

基于 GRNN 神经网络模型结合气溶胶消光系数和气象要素评估颗粒物质量浓度

莫祖斯¹, 卜令兵^{1*}, 王勤¹, 林雪飞¹, Samuel A. Berhane¹, 杨彬², 邓晨², 李智²¹南京信息工程大学大气物理学院, 气象灾害预警与评估协同创新中心, 江苏 南京 210044;²南京牧镭激光科技有限公司, 江苏 南京 210038

摘要 大气颗粒物是最重要的空气污染物之一, 会对人类健康产生负面影响。激光雷达探测是实现颗粒物分布高精度测量的可行手段。气溶胶消光系数在一定程度上能反映气溶胶质量浓度的相对大小, 气象要素对消光系数和质量浓度的影响不容忽视。本团队利用反演得到的消光系数, 结合地面温度、相对湿度、风速、地面气压等地面气象要素, 与 PM_{2.5}、PM₁₀ 质量浓度建立数据集; 通过主成分分析法计算数据特征, 基于广义回归神经网络(GRNN)对 PM_{2.5}、PM₁₀ 质量浓度建立评估模型。GRNN 模型得到的 PM_{2.5} 和 PM₁₀ 质量浓度的评估值与真实值的相关系数分别为 0.86 和 0.85, 均方根误差(RMSE)分别为 2.58 μg/m³ 和 10.84 μg/m³, 平均绝对误差(MAE)分别为 0.81 μg/m³ 和 1.53 μg/m³。将 GRNN 模型应用于激光雷达扫描模式下, 对南京市浦口区颗粒物质量浓度的水平分布进行了评估, 评估值和实际站点测量值的一致性较好, 进一步验证了 GRNN 模型用于颗粒物质量浓度评估的有效性。

关键词 遥感; 气溶胶消光系数; 颗粒物质量浓度; 神经网络; 气象要素

中图分类号 X831; P412.25; P407.5

文献标志码 A

DOI: 10.3788/CJL202249.1710001

1 引言

气溶胶是悬浮在大气中的固液混合颗粒物, 尺寸一般在 0.001~100 μm 之间, 是影响城市空气质量的重要指标^[1-3]。一方面, 当空气中的气溶胶浓度过大时, 气溶胶颗粒可以进入人体, 甚至进入肺部或肺泡, 其携带的有毒物质会危害人类健康^[4-5]; 另一方面, 大气气溶胶污染是造成雾霾的根本原因^[6-8], 大气能见度的降低会导致交通事故的发生率增加。此外, 气溶胶污染对大气遥感定量分析也有着巨大影响^[9-13]。

直接测量 PM_{2.5}、PM₁₀ 质量浓度的仪器是颗粒物检测器, 检测方法主要有微动平衡法和 β 射线吸收法^[14], 这些方法均基于取样观测原理进行检测。徐敬等^[15]将颗粒物监测仪放置在楼顶, 对北京 PM_{2.5} 的质量浓度和化学成分特征进行了统计分析; Wu 等^[16]将定点采样仪分别放置在 325 m 和 250 m 高的气象塔上, 研究了 PM_{2.5} 质量浓度的垂直分布。这些方法在区域大尺度场景数据的获取上有所欠缺。Peng 等^[17]将颗粒物监测仪搭载在无人机上, 对颗粒物的垂直分布进行了观测, 但该方法的探测高度和范围有限。

激光雷达是解决区域大尺度场景气溶胶观测的有效手段^[18-19]。张辉等^[20]利用激光雷达和 PM_{2.5} 监测仪, 对气溶胶后向散射系数、颗粒物质量浓度和相对湿度数据进行拟合, 证明了相对湿度是影响颗粒物浓度反演精度的重要因素; 陶宗明等^[21]利用侧向散射激光雷达、PM_{2.5} 监测仪, 对合肥地区的 PM_{2.5} 质量浓度垂直分布廓线进行了研究。

上述研究主要考虑利用激光雷达消光系数的变化信息, 或者进一步结合温度、相对湿度等误差影响因素, 通过反演来获取 PM_{2.5} 和 PM₁₀ 的质量浓度。需要指出的是, 韩道文等^[22]认为除了温度、相对湿度以外, 气压和风场等气象要素同样对颗粒物质量浓度分布有显著影响。要想提高基于激光雷达测量气溶胶质量浓度的精度, 需要同时综合气象参数等影响颗粒物质量浓度的各类因素, 开展特定预测模型及其算法的研究。白盛楠等^[23]搭建了多变量的 LSTM 循环神经网络 PM_{2.5} 预测模型, 并用该模型实现了 PM_{2.5} 日值质量浓度的准确预测; 章雅楠等^[24]综合气象要素和颗粒物质量浓度, 建立了基于小波神经网络的 PM_{2.5} 质量浓度预测模型, 该模型弥补了传统观测数据在

收稿日期: 2021-11-15; 修回日期: 2021-12-18; 录用日期: 2022-01-14

基金项目: 国家自然科学基金(42175145)

通信作者: *lingbingbu@nuist.edu.cn

预测能力上的不足。以上研究证明了将神经网络用于评价颗粒物质量浓度是可行的。因此,如果能将激光雷达反演的消光系数与影响颗粒物质量浓度分布的气象因素结合起来,并基于神经网络建立一个质量浓度空间分布的评价模型,就可以发挥激光雷达在 PM_{2.5}、PM₁₀ 质量浓度监测方面覆盖范围广、时空分辨率高的优势。

本文基于广义回归神经网络结构 (GRNN),利用激光雷达得到的近地表消光系数,在综合考虑风速、地面温度、相对湿度、地面气压等气象因素对颗粒物质量浓度影响的前提下,结合近地表 PM_{2.5}、PM₁₀ 质量浓度数据,建立 GRNN 评估模型,并输出气溶胶质量浓度空间分布的模型评估结果;最后将该模型应用于激光雷达水平扫描监测颗粒物质量浓度的场景,通过将模型的评估结果与真实站点的测量结果进行对比,验证了所建模型的有效性。

2 数据和方法

2.1 数据

2.1.1 气溶胶消光系数

实验中采用的雷达是米散射体制的激光雷达,其系统架构如图 1 所示。基于激光雷达方程,采用一种改进的消光系数反演斜率算法,获得大气气溶胶消光系数^[25-27]。实验地点是中国江苏省南京市栖霞区 (32.09°N, 118.91°E),激光雷达放置在 30 m 的高楼上,水平发射 532 nm 的激光,激光传输路径经过南京市国家环境空气质量监测点仙林大学城站(以下简称“环境监测站”),激光雷达与该环境监测站相距 0.84 km,这一距离大于激光雷达收发视场不完全重叠区长度。取激光雷达反演产品的消光系数,并将其与环境监测站测得的 PM_{2.5} 和 PM₁₀ 的小时质量浓度进行时空匹配。

