

## 基于深度学习反演区域气溶胶光学厚度

梁天辰, 孙林\*, 王永吉

山东科技大学测绘科学与工程学院, 山东 青岛 266590

**摘要** 为了解决现有陆地气溶胶光学厚度(AOD)反演算法精度和空间分辨率较低的问题,基于深度学习的思想,使用深度置信神经网络(DBN),实现了具有 30 m 空间分辨率的陆地气溶胶光学厚度反演。算法的训练样本包括全球长时间序列的 AERONET 站点数据以及在时空上与之对应的 Landsat8 OLI 的观测几何数据和表观反射率数据。为了保证反演的精度和稳定性,研究了 AERONET 站点数据的处理方法、卫星与站点数据的时空匹配方法以及 DBN 结构的设置。使用独立于训练样本的 AERONET 站点数据,对不同地表类型的 550 nm 处的 AOD 估算结果进行了整体验证,并对研究区域进行小尺度精度验证。结果表明,该方法的均方根误差和平均绝对误差分别为 0.11 与 0.072。该方法打破了现有的气溶胶光学厚度反演方法依赖于其他遥感产品或者其他时相数据的局面,有效提高了气溶胶光学厚度反演的效率和空间分辨率。

**关键词** 大气光学; 气溶胶光学厚度; 深度学习; Landsat8 OLI 数据

**中图分类号** P407

**文献标志码** A

**doi:** 10.3788/AOS202141.0401002

## Retrieval of Regional Aerosol Optical Depth Using Deep Learning

Liang Tianchen, Sun Lin\*, Wang Yongji

*College of Geomatics, Shandong University of Science and Technology, Qingdao, Shandong 266590, China*

**Abstract** To solve the problem that there exist low precision and spatial resolution in the retrieval algorithm of land aerosol optical depth (AOD), a deep learning-based deep belief network (DBN) is proposed to realize the retrieval of land AOD with a spatial resolution of 30 m. The training samples for the algorithm include the AERONET site data with global long time series as well as the observation geometric data and apparent reflectivity data from Landsat 8 OLI which are corresponding to the former in space and time. To ensure the estimation accuracy and stability of retrieval, the process method for the AERONET site data, the spatial-temporal matching method for satellite and site data, and the setting of the DBN structure are investigated. The AERONET site data, independent of the training samples, are used to test the AOD estimation results at 550 nm for different surface types as a whole. In addition, the small-scale accuracy verification is conducted in the study area. The results demonstrate that the root mean square error and the mean absolute error of the proposed method are 0.11 and 0.072, respectively. The proposed method can break the situation in which the retrieval of AOD based on the existing methods relies excessively on other remote sensing products or time-phase data, and it effectively improves the efficiency and spatial resolution in the retrieval of AOD.

**Key words** atmospheric optics; aerosol optical depth; deep learning; Landsat8 OLI data

**OCIS codes** 010.1100; 010.1110; 010.1290; 010.0280

## 1 引言

大气气溶胶通过辐射强迫和化学扰动在地球辐

射收支中发挥重要作用。大气气溶胶粒子通过对太阳辐射和地球长波辐射的散射、吸收,直接影响天气系统的辐射能量收支。其次,气溶胶可以作为凝结

收稿日期: 2020-08-04; 修回日期: 2020-08-30; 录用日期: 2020-09-25

基金项目: 国家自然科学基金(41771408)、山东省自然科学基金(ZR201702210379)

\* E-mail: sunlin6@126.com

核影响云的辐射特性,增加云量,间接影响气候<sup>[1]</sup>。气溶胶作为一种主要的空气污染物,严重危害人类健康<sup>[2-5]</sup>。与此同时,它影响卫星成像的质量,使图像模糊,从而影响地表信息的精确提取,是遥感数据大气校正的重要输入参数之一。因此,气溶胶信息的准确获取在气候变化、环境污染监测和卫星数据利用中具有重要作用。其中,气溶胶光学厚度(Aerosol Optical Depth, AOD)是卫星遥感监测的重要参数之一,是监测气溶胶的重要指标。目前,许多卫星在气溶胶影响环境和气候的研究中发挥了重要作用,例如 Moderate Resolution Imaging Spectroradiometer (MODIS)、the Visible Infrared Imaging Radiometer Suite (VIIRS) 和 the Multiangle Imaging Spectro Radiometer (MISR) 等。相较于这些卫星,Landsat8 具有更高的空间分辨率,能够提供更多的局部细节,有利于区域性气溶胶研究。

