

2014—2023年东亚地区沙尘气溶胶质量浓度再分析数据集

金建炳^①,李德昊^①,庞米杰^②,程喆琪^①,徐灿杰^①,廖宏^{①*}

① 南京信息工程大学 环境科学与工程学院/气候系统预测与变化应对全国重点实验室/江苏省大气环境监测与污染控制高技术研究重点实验室/江苏省大气环境与装备技术协同创新中心,江苏 南京 210044;

② 代尔夫特理工大学 应用数学学院,荷兰 代尔夫特 2628 CD

* 联系人, E-mail: hongliao@nuist.edu.cn

2025-11-02 收稿, 2025-12-15 接受

国家重点研发计划政府间国际科技创新合作重点专项(2024YFE0113700)

摘要 沙尘是中国北方典型的灾害天气。构建长时间尺度东亚地区高分辨率沙尘气溶胶质量浓度再分析数据集,是深化理解中国沙尘天气发生机理和提升多尺度预报水平的数据基础。受到风蚀起沙过程参数化方案、长距离输送误差等限制,当前沙尘模拟结果存在较大不确定性。鉴于此,本研究在前期开发的沙尘同化系统基础上,集成地面PM₁₀质量浓度、卫星气溶胶光学厚度(aerosol optical depth, AOD)观测非沙尘组分偏差校正技术,以及适用于沙尘气溶胶强度、位置误差协同校正的有效时刻偏移卡尔曼滤波同化算法(valid time shift ensemble Kalman filter, VTS-EnKF),建立了10 a(2014—2023年)东亚地区春季(3—5月)逐3 h的沙尘气溶胶三维质量浓度再分析数据集,分辨率为0.25°×0.25°。在此基础上,分析了所建立的再分析数据集相较于MERRA-2(modern-era retrospective analysis for research and applications version 2)沙尘再分析产品的优势,同时讨论了过去10 a东亚地区春季沙尘天气的月、年际变化趋势。

关键词

沙尘;
再分析数据;
数据同化

沙尘是中国北方典型的灾害天气。阿拉善-河西走廊、腾格里沙漠及蒙古国南戈壁沙漠是东亚主要沙源地,年释放总量超800 Mt(张小曳,2001)。春季由西伯利亚高压与蒙古气旋的协同作用导致的近地面大风,配合强烈垂直对流,形成了典型的“搬运-悬浮-传输”三维动力结构,是引起沙尘天气的主要原因(Yin et al.,2022)。在全球变暖背景下,蒙古高原春季降水大幅减少,蒙古国土地超75%不同程度荒漠化,从而加剧了跨境传输到中国的沙尘天气(Yuan et al.,2024)。以2023年春季为例,中国北方共出现17次沙尘天气,其中5次达沙尘暴及以上强度。其中3月19日强沙尘暴影响中国20个省份,华北局部地区PM₁₀质量浓度小时峰值超过9 000 μg·m⁻³,给人民健康、农业、交通等造成了严重危害(尹志聪等,2023),并入选“2023年国内十大天气气候事件”。因此,迫切需要构建沙尘天气历

史数据库,深化理解沙尘天气发生机理和提升预报水平,为我国人民生命财产安全和中蒙两国的“一带一路”经济建设提供精准的气象服务支撑。

大气模式是当前模拟起沙排放、沙尘区域输送、沉降等过程,解析沙源地植被生态、天气系统等对沙尘天气形成影响机制,量化评估沙尘气溶胶时空动态分布及其对生态与人民健康影响的关键工具。然而受沙尘排放过程模拟不确定性的限制(Escribano et al.,2016;Di Tomaso et al.,2017),以及长距离输送过程位置误差等限制(Jin et al.,2021,Pang,et al.,2024),目前各个模式在模拟沙尘天气发生时间、频率、沙尘浓度等方面仍然存在较大误差,沙尘浓度与实际观测差异甚至达到了两个数量级(Niu et al.,2008;Huneeus et al.,2011;Tian et al.,2021;Jin et al.,2022)。

同时大气监测技术的迅猛发展,为沙尘气溶胶

引用格式: 金建炳,李德昊,庞米杰,等,2026.2014—2023年东亚地区沙尘气溶胶质量浓度再分析数据集[J].大气科学学报,49(1):179-195.