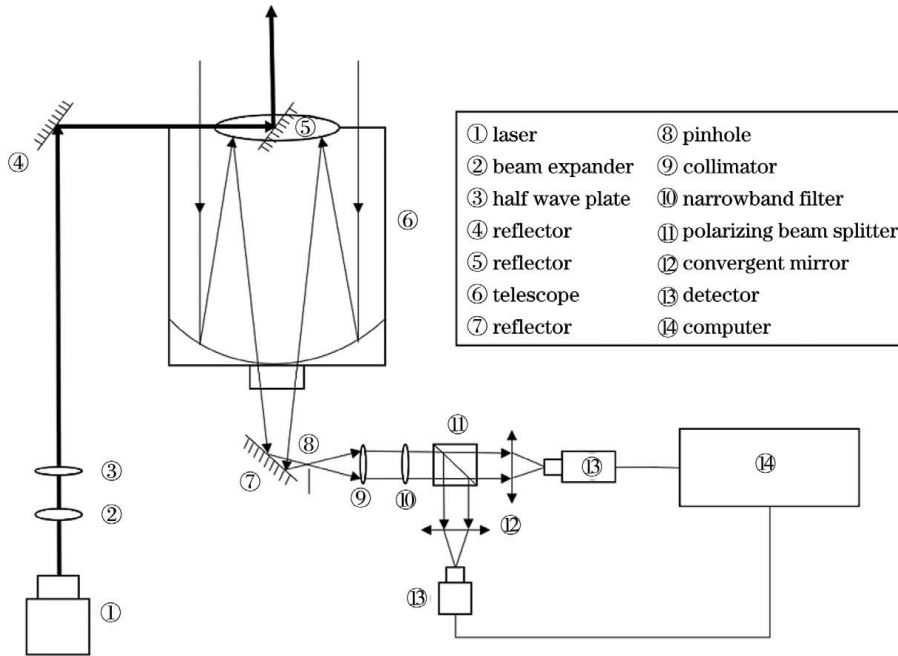


图 1 激光雷达系统架构

Fig. 1 A systematic framework of lidar

利用 Collis 斜率法结合 Fernald 后向积分法反演水平方向上的消光系数。假定水平方向上每一小段的大气是均匀的,然后利用 Collis 法反演各个小段的水平消光系数^[28]。反演公式为

$$\alpha_1(Z_c) = -\frac{1}{2} \frac{dS(Z)}{dZ}, \quad (1)$$

式中: $\alpha_1(Z_c)$ 为参考点 Z_c 处气溶胶的消光系数; $S(Z)$ 表示对距离修正回波信号取自然对数后的结果, $S(Z) =$

$\ln[X(Z)]$; $X(Z)$ 为距离修正回波信号; Z 为距离。

假定水平状态下气溶胶的消光散射比 $S_1 = \alpha_1(Z) / \beta_1(Z)$ 为常数 50,其中 $\beta_1(Z)$ 为气溶胶的后向散射系数。假定水平状态下大气分子的消光系数是常数,消光散射比 $S_2 = \alpha_2(Z) / \beta_2(Z)$ 为常数 $8\pi/3$,其中 $\alpha_2(Z)$ 为大气分子的消光系数,可由标准大气模式得到。

结合 Fernald 后向反演公式可以得到气溶胶消光系数随距离的变化函数为

$$\alpha_1(Z) = -\frac{3S_1}{8\pi} \alpha_2 + \frac{X(Z) \exp \left[2 \left(\frac{3S_1}{8\pi} - 1 \right) \alpha_2 (Z_c - Z) \right]}{\frac{X(Z_c)}{3S_1} + 2 \int_z^{Z_c} X(Z) \exp \left[2 \left(\frac{3S_1}{8\pi} - 1 \right) \alpha_2 (Z_c - Z) \right] dZ}, \quad (2)$$

式中： $\alpha_1(Z)$ 为气溶胶的消光系数； S_1 为气溶胶的消光散射比； $\alpha_2(Z_c)$ 为参考点 Z_c 处大气分子的消光系数。

2.1.2 地面要素资料

PM₁₀ 和 PM_{2.5} 小时质量浓度数据来自中华人民共和国生态环境部官网 (<http://www.mee.gov.cn/>)^[29]；地面温度、相对湿度、风速、地面气压和能见度等气象要素数据集来自国家卫星气象中心 (<http://data.cma.cn/>)。

2.2 方法

如图 2 所示，基于消光系数和地面气象要素，利用神经网络对颗粒物的质量浓度进行评估。评估过程主要包括 5 个步骤：1) 利用 Collis 斜率法结合 Fernald 法反演得到消光系数，搜集小时地面温度、相对湿度、风速、地面气压等基本要素，以及环境监测站 PM_{2.5}、PM₁₀ 的小时质量浓度，整理对应样本；2) 对样本数据进行预处理，包括异常值处理、剔除雨天天气数据、缺失值剔除，最终得到自变量为消光系数、地面温度、相对湿度、风速和地面气压，因变量为 PM_{2.5}、PM₁₀ 质量浓度的小时均值；3) 将各个要素数据进行标准化处理，使每个特征的分布基本趋于一致（这样每个特征的梯度就会下降得很快，同时能加快模型收敛），并将标准

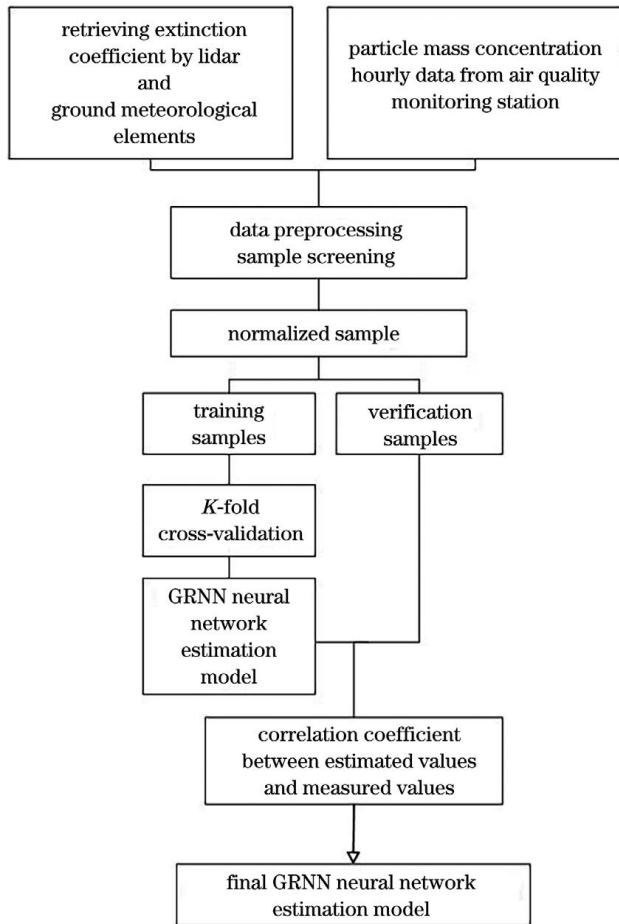


图 2 基于消光系数、气象要素，利用 GRNN 评估 PM_{2.5} 和 PM₁₀ 质量浓度的流程图

Fig. 2 Flow chart for estimating PM_{2.5} and PM₁₀ mass concentrations with GRNN based on extinction coefficient and meteorological elements

化后的数据随机分成训练样本和验证样本；4) 对训练样本进行 K 折交叉验证处理^[30]，将样本分割成 4 个子集，其中 1 个子集被保留作为验证模型的数据，其他 3 个子集用来训练（重复用随机产生的子样本进行训练和验证，对每一次训练结果进行误差验证）；5) 依据相关系数不断训练，输出较优的 GRNN 评估模型。

如图 3 所示，GRNN 神经网络由 4 层构成^[31]，依次为输入层、模式层、求和层和输出层。假设样本特征集为 $\{T_1, T_2, \dots, T_m\}$ ，单个样本的维度为 n ，即 $T_i = [x_1, x_2, \dots, x_n]$ ；同时假设标签集为 $\{B_1, B_2, \dots, B_m\}$ ，单个标签的维度为 k ，即 $B_i = [y_1, y_2, \dots, y_k]$ 。