现有多种基于物理模型、结构函数和机器学习等方法的气溶胶光学厚度反演算法。其中应用较为广泛的是基于物理模型方法的气溶胶光学厚度反演算法,其利用辐射传输模型模拟地表、大气和大气顶层的关系,从而构建查找表并从中检索气溶胶光学厚度,例如浓密植被算法<sup>[6]</sup>和深蓝算法<sup>[7]</sup>。气溶胶算法的关键就是如何精准确定地表反射率,浓密植被法通过统计学方法,发现红、蓝波段的地表反射率较低且与中红外波段的地表反射率存在明显的线性关系,进而确定了红、蓝波段的地表反射率;深蓝算法利用 SeaWiFS 地表反射率产品,通过最小值合成法,构建了地表反射率库以提供红蓝波段地表反射率,进而求解气溶胶光学厚度。这些算法被成功地应用于 MODIS 传感器中并用于生产全球气溶胶产品(MOD04)。MOD04 与地基观测结果具有较高的相关性且适合大尺度应用,但 MOD04 的空间分辨率仅为 3 km 和 10 km,不利于小尺度分析。同时这些算法是建立在一些参数(例如地表反射率)假设基础上的,这些假设会在某些区域产生一定的误差,从而影响气溶胶光学厚度反演的精度。例如,浓密植被法确定的地表反射率的统计关系只在暗黑像元处稳定,在城市高亮地表处有较大偏差。

研究者针对地表反射率确定问题进行了诸多改进,提出了多个高精度的气溶胶光学厚度反演算法,例如 Multi Angle Implementation of Atmospheric Correction(MAIA)算法<sup>[8]</sup>和 Simplified Aerosol Retrieval Algorithm (SARA)<sup>[9]</sup>。MAIA 算法通

过分析 MODIS 的长时间序列数据来确定地表反射率,此算法在反演 AOD 的同时也能反演细粒子分布,最终生成具有 1 km 分辨率的气溶胶产品,有效地提高了 AOD 的反演精度与分辨率。SARA 基于 MODIS 地表反射率产品(MOD09GA),提供了地表反射率,通过提取分析研究区域对应的 AERONET 站点参数,改进了区域气溶胶模型。此算法成功应用于京津冀以及香港地区。但是,在 SARA 中,下垫面被假设为朗伯表面且气溶胶参数在同一区域中被假设为短时间保持不变,这使得 SARA 在不同地区的应用会出现偏差。

为了避免不同的参数假设,Li 等<sup>[10]</sup>直接使用站点和卫星数据,利用自回归综合移动平均(Autoregressive Integrated Moving Average, ARIMA)法,进行了长期气溶胶趋势分析并预测了气溶胶光学厚度<sup>[11]</sup>。结果表明,利用时间序列数据进行建模并预测气溶胶光学厚度是可行的,该方法在区域范围内表现出良好的精度。此外,Radosavljevic 等<sup>[12-13]</sup>利用神经网络直接学习卫星观测气溶胶光学厚度与地面观测气溶胶光学厚度之间的关系。相较于物理模型,深度学习方法能够获得更高的精度。Huttunen 等<sup>[14]</sup>对比分析了查找表方法、非线性回归方法和四种机器学习方法的优劣,结果表明,大多数机器学习方法能够得到更高精度的气溶胶光学厚度反演结果。

卫星表观反射率与气溶胶光学厚度之间的关系是一个复杂的非线性物理问题,但机器学习是一个强大的学习工具,可以基于样本数据进行学习并实现对复杂方程式的精确求解,已被用于多类时间序列预测中。因此,本文利用深度学习算法,直接从 Landsat8 陆地成像仪(Operational Land Imager, OLI)影像中估算气溶胶光学厚度。具体来说,本文利用概率学习模型即深度置信网络(Deep Belief Network, DBN)<sup>[15]</sup>,建立了地表气溶胶光学厚度与卫星大气顶部(Top of Atmosphere, TOA)反射率、观测角度之间的关系。为了使样本更具代表性并且能够充分学习非线性关系,本文选取分布在全球且具有长时间序列的 AERONET 站点数据,根据时间和空间信息,将其与 Landsat8 表观反射率数据匹配,从而形成模型的训练样本。在样本训练中,Landsat8 提供 TOA 反射率以表征地表大气耦合的影响,利用归一化差分植被指数(Normalized Difference Vegetation Index, NDVI)来区分不同地表类型,利用观测几何数据表征各种成像条件。本

文利用 DBN 对卫星数据和地表测量的 AOD 之间的关系进行训练建模,从而检索特定区域的气溶胶光学厚度。对 2019 年所有站点数据进行了预测并验证了它们的精度,对研究区域的小尺度气溶胶光学厚度进行了更精细的验证。

## 2 数据及研究区域

### 2.1 研究区域

本文选取美国西北部的爱达荷州以及英格兰东南部的大伦敦作为研究区域。爱达荷州的地势起伏较大,森林茂密,植被丰富,森林覆盖率达到 33%,并且水源丰富,其境内的 AERONET-DRAGON 网络可以对 AOD 估算质量进行验证。大伦敦地区的下垫面以城市人造地表为主,植被覆盖率约为 13%;其地势较低,以平原为主,地形平坦;同时,伦敦受北大西洋暖流和西风的影响,气候湿润且多雨雾。

