimization//*2018 Second International Conference on Intelligent Computing and Control Systems (ICICCS)*. Madurai: IEEE: 45-51 [DOI: 10.1109/ICCONS.2018.8662969]
- Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez A N, Kaiser Ł and Polosukhin I. 2017. Attention is all you need//*Proceedings of the 31st International Conference on Neural Information Processing Systems*. Long Beach, California : Curran Associates Inc.: 6000-6010
- Wang X M. 2020. Analysis of Spatial and Temporal Variation of PM_{2.5} Meteorological Influencing Factors in Anhui Province. Hefei: Anhui Agricultural University (王学梅. 2020. 安徽省PM_{2.5}时空变化及其气象影响因素分析. 合肥: 安徽农业大学)
- Wei J, Li Z Q, Cribb M, Huang W, Xue W H, Sun L, Guo J P, Peng Y R, Li J, Lyapustin A, Liu L, Wu H and Song Y M. 2020. Improved 1 km resolution PM_{2.5} estimates across China using enhanced space - time extremely randomized trees. *Atmospheric Chemistry and Physics*, 20(6): 3273-3289 [DOI: 10.5194/acp-20-3273-2020]
- Xiao Q Y, Wang Y J, Chang H H, Meng X, Geng G N, Lyapustin A and Liu Y. 2017. Full-coverage high-resolution daily PM_{2.5} estimation using MAIAC AOD in the Yangtze River Delta of China. *Remote Sensing of Environment*, 199: 437-446 [DOI: 10.1016/j.rse.2017.07.023]
- Xie Y Y, Wang Y X, Zhang K, Dong W H, Lv B L and Bai Y Q. 2015. Daily estimation of ground-level PM_{2.5} concentrations over Beijing using 3 km resolution MODIS AOD. *Environmental Science and Technology*, 49(20): 12280-12288 [DOI: 10.1021/acs.est.5b01413]
- Yang L J, Xu H Q and Jin Z F. 2018. Estimation of ground-level PM_{2.5} concentrations using MODIS satellite data in Fuzhou, China. *Journal of Remote Sensing*, 22(1): 64-75 (杨立娟, 徐涵秋, 金致凡. 2018. MODIS 卫星遥感估计福州地区近地面PM_{2.5}浓度. 遥感学报, 22(1): 64-75) [DOI: 10.11834/jrs.20186501]
- Ye Z R, Fu Y Y, Gan M Y, Deng J S, Comber A and Wang K. 2019. Building extraction from very high resolution aerial imagery using joint attention deep neural network. *Remote Sensing*, 11(24): 2970 [DOI: 10.3390/rs11242970]
- Zang L, Mao F Y, Guo J P, Wang W, Pan Z X, Shen H F, Zhu B and Wang Z M. 2019. Estimation of spatiotemporal PM_{1.0} distributions in China by combining PM_{2.5} observations with satellite aerosol optical depth. *Science of the Total Environment*, 658: 1256-1264 [DOI: 10.1016/j.scitotenv.2018.12.297]
- Zhang Y and Li Z Q. 2013. Estimation of PM_{2.5} from fine-mode aerosol optical depth. *Journal of Remote Sensing*, 17(4): 929-943 (张莹, 李正强. 2013. 利用细模态气溶胶光学厚度估计PM_{2.5}. 遥感学报, 17(4): 929-943) [DOI: 10.11834/jrs.20133063]

Zhu J A. 2020. Research on Visual Object Tracking Algorithm based on Deep Learning. Changchun: University of Chinese Academy of Sciences (Changchun Institute of Optics, Fine Mechanics and

Physics, Chinese Academy of Sciences) (朱均安. 2020. 基于深度学习的视觉目标跟踪算法研究. 长春: 中国科学院大学(中国科学院长春光学精密机械与物理研究所))

Spatiotemporal estimation of PM_{2.5} using attention-based deep neural network

CHEN Binjie¹, YE Yang¹, LIN Yi², YOU Shixue¹, DENG Jinsong^{1,3,4}, YANG Wu^{1,4}, WANG Ke¹

1.College of Environment and Resource Sciences, Zhejiang University, Hangzhou 310058, China;

2.College of Environment, Zhejiang University of Technology, Hangzhou 310014, China;

3.Zhejiang Ecological Civilization Academy, Anji 313300, China;

4.Center for Intelligent Ecology and Sustainability, Zhejiang University, Hangzhou 310058, China

Abstract: PM_{2.5}, as the primary indicator of environmental quality, not only affects the occurrence of haze but also is closely related to public health and has raised great attention recently. Although PM_{2.5} ground monitoring stations are expanding, they are still on the sparse side to identify the spatiotemporal heterogeneity of PM_{2.5} concentrations. With the development of remote sensing technology, satellite-based Aerosol Optical Depth (AOD) data provide an effective way to estimate large-scale PM_{2.5} concentrations. This study aims to develop a novel deep neural network model for estimating PM_{2.5} concentrations in the Yangtze River Delta (YRD).

In addition to satellite remote sensing AOD data, meteorological factors, digital elevation model data, normalized different vegetation index data, and the lunar calendar day representing Chinese production and living habits were integrated into the proposed attention-based Self-Adaptive Deep Neural Network (SADNN) in this study to estimate PM_{2.5} concentrations in the YRD region from 2015 to 2020. Five-fold cross-validation was executed to evaluate the estimation accuracy of the SADNN. The multiple linear regression and random forest models were applied to compare with the SADNN.

The cross-validation results showed the proposed SADNN model had a high coefficient of determination value of 0.85 and a slope of 0.86, which were highly consistent with ground-level observations. The results also showed better performance than those of multiple linear regression and random forest models. The results for the spring festival in 2016 demonstrated the effectiveness of integrating the lunar calendar day and attention module into the model. The spatiotemporal patterns of PM_{2.5} in the YRD were as follows: PM_{2.5} concentrations were high in the north and low in the south, and the coastal and mountainous areas were better than inland and plain areas, respectively. On seasonal scales, winter was the most polluted season, while summer was the best. The overall PM_{2.5} concentration in the YRD showed a decreasing trend from 2015 to 2020, especially in Shanghai Municipality, with the decreasing speed of 3.30 μg/(m³·a), following the Jiangsu Province (2.65 μg/(m³·a)). Zhejiang Province and Anhui Province had a lower decreasing speed of less than 2 μg/(m³·a), and Anhui Province needed more effort and attention to improve the air quality due to its overall high PM_{2.5} concentrations.

In conclusion, applying satellite remote sensing data and the proposed SADNN model to accurately estimate spatially continuous PM_{2.5} concentrations can greatly make up for the lack of ground monitoring stations and scientific guide for environmental policy planning and implementation.

Key words: remote sensing, Aerosol Optical Depth (AOD), deep learning, attention module, the Yangtze River Delta, air quality

Supported by National Key Research and Development Program of China (No. 2020YFC1807500); Key Research and Development Program of Zhejiang Province (No. 2022C03078)