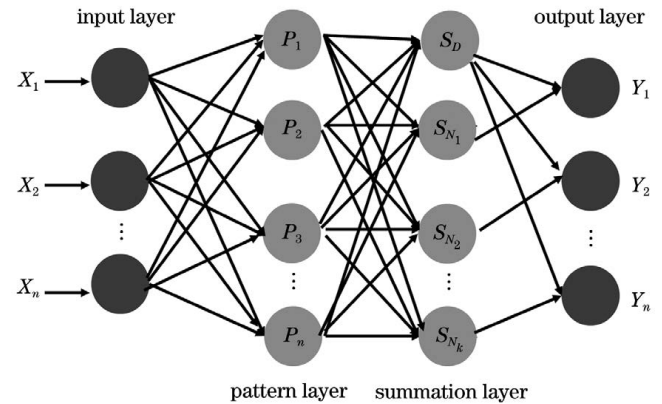


图 3 GRNN 架构

Fig. 3 GRNN structure

模式层中的节点个数等于训练样本的个数，第 i 个测试样本 T_i 与第 j 个测试样本 T_j 之间的高斯函数取值为

$$\text{Gauss}(T_i - T_j) = \exp\left(-\frac{\|T_i - T_j\|^2}{2\delta^2}\right), \quad (3)$$

式中： δ 为高斯函数的宽度系数，即光滑因子，该参数取值的大小会影响神经网络对样本的逼近程度，光滑因子越小，逼近性越好，光滑因子越大，样本误差越大。

求和层节点个数等于输出样本维度加 $(k + 1)$ 。求和层的输出分为两部分，第一个节点的输出为模式层输出的算术和，其余 k 个节点的输出为模式输出的加权和。假设对于测试样本，模式层的输出为 $\{g_1, g_2, \dots, g_m\}$ ，则求和层第一个节点的输出为

$$S_D = \sum_{i=1}^m g_i, \quad (4)$$

其余 k 个节点的输出为

$$S_{N_j} = \sum_{i=1}^m y_{ij} g_i, \quad j = 1, 2, \dots, k, \quad (5)$$

加权系数 y_{ij} 为第 i 个模式层节点对应的标签集中的第 j 个元素。输出层节点个数等于标签向量的维度，每个节点的输出等于对应的求和层输出与求和层第一个节点输出相除。

2.3 模型性能评价

为了评价所建立的 GRNN 神经网络模型的性能，对验证集的评价值和实际观测值进行比对分析。具体

的评价指标包括均方根误差 (RMSE)、相关系数 (R)、平均绝对误差 (MAE), 它们的计算公式分别为

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (V_a - V_b)^2}, \quad (6)$$

$$R = \frac{\frac{1}{n} \sum_{i=1}^n (V_a - V'_a)(V_b - V'_b)}{\sqrt{\sum_{i=1}^n (V_a - V'_a)^2} \sqrt{\sum_{i=1}^n (V_b - V'_b)^2}}, \quad (7)$$

$$MAE = \frac{1}{m} \sum_{i=1}^m |(V_a - V_b)|, \quad (8)$$

式中: n 为验证集中样本的个数; V_a 表示目标输出值, 即预测值; V'_a 表示目标输出平均值; V_b 表示实际输出值; V'_b 表示实际输出平均值。

3 结果与分析

3.1 基于 GRNN 评估 $PM_{2.5}$ 和 PM_{10} 质量浓度的模型验证

3.1.1 数据选取及相关性分析

选取 2021 年 4—8 月的消光系数、地面温度、相对湿度、风速、地面气压和颗粒物质量浓度观测数据作为模型训练样本。对数据质量进行严格控制, 具体控制方法为

- 1) 剔除阴雨天气消光系数数据样本;
- 2) 对消光系数进行小时平均, 并环境监测站提供的消光系数进行时间匹配, 采用 3 σ 原则进行异常值剔除;

3) 地面要素以及 $PM_{2.5}$ 和 PM_{10} 质量浓度也采用 3 σ 原则进行异常值剔除。

最终得到了 1451 个有效样本。表 1 为 $PM_{2.5}$ 和 PM_{10} 质量浓度与其他气象要素的相关系数统计, 其中, * 表示有统计学差异 ($P < 0.05$), ** 表示差异显著 ($P < 0.01$), *** 表示差异极显著 ($P < 0.001$)^[32], P 值表示原假设为真时, 样本观察结果或更极端结果出现的概率。通常情况下, 显著性结果达到 0.05 水平或 0.01 水平, 才可以说数据之间具备差异显著或是极显著。

PM_{10} 和 $PM_{2.5}$ 的质量浓度与消光系数呈正相关, 相关系数分别达到了 0.6339 和 0.5102, 其 P 值均小于 0.001。 PM_{10} 和 $PM_{2.5}$ 的质量浓度与地面温度呈负相关, 这是因为在高温下气溶胶粒子混合均匀, 其扩散能力较强, 边界层内的气溶胶浓度不会很大。相对湿度的增大会导致吸湿性和可溶性气溶胶颗粒的尺寸增加。近地表风场会影响污染物的传输和扩散, 风速的增大将加速大气中污染物的稀释和扩散。地面气压对颗粒物质量浓度的影响不明显, 这可能是因为地面气压本身在稳定的天气下变化不大^[22]。

表 1 自相关性分析 (表中 EX 为气溶胶的消光系数, TE 为地面温度, RH 表示相对湿度, WS 表示风速, PS 表示地面气压)

Table 1 Autocorrelation analysis (EX represents aerosol extinction coefficient, TE represents temperature, RH represents relative humidity, WS represents wind speed, and PS represents surface pressure)

| Parameter | Mass concentration of $PM_{2.5}$ | Mass concentration of PM_{10} | EX | TE | RH | WS | PS |
|----------------------------------|----------------------------------|---------------------------------|-----------|------------|------------|------------|----------|
| Mass concentration of $PM_{2.5}$ | 1 | 0.7394*** | 0.6339*** | -0.1877*** | -0.2894*** | -0.2067*** | 0.1146** |
| Mass concentration of PM_{10} | | 1 | 0.5102*** | -0.1372*** | 0.1005*** | -0.3139*** | 0.0838** |
| EX | | | 1 | -0.1817*** | 0.4662*** | -0.2241*** | 0.0257** |
| TE | | | | 1 | -0.3212*** | 0.2272** | -0.0177* |
| RH | | | | | 1 | -0.3284* | -0.1311* |
| WS | | | | | | 1 | -0.0717 |
| PS | | | | | | | 1 |

因此, 本文将激光雷达消光系数、地面温度、相对湿度、风速和地面气压作为自变量, $PM_{2.5}$ 和 PM_{10} 的质量浓度作为因变量, 建立 GRNN 模型。

3.1.2 GRNN 模型的验证

随机取样并按比例分配训练样本和验证样本。为方便对样本进行四折交叉验证, 最终取 1200 个样本作为训练集, 251 个样本作为验证集, 每次训练时都随机取样。在建立 GRNN 模型过程中, 为取到一个最优的 Spread 参数, 即 Spread 值在 0~2 之间, 通过四折交叉验证下均方差 (MSE) 指标的最小值来确定最优的 Spread 参数。四折交叉验证结果如图 4 所示, 可以看

出, 最优 Spread 值为 0.08。

图 5 是 GRNN 模型的验证结果, 其中图 5(a) 是 $PM_{2.5}$ 和 PM_{10} 质量浓度评估值与实际值的比较, 图 5(b) 是 $PM_{2.5}$ 质量浓度评估值与实际值的相关性, 图 5(c) 是 PM_{10} 质量浓度评估值与实际值的相关性。结合表 2 可以看出, GRNN 模型对 $PM_{2.5}$ 、 PM_{10} 质量浓度的评估与实际值比较符合, $PM_{2.5}$ 和 PM_{10} 质量浓度评估值与实际值的相关系数 (R) 分别为 0.86 和 0.85, 均方根误差 (RMSE) 分别为 2.58 $\mu\text{g}/\text{m}^3$ 和 10.84 $\mu\text{g}/\text{m}^3$, 平均绝对误差 (MAE) 分别为 0.81 $\mu\text{g}/\text{m}^3$ 和 1.53 $\mu\text{g}/\text{m}^3$ 。