### 2.2 地面气溶胶光学厚度的测量

AERONET(Aerosol Robotic Network)是一个全球地基气溶胶观测网络,可提供稳定且易于获取的气溶胶光学特性数据库。它具有较高的时间分辨率,每隔 15 min 便提供一次观测结果,并在全球范围内设置有 400 个以上有效观测站点,常被用于气溶胶研究和卫星遥感气溶胶反演验证等。AERONET 提供三个级别的 AOD 观测数据:Level 1.0(未筛选过的产品),Level 1.5(云剔除产品)以及 Level 2.0(云剔除和质量控制产品)。AERONET 新的 V3 版本数据使用自动化的近实时质量控制算法,AOD 计算结果更加精确(不确定度为 0.01~0.02),并利用自动云筛选代替了先前的手动质量控制,Level 2.0 数据更加稳定<sup>[16]</sup>。本文使用 Level 2.0 站点数据进行训练样本收集,在全球范围内挑选 455 个 AERONET 站点以尽可能涵盖更多的气溶胶类型。本文需要收集波长 550 nm 处的 AOD 以组成训练样本并验证模型估算精度,但 AERONET 站点缺少波长 550 nm 的数据,因此使用 Ångström 指数<sup>[17]</sup>的方法进行波段转化,具体公式为

$$\tau_a(\lambda) = \beta\lambda^{-a}, \quad (1)$$

$$a = -\frac{\ln\left[\frac{\tau_a(\lambda_1)}{\tau_a(\lambda_2)}\right]}{\ln\left(\frac{\lambda_1}{\lambda_2}\right)}, \quad (2)$$

$$\beta = \frac{\tau_a(\lambda_1)}{\lambda_1^{-a}}, \quad (3)$$

式中: $\tau_a(\lambda)$ 为波长为 $\lambda$ 的气溶胶光学厚度; $a$ 、 $\beta$ 分别为 Ångström 指数和大气浑浊度系数。分别使用受水汽影响较小且与 550 nm 最接近的 500 nm 和 675 nm 作为波长 $\lambda_1$ 、 $\lambda_2$ ,求解两系数值 $a$ 和 $\beta$ 。

### 2.3 卫星数据

Landsat8 搭载有先进的 Operational Land Imager (OLI)与 Thermal Infrared Sensor (TIRS)传感器,其中 OLI 的空间分辨率为 30 m,卫星每 16 d 就可以实现一次全球覆盖,多光谱波段覆盖了可见光到短波红外波段,波段设置如表 1 所示,其中 NIR 为近红外波段,SWIR 为短波近红外波段,Pan 为全色波段,Cirrus 为卷云波段。

表 1 Landsat8 OLI 数据

Table 1 Landsat8 OLI data

Band No.	Spectral range / $\mu\text{m}$	Spatial resolution /m
B1 (Coastal)	0.43–0.45	30
B2 (Blue)	0.45–0.51	30
B3 (Green)	0.53–0.59	30
B4 (Red)	0.64–0.67	30
B5 (NIR)	0.85–0.88	30
B6 (SWIR 1)	1.57–1.65	30
B7 (SWIR 2)	2.11–2.29	30
B8 (Pan)	0.50–0.68	15
B9 (Cirrus)	1.36–1.38	30

OLI 不仅具有较高的空间分辨率,而且波段覆盖广,其在气溶胶监测方面得到广泛应用。Google Earth Engine (GEE)是一个基于云的行星规模地理空间分析平台,在云端就可以实现全球尺度遥感数据的处理。本文利用 GEE,在全球范围内提取了 2013—2019 年的 Landsat8 表观反射率数据。Landsat8 数据含有 B1~B7 波段的 TOA 反射率数据,波长从可见光到近红外的辐射光谱、相应的观测几何数据(太阳高度角和相对方位角)、NDVI 数据和质量控制波段的数据。NDVI 计算公式<sup>[18]</sup>为

$$N_D = \frac{B_5 - B_4}{B_5 + B_4} \quad (4)$$

式中: $N_D$ 为 NDVI; $B_4$ 、 $B_5$ 分别为 Landsat8 卫星 B4、B5 波段的反射率。计算后的 NDVI 按月合成月数据集,将其代入 AOD 估算模型以表征不同的土地覆盖类型。另外,在质量评估波段(Quality Assessment Band, QA),通过受仪器或云层影响的信号标示像元,能够快速识别污染像元。本文基于 Landsat8 卫星 2013—2018 年的数据进行样本收集,并利用 QA 波段进行云像元的剔除,其中 2018 年部分数据用于生成独立验证集,2019 年的数据用

于 AOD 估算。

## 2.4 数据时空匹配

为了进行时间匹配,本文选择卫星过境前后 30 min 内的 AERONET 站点测量值与卫星数据进行匹配。若区间内出现多组站点的测量值,则本文采取线性插值得到当前时刻的 AOD:

$$\tau = \tau_a \frac{t_b - t}{t_b - t_a} + \tau_b \frac{t - t_a}{t_b - t_a}, \quad (5)$$

式中: $\tau$  为成像时刻的 AOD; $t_a$ 、 $t_b$  为与成像时刻最为接近的两个时间点; $\tau_a$ 、 $\tau_b$  分别为对应时刻的 AOD; $t$  为 Landsat8 卫星成像时刻。在空间匹配方面,本文利用 AERONET 站点的经纬度信息,通过影像投影进行逐像元匹配。本文共收集 20003 条训练样本。