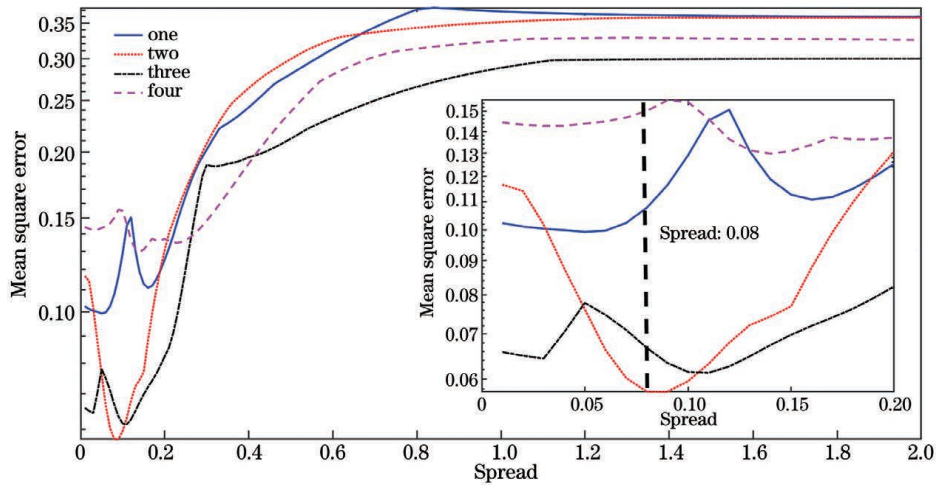


图 4 四折交叉验证结果

Fig. 4 Result of four fold cross validation

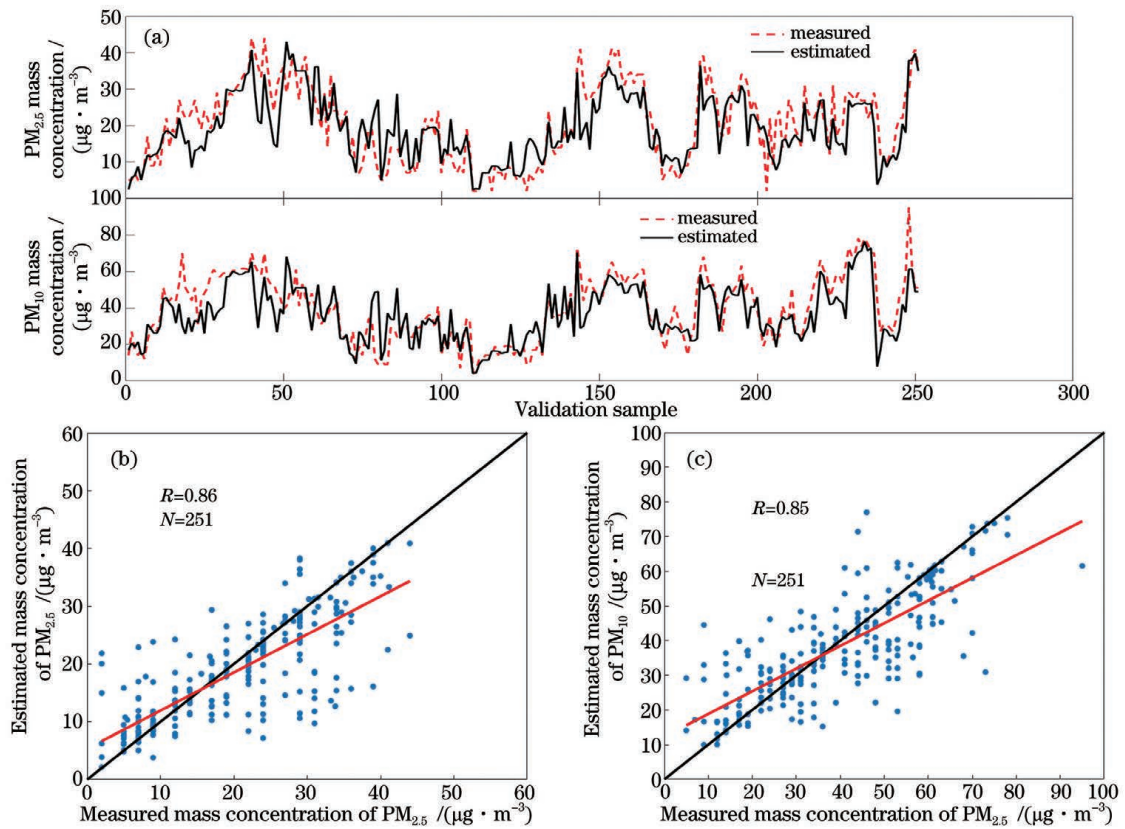


图 5 GRNN 模型验证结果。(a)PM_{2.5} 和 PM₁₀ 质量浓度评估值与实际值的比较;(b)PM_{2.5} 评估值与实际值的相关性;
(c)PM₁₀ 评估值与实际值的相关性

Fig. 5 Validation of GRNN model. (a) Comparison between estimated and measured mass concentrations of PM_{2.5} and PM₁₀ ;
(b) correlation between estimated and measured mass concentration of PM_{2.5} ; (c) correlation between estimated and measured mass concentration of PM₁₀

表 2 GRNN 模型的性能指标

Table 2 Performance indicators of GRNN model

| Particulate matter | Quantity of training sample | Quantity of validation sample | R | MAE / ($\mu\text{g} \cdot \text{m}^{-3}$) | RMSE / ($\mu\text{g} \cdot \text{m}^{-3}$) |
|--------------------|-----------------------------|-------------------------------|------|---|--|
| PM ₁₀ | 1200 | 251 | 0.86 | 1.53 | 10.84 |
| PM _{2.5} | 1200 | 251 | 0.85 | 0.81 | 2.58 |

3.2 GRNN 评价 $PM_{2.5}$ 和 PM_{10} 质量浓度的应用

如图 6 所示,将激光雷达 A 置于一建筑物上,进行连续水平扫描观测,其放置高度为 30 m 左右,仰角为 0.5° 。因受到高大建筑物的遮挡,其扫描角度范围

限制在西南 $215^\circ \sim$ 东南 110° ,有效探测半径为 3 km (以 A 为圆心),扫描周期为 20 min。基于该激光雷达反演的消光系数,结合地面气象要素,采用 GRNN 模型对 $PM_{2.5}$ 和 PM_{10} 的质量浓度进行评估。

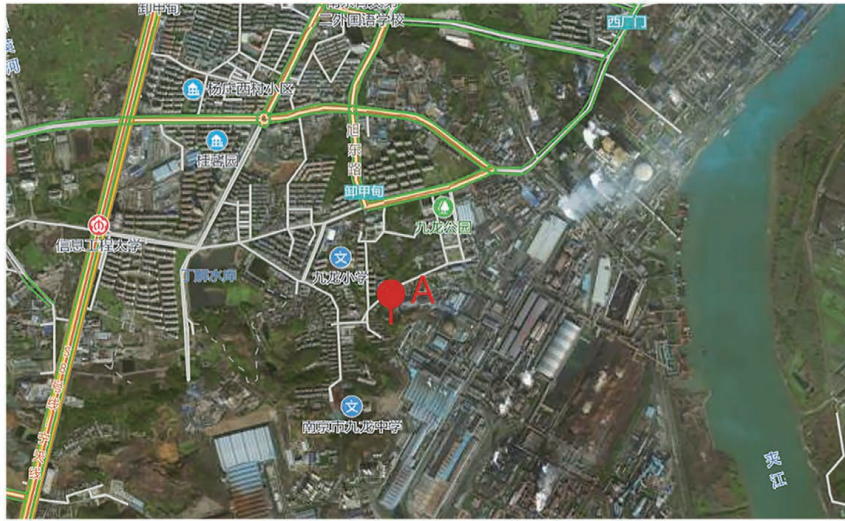


图 6 激光雷达观测点位分布图(点 A 是用于水平扫描的激光雷达)

Fig. 6 Lidar observation location distribution map (dot A is the locations of polarimetric lidar for horizontal scanning)

激光雷达 A 进行了连续一天的扫描,获得了该区域内消光系数的水平分布。本团队基于 GRNN 神经网络训练模型,结合地面温度、相对湿度、风速、地面气压等气象要素进行了 $PM_{2.5}$ 和 PM_{10} 质量浓度的评估。

PM_{10} 质量浓度的评估结果,结合图 9 所示的环境监测站提供的浦口区 PM_{10} 和 $PM_{2.5}$ 的质量浓度以及地面基本气象要素变化可以看出:当天受到主导风西南风、西风的影响,工业生产区内颗粒物有向东北方向扩散的趋势;5月9日 00:00—01:00 时,浦口区附近风力

图 7 和图 8 分别为激光雷达扫描区内 $PM_{2.5}$ 和

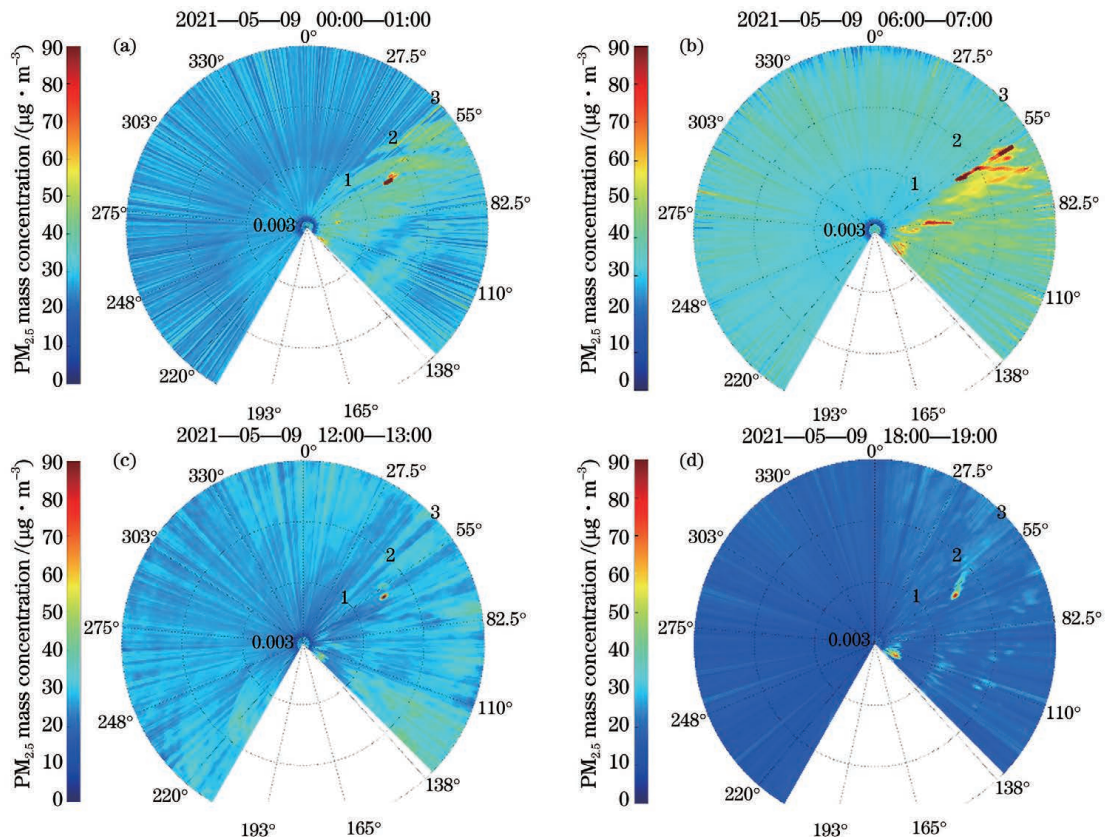


图 7 2021 年 5 月 9 日不同时间段水平扫描下 $PM_{2.5}$ 质量浓度评估结果的平均值

Fig. 7 Average $PM_{2.5}$ mass concentration estimation under horizontal scanning in different periods on May 9, 2021

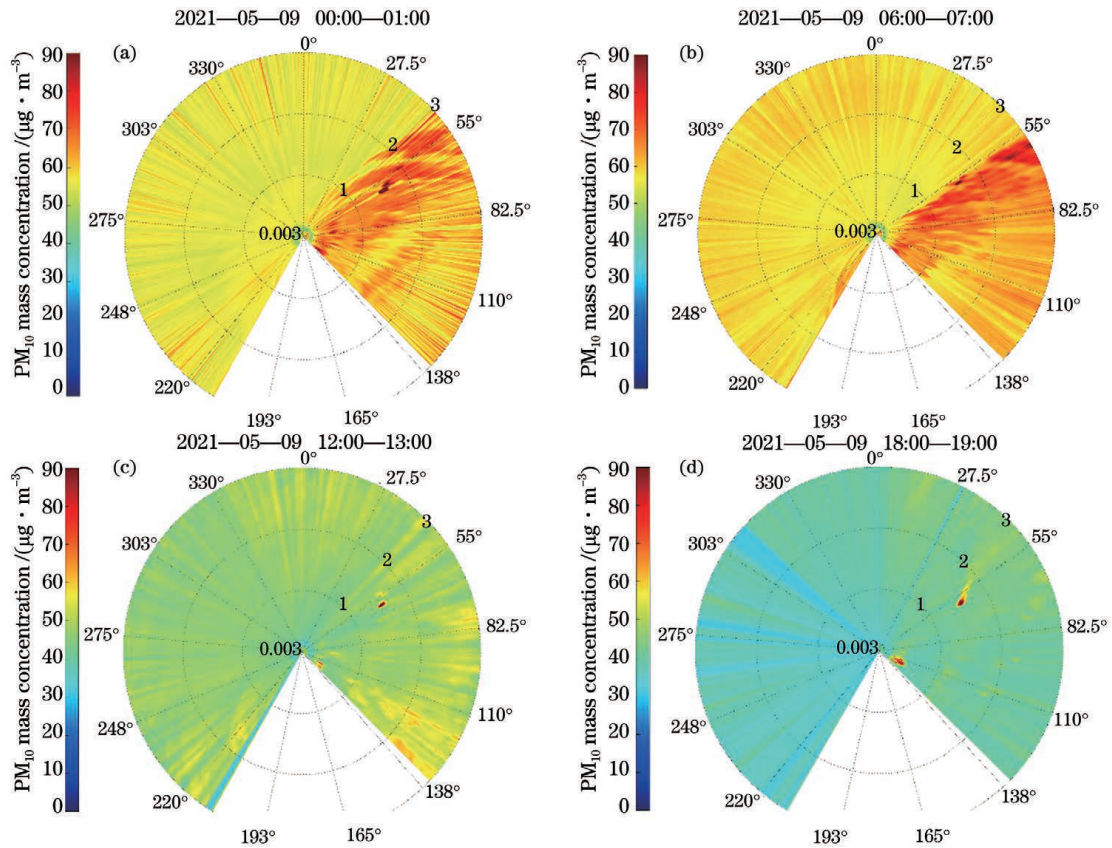


图 8 2021 年 5 月 9 日不同时间段水平扫描下 PM₁₀ 质量浓度评估结果的平均值

Fig. 8 Average PM₁₀ mass concentration estimation under horizontal scanning in different periods on May 9, 2021