## 3 DBN 的气溶胶光学厚度估算

Landsat8 卫星 B7 波段的波长较长且受大气影响较小,B2 和 B4 波段信息可以表征蓝、红通道上的地表与大气耦合反射率,因此地表反射率可以通过 B2、B4、B7 三个波段之间的统计经验关系进行分离<sup>[6]</sup>。Landsat8 卫星新增加的 B1 波段相较于蓝波段对气溶胶更加敏感,更易受到气溶胶的影响。因此,可以从这些输入变量中得到大气对卫星 TOA 反射率信号的影响程度;剩余的波段能够增加更多的信息,有利于模型学习其中的关系。但是,AOD、卫星 TOA 反射率、观测角度以及 NDVI 之间的关系非常复杂。深度学习在拟合非线性和复杂关系方面具有很大潜力,可以更好地表征这种关系。

本文利用 DBN 拟合参数之间的关系。DBN 是一种概率生成模型,它通过采用逐层训练的方式,解决了深层次神经网络的优化问题,为整个网络赋予了较好的初始权值,使得网络只要经过微调就可以实现最优解。其中,逐层训练的关键是深度波尔兹曼机 (Restricted Boltzmann Machine, RBM),RBM 包含了一个用于输入训练数据的显层 ( $v$ ),还有一层作为特征检测器的隐含层 ( $h$ ),层内无连接,层与层之间全连接,这大大减少了计算量和计算难度。RBM 的训练过程实际上就是求解最大概率分布,由于这个分布的决定性因素在于权值  $W$ ,因此训练 RBM 的目标就是寻找最佳的权值。在 DBN 中,前一层的 RBM 隐含层连接着后一层的可见层,多层堆叠的 RBM 组合后通过连接一层后向传播 (Back-Propagation, BP) 层<sup>[19]</sup>来进行微调输出,后向传播层不仅可以应用于预测也可以应

用于分类,本文利用其估算 AOD。图 1 显示了本文所使用的 DBN 网络结构;理论上 RBM 层越多越容易预先学习参数之间的关系,但过多的层容易导致过拟合。神经元个数一般通过经验公式<sup>[20]</sup>获得:

$$N \in (2\sqrt{n} + \mu, 2n + 1), \quad (6)$$

式中: $N$  为神经元个数; $n$  为特征个数; $\mu$  为输出变量个数。为了充分学习参数之间的关系并防止过拟合,本文使用了两层 RBM 层,隐含层的神经元个数设置为 20。利用深度置信神经网络估算 AOD 主要分为三大步骤,流程图如图 2 所示。

1) 将卫星 TOA 反射率、观测角度和 NDVI 输入 DBN 模型,使用 RBM 对收集的样本进行逐层的非监督预训练,非监督预训练能够提取出必要特征并将其传递给下一层 RBM。

2) 利用步骤 1) 的结果生成初始权重并计算得到 AOD; 随后计算模型预估的 AOD 与 AERONET 实际站点数据间的均方根误差 (MSE),将误差返回 BP 层并进行反复微调,直到模型达到满意的性能。

3) 模型在使用前需要进行交叉验证,精度满足后利用其估算没有地面监测站地区的 AOD 值,从而进一步得到 Landsat8 卫星的气溶胶光学厚度分布。

本文采用十折交叉验证 (10-fold cross-validation) 技术<sup>[21]</sup>来检验模型的拟合和估算能力,采用基于样本的交叉验证<sup>[22]</sup>创建独立的验证数据集。在基于训练样本的交叉验证中,本文将训练数据分为 10 个子集,其中 9 个用来训练,剩余的 1 个用来验证;并且在 2018 年的站点数据中,挑选与本文研究区域具有类似气溶胶特性的站点数据以形成独立的验证数据集,进而进行交叉验证。本文使用相关系数、均方根误差以及平均绝对误差来评价验证精度。

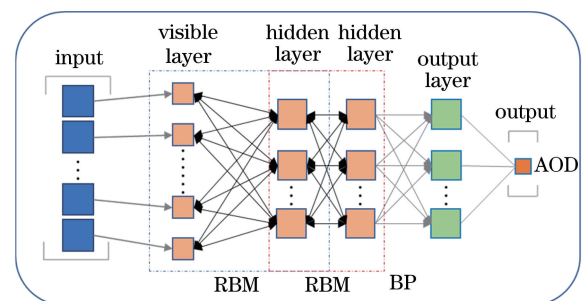


图 1 DBN 架构

Fig. 1 Structural diagram of DBN

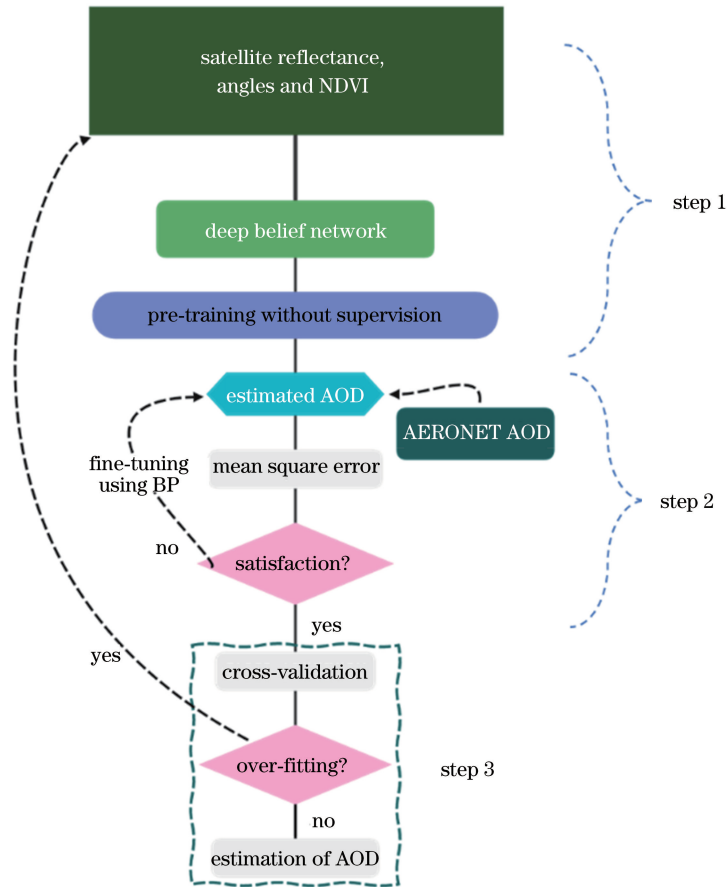


图 2 DBN 的流程图

Fig. 2 Flow chart of DBN

## 4 实验结果与分析

### 4.1 AERONET 站点验证

为了初步验证 AOD 模型的估算能力,首先利用 2019 年收集的全部数据进行 550 nm 处的 AOD 估算并利用 AERONET 站点进行精度验证。本文在 500 m 分辨率的尺度下验证 AOD 的估算精度,由于估算的是 30 m 分辨率下的 AOD,需要进行像元合成。为了尽可能减少云对验证结果的影响,本文采用如下步骤进行像元合成:1)以站点为中心,取大小为 17 pixel × 17 pixel 的窗口;2)将所有像元按照 B2 波段(蓝波段)的数值从小到大进行排序;3)删除蓝波段反射率最高的像元(占比 30%)和蓝波段最暗的像元(占比 20%);4)剩余像元按照 QA 波段信息进行云判断,非云的像元取平均即为最终结果。本文使用卫星过境前后 30 min 的站点数据进行验证,并利用皮尔逊积矩相关系数( $R$ )、均方根误差( $R_M$ )、平均绝对误差( $M_A$ )和 MODIS 期望误差( $E_E$ )来评价精度。计算公式分别为

$$R = \frac{\text{Cov}(A_{\text{prediction}}, A_{\text{AERONET}})}{\sqrt{\text{Var}[A_{\text{AERONET}}]\text{Var}[A_{\text{prediction}}]}}, \quad (7)$$

$$R_M = \sqrt{\frac{1}{n} \sum_{i=0}^n (A_{\text{prediction}} - A_{\text{AERONET}})^2}, \quad (8)$$

$$M_A = \frac{1}{n} \sum_{i=0}^n |A_{\text{prediction}} - A_{\text{AERONET}}|, \quad (9)$$

$$E_E = \pm (0.05 + 0.15 \times A_{\text{AERONET}}), \quad (10)$$

式中: $A_{\text{prediction}}$  为模型估计结果; $A_{\text{AERONET}}$  为站点观测结果; $\text{Cov}(\cdot, \cdot)$  为协方差; $\text{Var}(\cdot)$  为方差。本文在全球范围内共匹配到 233 个站点的数据,共计 1962 条有效观测数据。图 3 为反演结果与站点结果的密度散点图,约有 69.5% 的数据落在  $E_E$  误差线内,均方根误差和平均绝对误差分别为 0.11 和 0.07。整体预测结果显示了较高的精度,数据点密度最为集中的地方十分靠近  $y = x$  回归线,这初步证明了方法的可行性;但是在 AOD 小于 0.2 的情况下,本文预测结果偏高,这可能是由于 Landsat8 卫星较长的回访周期仅提供有限的观测数据,再加上 AERONET 站点在全球的分布不均匀,从而部分地区的 AOD 预测受到影响。对全球 AERONET

站点测量值与模型估算 AOD 的  $R_M/M_A$  进行了分析,具体统计结果如表 2 所示。在所有 AERONET 站点测量值中,约有 78.1% (182/233) 的  $R_M$  小于 0.12,  $R_M$  值较小的站点大多分布在北美、南美和欧洲,这些地区的站点分布相对密集,提供了大量的训练样本数据,且这些地区的气溶胶特征的年变化和季节变化较小,从而可以获得更好的预测效果;相反,在印度、非洲沙漠和东南亚等国家,  $R_M$  的值基本超过了 0.12, 这些地区 AERONET 站点的分布稀疏,没有充足的训练样本支撑并且此地区大多为发展中国家,空气污染较为严重,气溶胶在空间中的变化剧烈。在所有站点测量值中,大约 76.4% (178/233) 的  $M_A$  小于 0.09, 这些站点主要集中在北美、南美、欧洲和大洋洲地区。  $M_A$  较高的站点主要集中在印度和东南亚等国家, 这些地区拥有复杂的高亮人造地表, 并且人类活动的加剧迫使气溶胶特征变化剧烈。相比之下,在北美、南美和欧洲等拥有大量观测点的地区,稳定的气溶胶特征有利于精度的提高。