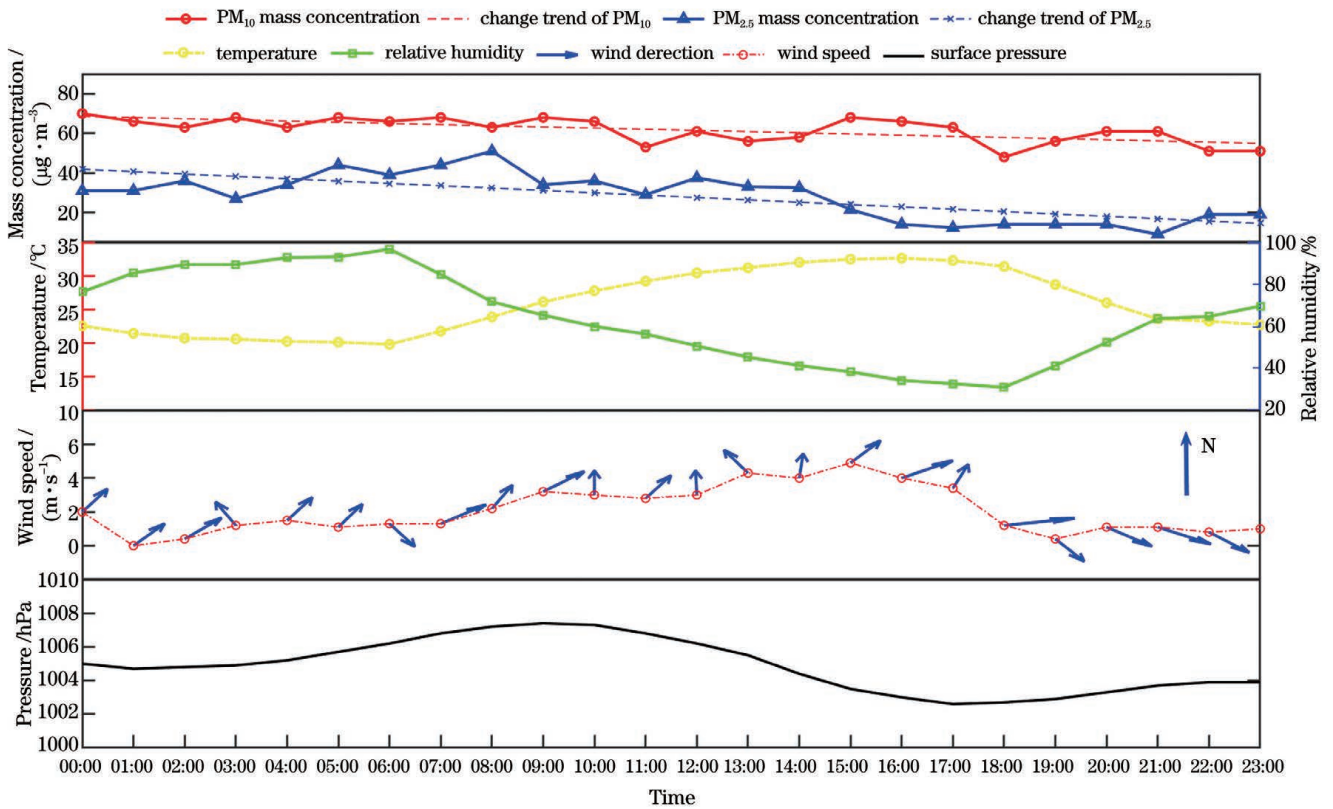


图 9 2021 年 5 月 9 日浦口区 PM_{2.5} 质量浓度、PM₁₀ 质量浓度、地面温度、相对湿度、风向风速、地面气压的小时平均值变化
Fig. 9 Change in hourly averages of PM_{2.5} mass concentration, PM₁₀ mass concentration, temperature, relative humidity, wind direction and speed, and surface pressure in Pukou District on May 9, 2021

约为 0~2 m/s,受西南风影响,激光雷达东侧和东北侧的颗粒物质量浓度较高(激光雷达 A 东侧为工业生产区,该区域内散布着公路、炼钢厂、水泥厂等),大部分区域 $PM_{2.5}$ 和 PM_{10} 的质量浓度分别超过了 $50 \mu\text{g}/\text{m}^3$ 和 $80 \mu\text{g}/\text{m}^3$,而激光雷达西侧的颗粒物质量浓度较低, $PM_{2.5}$ 和 PM_{10} 的质量浓度分别约为 $30 \mu\text{g}/\text{m}^3$ 和 $60 \mu\text{g}/\text{m}^3$;06:00—07:00 时,激光雷达 A 东侧、东北侧颗粒物质量浓度比 00:00—01:00 时增加明显,大部分区域 $PM_{2.5}$ 和 PM_{10} 的质量浓度分别超过 $60 \mu\text{g}/\text{m}^3$ 和 $90 \mu\text{g}/\text{m}^3$,而西侧 $PM_{2.5}$ 和 PM_{10} 的质

量浓度分别约为 $50 \mu\text{g}/\text{m}^3$ 和 $70 \mu\text{g}/\text{m}^3$;12:00—13:00 时,随着气温升高和风速加强,扫描区域内整体颗粒物质量浓度降低, $PM_{2.5}$ 和 PM_{10} 的质量浓度分别约为 $30 \mu\text{g}/\text{m}^3$ 和 $60 \mu\text{g}/\text{m}^3$;18:00—19:00 时,扫描区域内整体 $PM_{2.5}$ 和 PM_{10} 的质量浓度分别约为 $20 \mu\text{g}/\text{m}^3$ 和 $50 \mu\text{g}/\text{m}^3$ 。

GRNN 模型得到的 $PM_{2.5}$ 和 PM_{10} 颗粒物质量浓度评估值与环境监测站提供的颗粒物质量浓度小时均值的对比如图 10 所示,可以看出,模型值与站点实测值的变化趋势较为一致。

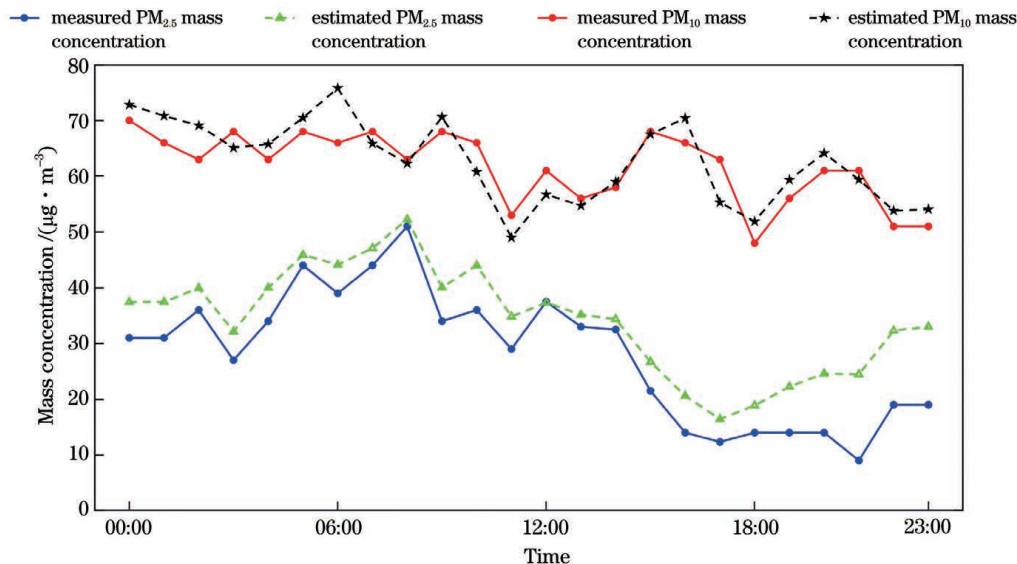


图 10 基于 GRNN 模型的 $PM_{2.5}$ 和 PM_{10} 评估值与实际测量值的对比

Fig. 10 Comparison between GRNN model estimated mass concentrations of $PM_{2.5}$ or PM_{10} and measured values