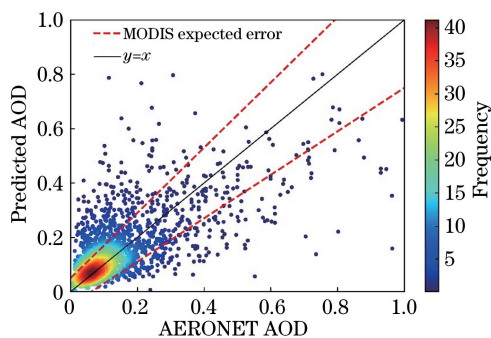


图 3 AOD 预测值与 AERONET 站点观测值的密度散点图  
Fig. 3 Scatter density map of AOD prediction value and AERONET site observation value

为了评价算法的区域性估算精度,本文获取了实验区域内 2019 年共 10 个站点的数据并进行区域精度评价,其中大伦敦地区 2 个,爱达荷州 8 个。图 4 为研究区域站点测量值与模型估算结果的对比图,可以看出,模型估算结果和站点测量值之间有着较高的相关系数,约为 0.79, 44 个验证数据具有较低的均方根误差和平均绝对误差,分别为 0.045 和 0.034;爱达荷州的 AERONET-DRAGON 网络表现出较高精度,证明了本文算法在浓密植被地区的 AOD 估算能力,该算法在伦敦地区也具有较高精度。这表明本文算法可以在研究区域内估算高分辨率的 AOD,并具有较高的精度。

表 2 误差统计

Table 2 Error statistics

Position	R	$R_M$	$M_A$	$E_E$
Europe	0.45	0.083	0.059	73.12
North America	0.53	0.077	0.049	79.36
South America	0.79	0.069	0.048	80.05
Africa	0.70	0.152	0.098	64.77
Oceania	0.73	0.053	0.041	84.58
Asia	0.75	0.142	0.100	61.58

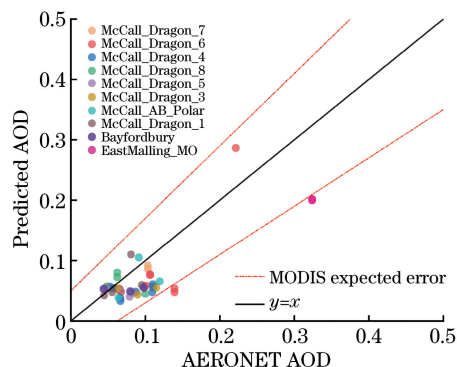


图 4 模型预测结果与不同站点实测值的对比

Fig. 4 Comparison between model prediction results and measurement results at different sites

#### 4.2 气溶胶的空间分布

在所选研究区域内评估气溶胶的空间分布和变化情况。为了估算更高精度的气溶胶光学厚度,本文首先利用 Sun 等<sup>[23]</sup>提出的基于 GlobeLand30 数据库的云检测算法进行云水像元的剔除,同时使用归一化冰雪指数 (NDSI,  $N_s$ ) 进一步去除雪体。NDSI 的计算公式为

$$N_s = \frac{B_3 - B_6}{B_3 + B_6}, \quad (11)$$

式中:  $B_3$  为绿波段反射率;  $B_6$  为短波近红外波段的反射率。

图 5 为不同研究区域的真彩色图像与 AOD 估算图的空间分布。从图 5(a)真彩色图像可以看出,在影像的左边以及大伦敦地区(箭头所指区域)有大范围的薄雾,本文所得结果具有较好的空间连续性。如图 5(b)所示,本文算法很好地捕捉了大伦敦地区(箭头所指区域)薄雾的发生情况。图 5(c)所示为北美研究区域,此地区具有较多的浓密植被,预测结果十分稳定且分布连续。本文算法能够反演 30 m 空间分辨率下的气溶胶影像,较高的空间分辨率适合小尺度研究,可以提供更多的局部细节,有利于大气监测。

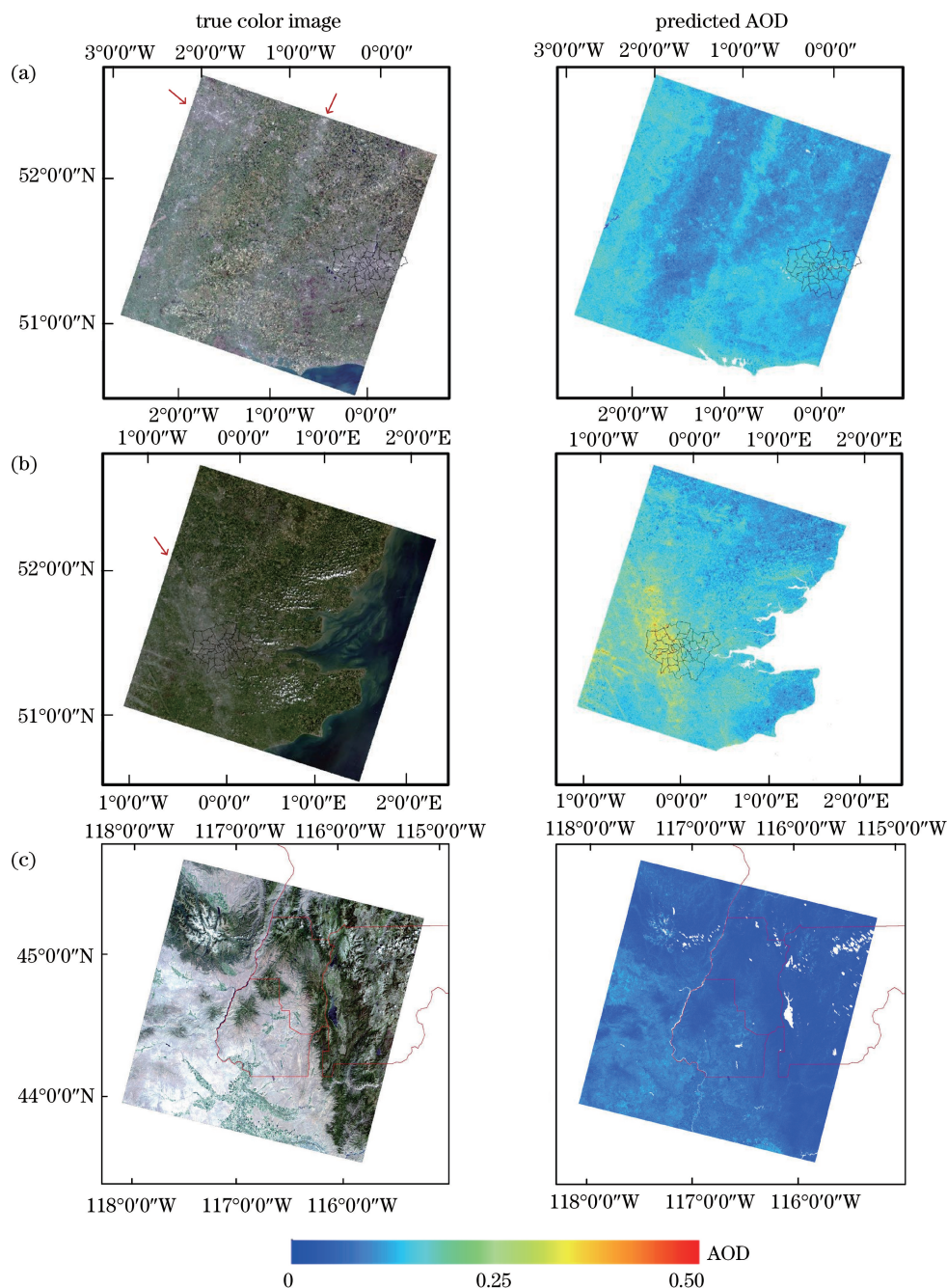


图 5 气溶胶光学厚度空间分布图

Fig. 5 Spatial distribution of aerosol optical depth

## 5 结 论

提出了一种基于 DBN 的 Landsat8-OLI 气溶胶反演算法。使用 GEE 在全球范围内收集训练样本, 将 2013—2018 年的 AERONET 站点数据与 Landsat8 表观反射率数据进行匹配以形成训练样本, 并利用 2019 年的影像数据进行全球站点的预测反演。结果表明, 在全球范围内, 所提算法可以预测出具有较高精度的 AOD, 均方根误差和平均绝对误

差分别为 0.11 和 0.07, 约有 69.5% 的数据落在  $E_E$  误差线内。在小尺度范围内进行了 AOD 的验证, 结果表明, AOD 预测结果和站点测量结果之间有着较高的相关系数(0.79), 预测结果具有较低的均方根误差和平均绝对误差, 分别为 0.045 和 0.034。验证结果证明了所提算法具有较高的精度且可靠性好。最后对气溶胶的空间分布进行了分析, 所提算法能够在浓密植被区域提供稳定、连续的气溶胶空间分布结果, 并能够在城市高亮地表精准地识别薄

雾。与传统方法相比,所提算法仅利用单相卫星遥感数据即可实现气溶胶光学厚度反演,无需复杂的物理建模,简化了反演过程。所提算法仍存在一些不足,由于 Landsat8 卫星的回访周期较长,部分地区站点数据的不充足会造成训练样本分布不均匀,东南亚及非洲沙漠等地区的气溶胶评估结果变得不可靠。这些地区气溶胶变化剧烈,在这些区域反演出具有高分辨率和高精度的气溶胶光学厚度是未来着重的发展方向。