4 结 论

本文基于 GRNN 神经网络模型,利用一台激光雷达观测反演得到的气溶胶消光系数,结合地面温度、相对湿度、风速、地面气压数据集以及地表 $PM_{2.5}$ 、 PM_{10} 质量浓度小时均值建立相关模型。采用主成分分析筛选模型变量,对训练数据进行 K 折交叉验证,以均方差指标最小为依据选取最优滑动因子,并对输出模型进行了验证分析。验证结果显示: $PM_{2.5}$ 和 PM_{10} 质量浓度评价值与实际值的相关系数分别达到 0.86 和 0.85,均方根误差分别为 $2.58 \mu\text{g}/\text{m}^3$ 和 $10.84 \mu\text{g}/\text{m}^3$,平均绝对误差分别为 $0.81 \mu\text{g}/\text{m}^3$ 和 $1.53 \mu\text{g}/\text{m}^3$ 。利用激光雷达对南京市浦口区进行水平扫描观测,结合本地地面温度、相对湿度、风速和地面气压以及消光系数,用 GRNN 模型评估 $PM_{2.5}$ 和 PM_{10} 的质量浓度,并将评估结果与环境监测站提供的颗粒物质量浓度进行对比,两者的变化趋势比较一致,进一步验证了 GRNN 模型评估结果的准确性。GRNN 模型应用于激光雷达监测颗粒物质量浓度具有可行性,为颗粒物的大范围、高时空分辨率监测提供了新途径。

参 考 文 献

- [1] Pui D Y H, Chen S C, Zuo Z L. $PM_{2.5}$ in China: measurements, sources, visibility and health effects, and mitigation[J]. *Particology*, 2014, 13: 1-26.
- [2] Chan C K, Yao X H. Air pollution in mega cities in China[J]. *Atmospheric Environment*, 2008, 42(1): 1-42.
- [3] Kan H D, Chen R J, Tong S L. Ambient air pollution, climate change, and population health in China[J]. *Environment International*, 2012, 42: 10-19.
- [4] Vandyck T, Keramidas K, Kitous A, et al. Air quality co-benefits for human health and agriculture counterbalance costs to meet Paris Agreement pledges [J]. *Nature Communications*, 2018, 9(1): 4939.
- [5] Zhang Q, Jiang X, Tong D, et al. Transboundary health impacts of transported global air pollution and international trade [J]. *Nature*, 2017, 543(7647): 705-709.
- [6] Wang Y S, Yao L, Wang L L, et al. Mechanism for the formation of the January 2013 heavy haze pollution episode over central and Eastern China [J]. *Science China Earth Sciences*, 2014, 57(1): 14-25.
- [7] Chan K L. Biomass burning sources and their contributions to the local air quality in Hong Kong [J]. *Science of the Total Environment*, 2017, 596/597: 212-221.
- [8] Chan K L, Chan K L. Aerosol optical depths and their contributing sources in Taiwan [J]. *Atmospheric Environment*, 2017, 148: 364-375.
- [9] Taylor M, Retalis A, Flocas H A. Particulate matter estimation from photochemistry: a modelling approach using neural

- networks and synoptic clustering[J]. *Aerosol and Air Quality Research*, 2016, 16(9): 2067-2084.
- [10] Han G, Ma X, Liang A L, et al. Performance evaluation for China's planned CO₂-IPDA[J]. *Remote Sensing*, 2017, 9(8): 768.
- [11] Dong Y N, Du B, Zhang L P, et al. Hyperspectral target detection via adaptive information—theoretic metric learning with local constraints[J]. *Remote Sensing*, 2018, 10(9): 1415.
- [12] Dong Y N, Du B, Zhang L P, et al. Dimensionality reduction and classification of hyperspectral images using ensemble discriminative local metric learning[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2017, 55(5): 2509-2524.
- [13] Zhang R H, Li Q, Zhang R N. Meteorological conditions for the persistent severe fog and haze event over Eastern China in January 2013[J]. *Science China Earth Sciences*, 2014, 57(1): 26-35.
- [14] 傅敏宁, 郑有飞, 徐星生, 等. PM_{2.5} 监测及评价研究进展[J]. *气象与减灾研究*, 2011, 34(4): 1-6.
Fu M N, Zheng Y F, Xu X S, et al. Advances of study on monitoring and evaluation of PM_{2.5} [J]. *Meteorology and Disaster Reduction Research*, 2011, 34(4): 1-6.
- [15] 徐敬, 丁国安, 颜鹏, 等. 北京地区 PM_{2.5} 的成分特征及来源分析[J]. *应用气象学报*, 2007, 18(5): 645-654.
Xu J, Ding G A, Yan P, et al. Component characteristics and sources identification of PM_{2.5} in Beijing[J]. *Journal of Applied Meteorological Science*, 2007, 18(5): 645-654.
- [16] Wu Z L, Liu A X, Zhang C C, et al. Vertical distribution feature of PM_{2.5} and effect of boundary layer in Tianjin [J]. *Urban Environment & Urban Ecology*, 2009, 22(4): 24-29.
- [17] Peng Z R, Wang D S, Wang Z Y, et al. A study of vertical distribution patterns of PM_{2.5} concentrations based on ambient monitoring with unmanned aerial vehicles: a case in Hangzhou, China[J]. *Atmospheric Environment*, 2015, 123: 357-369.
- [18] 赵强, 苏红超, 易明建, 等. 激光雷达探测建筑基坑开挖阶段气溶胶水平分布[J]. *中国激光*, 2021, 48(20): 2010001.
Zhao Q, Su H C, Yi M J, et al. Aerosol horizontal distribution detected by lidar in excavation stage of construction site foundation pit [J]. *Chinese Journal of Lasers*, 2021, 48(20): 2010001.
- [19] 杨昊, 谢晨波, 方志远, 等. 星-地激光雷达联合观测合肥地区的气溶胶垂直分布[J]. *中国激光*, 2021, 48(1): 0110001.
Yang H, Xie C B, Fang Z Y, et al. Joint observations of vertical distribution of aerosols in Hefei area by spaceborne and ground-based lidars [J]. *Chinese Journal of Lasers*, 2021, 48(1): 0110001.
- [20] 张辉, 陶宗明, 麻晓敏, 等. 合肥地区 PM_{2.5} 质量浓度与气溶胶后向散射系数之间吸湿因子的拟合[J]. *中国激光*, 2018, 45(7): 0704006.
Zhang H, Tao Z M, Ma X M, et al. Fitting of hygroscopic factor between PM_{2.5} mass concentration and aerosol backscattering coefficient in Hefei area [J]. *Chinese Journal of Lasers*, 2018, 45(7): 0704006.
- [21] 陶宗明, 麻晓敏, 刘东, 等. 2014 年合肥西郊 PM_{2.5} 质量浓度廓线统计分布[J]. *光学学报*, 2016, 36(6): 0601001.
Tao Z M, Ma X M, Liu D, et al. Statistical distribution of PM_{2.5} mass concentration profiles at west suburb of Hefei City in 2014 [J]. *Acta Optica Sinica*, 2016, 36(6): 0601001.
- [22] 韩道文, 刘文清, 张玉钧, 等. 影响气溶胶质量浓度垂直分布的气象因素分析[C]//光子科技创新与产业化——长三角光子科技创新论坛暨 2006 年安徽博士科技论坛论文集. 合肥: 安徽省科学技术协会, 2006: 445-448.
Han D W, Liu W Q, Zhang Y J, et al. Analysis of meteorological elements, such as temperature, relative humidity, wind speed, and surface pressure, affecting the vertical distribution of aerosol mass concentration [C]//*Photonic Science and Technology Innovation and Industrialization: Proceedings of Yangtze River Delta Photonic Science and Technology Innovation Forum and Anhui Doctoral Science and Technology Forum in 2006*. Hefei: Anhui Association for Science and Technology, 2006: 445-448.
- [23] 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, 104.
Bai S N, Shen X L. PM_{2.5} prediction based on LSTM recurrent neural network [J]. *Computer Applications and Software*, 2019, 36(1): 67-70, 104.
- [24] 章雅楠, 孙建平. 基于小波神经网络的 PM_{2.5} 浓度预测模型[J]. *电力科学与工程*, 2020, 36(1): 55-61.
Zhang Y N, Sun J P. A PM_{2.5} concentration prediction model based on wavelet neural network [J]. *Electric Power Science and Engineering*, 2020, 36(1): 55-61.
- [25] Spinhirne J D. Micro pulse lidar [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 1993, 31(1): 48-55.
- [26] Cairo F, di Donfrancesco G, Adriani A, et al. Comparison of various linear depolarization parameters measured by lidar [J]. *Applied Optics*, 1999, 38(21): 4425-4432.
- [27] 迟如利, 刘东, 范爱媛, 等. 偏振-米散射激光雷达的研制[J]. *应用光学*, 2008, 29(5): 777-784.
Chi R L, Liu D, Fan A Y, et al. Development of polarization Mie-scattering lidar [J]. *Journal of Applied Optics*, 2008, 29(5): 777-784.
- [28] 赵文豪, 闫利, 王成义, 等. 基于三维大气探测激光雷达的大气颗粒物污染探测[J]. *武汉大学学报·信息科学版*, 2019, 44(10): 1436-1441.
Zhao W H, Yan L, Wang C Y, et al. Detection of atmospheric particulate matter pollution based on three-dimensional atmospheric detection LiDAR [J]. *Geomatics and Information Science of Wuhan University*, 2019, 44(10): 1436-1441.
- [29] Bi J R, Huang J P, Shi J S, et al. Measurement of scattering and absorption properties of dust aerosol in a Gobi farmland region of northwestern China—a potential anthropogenic influence [J]. *Atmospheric Chemistry and Physics*, 2017, 17(12): 7775-7792.
- [30] 胡局新, 张功杰. 基于 K 折交叉验证的选择性集成分类算法[J]. *科技通报*, 2013, 29(12): 115-117.
Hu J X, Zhang G J. K-fold cross-validation based selected ensemble classification algorithm [J]. *Bulletin of Science and Technology*, 2013, 29(12): 115-117.
- [31] Chasset P O. GRNN: general regression neural network [J]. *Revue De Physique Appliquée*, 2013, IV(6): 1321-1325.
- [32] 贾鹏程, 曹念文, 范广强, 等. 重污染过程的差分吸收激光雷达监测[J]. *激光与光电子学进展*, 2021, 58(9): 0901002.
Jia P C, Cao N W, Fan G Q, et al. Differential absorption lidar monitoring of heavy pollution process [J]. *Laser & Optoelectronics Progress*, 2021, 58(9): 0901002.