### 参 考 文 献

- [1] Pöschl U. Atmospheric aerosols: composition, transformation, climate and health effects [J]. *Angewandte Chemie International Edition*, 2005, 44 (46): 7520-7540.
- [2] Gauderman W J, Avol E, Gilliland F, et al. The effect of air pollution on lung development from 10 to 18 years of age [J]. *The New England Journal of Medicine*, 2004, 351(11): 1057-1067.
- [3] Künzli N, Kaiser R, Medina S, et al. Public-health impact of outdoor and traffic-related air pollution: a European assessment [J]. *The Lancet*, 2000, 356 (9232): 795-801.
- [4] Gehring U, Cyrus J, Sedlmeir G, et al. Traffic-related air pollution and respiratory health during the first 2 yrs of life [J]. *European Respiratory Journal*, 2002, 19(4): 690-698.
- [5] Volk H E, Lurmann F, Penfold B, et al. Traffic-related air pollution, particulate matter, and autism [J]. *JAMA Psychiatry*, 2013, 70(1): 71-77.
- [6] Kaufman Y J, Tanré D, Remer L A, et al. Operational remote sensing of tropospheric aerosol over land from EOS moderate resolution imaging spectroradiometer [J]. *Journal of Geophysical Research: Atmospheres*, 1997, 102 (D14): 17051-17067.
- [7] Hsu N C, Jeong M J, Bettenhausen C, et al. Enhanced deep blue aerosol retrieval algorithm: the second generation [J]. *Journal of Geophysical Research: Atmospheres*, 2013, 118 (16): 9296-9315.
- [8] Lyapustin A, Wang Y, Laszlo I, et al. Multiangle implementation of atmospheric correction (MAIAC): 2. Aerosol algorithm [J]. *Journal of Geophysical Research: Atmospheres*, 2011, 116(D3): D03211.
- [9] Bilal M, Nichol J E, Bleiweiss M P, et al. A simplified high resolution MODIS aerosol retrieval algorithm (SARA) for use over mixed surfaces [J]. *Remote Sensing of Environment*, 2013, 136: 135-145.
- [10] Li X K, Zhang C R, Zhang B, et al. A comparative time series analysis and modeling of aerosols in the contiguous United States and China [J]. *Science of the Total Environment*, 2019, 690: 799-811.
- [11] Li X K, Zhang C R, Li W D, et al. Exploring the trend, prediction and driving forces of aerosols using satellite and ground data, and implications for climate change mitigation [J]. *Journal of Cleaner Production*, 2019, 223: 238-251.
- [12] Radosavljevic V, Vucetic S, Obradovic Z. Aerosol optical depth retrieval by neural networks ensemble with adaptive cost function [EB/OL]. (2017-09-28) [2020-05-10]. <https://pdfs.semanticscholar.org/6e7a/00df19d6e2b3698073d1f96a12f16f3c5a89.pdf>.
- [13] Ristovski K, Vucetic S, Obradovic Z. Uncertainty analysis of neural-network-based aerosol retrieval [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2012, 50(2): 409-414.
- [14] Huttunen J, Kokkola H, Mielonen T, et al. Retrieval of aerosol optical depth from surface solar radiation measurements using machine learning algorithms, non-linear regression and a radiative transfer-based look-up table [J]. *Atmospheric Chemistry and Physics*, 2016, 16(13): 8181-8191.
- [15] Hinton G E, Osindero S, Teh Y W. A fast learning algorithm for deep belief nets [J]. *Neural Computation*, 2006, 18(7): 1527-1554.
- [16] Giles D M, Sinyuk A, Sorokin M G, et al. Advancements in the Aerosol Robotic Network (AERONET) Version 3 database: automated near-real-time quality control algorithm with improved cloud screening for Sun photometer aerosol optical depth (AOD) measurements [J]. *Atmospheric Measurement Techniques*, 2019, 12(1): 169-209.
- [17] Ångström A. The parameters of atmospheric turbidity [J]. *Tellus*, 1964, 16(1): 64-75.
- [18] Chen J M, Cihlar J. Retrieving leaf area index of boreal conifer forests using Landsat TM images [J]. *Remote Sensing of Environment*, 1996, 55(2): 153-162.
- [19] Rumelhart D E, Hinton G E, Williams R J. Learning representations by back-propagating errors [J]. *Nature*, 1986, 323(6088): 533-536.
- [20] Fletcher D, Goss E. Forecasting with neural networks [J]. *Information & Management*, 1993, 24 (3): 159-167.
- [21] Rodriguez J D, Perez A, Lozano J A. Sensitivity analysis of  $k$ -fold cross validation in prediction error estimation [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2010, 32 (3):



- 569-575.
- [22] Li T W, Shen H F, Zeng C, et al. Point-surface fusion of station measurements and satellite observations for mapping  $PM_{2.5}$  distribution in China: methods and assessment [ J ]. Atmospheric Environment, 2017, 152: 477-489.
- [23] Sun L, Zhou X Y, Wei J, et al. A new cloud detection method supported by GlobeLand30 data set [J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2018, 11 (10): 3628-3645.