Estimation of Particulate Matter Mass Concentration Based on Generalized Regression Neural Network Model Combining Aerosol Extinction Coefficient and Meteorological Elements

Mo Zusi¹, Bu Lingbing^{1*}, Wang Qin¹, Lin Xuefei¹, Samuel A. Berhane¹, Yang Bin²,
Deng Chen, Li Zhi²

¹ Collaborative Innovation Center on Forecast and Evaluation of Meteorological Disasters, School of Atmospheric Physics, Nanjing University of Information Science and Technology, Nanjing 210044, Jiangsu, China;

² Nanjing Mulei Laser Technology Co., Ltd., Nanjing 210038, Jiangsu, China

Abstract

Objective Atmospheric particulate matter is regarded as one of the most serious air pollutants that endanger human health. Lidar detection is a viable method for achieving high-precision particle distribution measurements. To some extent, the aerosol extinction coefficient reflects the relative size of aerosol mass concentration. Meteorological elements, such as temperature, relative humidity, wind speed, and surface pressure, have a significant impact on the extinction coefficient and mass concentration. In this study, a dataset of PM_{2.5} and PM₁₀ was created by combining the extinction coefficient obtained from lidar and surface meteorological elements such as temperature, relative humidity, wind speed, and surface pressure. The data characteristics were calculated using principal component analysis. The mass concentration of PM_{2.5} and PM₁₀ were estimated using the generalized regression neural network (GRNN) model. The results show that the correlation coefficients between the estimated mass concentrations obtained by the GRNN model and the true values collected from environmental monitoring stations for PM_{2.5} and PM₁₀ were 0.86 and 0.85, respectively. Moreover, the root mean square errors (RMSEs) were 2.58 μm/cm³ and 10.84 μm/cm³, and the mean absolute deviations (MAEs) were 0.81 μg/m³ and 1.53 μg/m³ for PM_{2.5} and PM₁₀, respectively. The GRNN model was applied to lidar scanning mode to evaluate the horizontal distribution characteristics of particle over Pukou District of Nanjing. The estimated mass concentrations of PM_{2.5} and PM₁₀ were consistent with the measured values, demonstrating the GRNN model's effectiveness in particle mass concentration evaluation.

Methods The aerosol extinction coefficient, wind speed, temperature, relative humidity, and surface pressure were used as input variables in the GRNN, and the PM_{2.5} and PM₁₀ mass concentrations were used as output variables for model training and verification. First, principal component analysis was used to calculate the characteristics of sample data, and the effects of various meteorological factors on aerosol mass concentration were fully considered. Furthermore, the better evaluation model was continuously trained by adjusting parameters. The GRNN evaluation model was used to evaluate the temporal and spatial changes of particulate matter mass concentration in the Pukou District of Nanjing. Simultaneously, it was verified and analyzed in conjunction with ground meteorological elements and data from environmental monitoring stations.

Results and Discussions The GRNN model was used to establish and validate an evaluation model of PM_{2.5} and PM₁₀ mass concentrations based on the extinction coefficient obtained from lidar and surface meteorological factors. The correlation coefficients between the estimated mass concentrations obtained by the GRNN model and the measured values were 0.86 and 0.85, the RMSE were 2.58 μg/m³ and 10.84 μg/m³, and the mean absolute errors (MAE) were 0.81 μg/m³ and 1.53 μg/m³ for PM_{2.5} and PM₁₀, respectively. The GRNN model was also used to estimate and analyze particulate matter mass concentration using lidar horizontal scanning. The trend of the model's evaluation values were consistent with the measured values from the environmental monitoring station.

Conclusions The aerosol extinction coefficient was calculated using the GRNN model and the lidar inversion result. The correlation model was constructed using the hourly mean value of mass concentration. The model variables were screened using principal component analysis, and *K*-fold cross-validation was performed on the training data. The optimal sliding factor was chosen using the minimum mean square error index. The output model's results were reliable, with correlation coefficients between predicted and actual values of 0.86 and 0.85 for PM_{2.5} and PM₁₀ mass concentrations, respectively. Moreover, the RMSEs were 2.58 μg/m³ and 10.84 μg/m³ and the average absolute errors were 0.81 μg/m³ and 1.53 μg/m³, for PM_{2.5} and PM₁₀ mass concentration, respectively. The Pukou District in Nanjing was observed horizontally using lidar, combined with local temperature, relative humidity, wind speed, surface pressure, extinction coefficient, the GRNN model was employed to estimate the mass concentrations of PM_{2.5} and PM₁₀. The correlation

analysis results between the estimated mass concentration of $PM_{2.5}$ and PM_{10} and the measured values collected from the environmental monitoring station further validated the GRNN model's accuracy. The results demonstrated that the model could be used to monitor $PM_{2.5}$ and PM_{10} mass concentrations using lidar, and it provided a new method of monitoring particulate matter pollution over a wide range with high temporal and spatial resolutions.

Key words remote sensing; aerosol extinction coefficient; particulate matter mass concentration; neural networks; meteorological elements