Jin J B, Li D H, Pang M J, et al., 2026. A decadal dust aerosol mass concentration reanalysis over East Asia during 2014—2023[J]. Trans Atmos Sci, 49(1): 179-195. DOI: 10.13878/j.cnki.dqkxxb.20251102008. (in Chinese).

等大气成分的研究提供了重要基础。中国自2013年起,建立起了覆盖全国范围的地面空气质量监测网络,其提供的逐小时 PM_{10} 观测为研究东亚沙尘天气提供了宝贵资料。另一方面,卫星遥感观测技术具备监测范围广、不受区域地形限制、感知信息多样性等特点,为沙尘气溶胶提供了多层、立体、多角度、全方位的观测信息(张鹏等,2018)。21世纪初以来,美国国家航空航天局组建了由Aqua、Terra等极轨卫星构成的午后星座(A-Train)地球观测系统,其Moderate Resolution Imaging Spectroradiometer(MODIS)气溶胶光学厚度(aerosol optical depth, AOD)观测资料广泛用于中国及全球的沙尘、人为源气溶胶的研究。

基于贝叶斯概率理论的数据同化方法为融合沙尘观测资料 and 先验信息(沙尘模式结果),形成时空覆盖连续、精确性高的沙尘再分析数据集提供了契机。国内外已开展了大量沙尘气溶胶数据同化的研究。其中一部分研究旨在通过同化卫星观测数据直接优化模式初始场(即沙尘质量浓度),提高沙尘预报水平。例如,中国气象科学研究院团队通过耦合三维变分数据同化和沙尘模式,构建了中国在线沙尘数值预报系统,通过同化地面能见度与风云二号卫星观测资料,对沙尘模式初值进行实时修正,大大提高了沙尘预测水平(Gong and Zhang, 2008; Niu et al., 2008; Wang et al., 2008; Zhou et al., 2008)。中国科学院大气物理研究所团队通过集合卡尔曼滤波同化地面 PM_{10} 质量浓度观测数据,同样对沙尘模式初始场进行了订正,实现了对沙尘气溶胶强度、区域输送的准确预报(Lin et al., 2008a, 2008b)。另有少数学者围绕个别沙尘暴案例,利用数据同化实现了沙尘排放反演和浓度场的协同优化。例如Yumimoto et al. (2008)和Di Tomaso et al. (2017)使用MODIS AOD分别对非洲和东亚的一次沙尘天气中的排放进行了优化。相较于其他局地人为源大气污染物数据同化,影响中国北方地区的沙尘因其观测偏差大(气溶胶非沙尘组分影响)、观测数据覆盖率低、长距离传输模拟误差不可避免、沙源地排放时空变化复杂度高限制,其沙尘数据同化存在较大挑战及提升空间。作者前期聚焦东亚沙尘,自主开发了基于四维变分数据同化的沙尘气溶胶排放反演与同化预报系统(Jin et al., 2018),且针对上述沙尘同化瓶颈设计了多项关键技术:1)设计了AOD质量控制(Jin et al., 2019b)、观测偏差(非沙尘组分)订正(Jin et al., 2019a)等前处理算法,大大提高

了卫星AOD、地面 PM_{10} 质量浓度等观测数据对沙尘气溶胶的精确定量,有效避免了同化分析过调整;2)编写了LOTOS-EUROS大气化学模式伴随算子,实现了沙尘气溶胶来源不同时空位置、不同排放部门的快速追踪,相比于传统扰动法更高效(Jin et al., 2020);3)提出了图形畸变数据同化(grid-distortion data assimilation),从图像处理角度出发,使用较低计算成本,实现了沙尘气溶胶强度、位置误差(因长距离输送引起)的可协同优化(Jin et al., 2021);4)设计了邻时重采样同化算法(VTS-EnKF),优化了沙尘背景协方差对排放、输送等多源误差的代表性(Pang et al., 2024);5)研究了东亚沙尘跨境传输中传统AOD、地面质量浓度观测同化的不足,以及垂直廓线资料同化的重要性,设计了垂直局地化方案(Pang et al., 2025)。凭借上述技术,作者通过同化卫星MODIS AOD、地面 PM_{10} 质量浓度等多源星地观测,对2015、2017、2021年春季多场超强沙尘天气实现了精细化模拟预报(Jin et al., 2022)。上述沙尘数据同化工作为获取高质量东亚沙尘排放再分析打下了基础,然而,所有研究均针对个别沙尘暴案例,未形成长时间尺度东亚地区高分辨率沙尘气溶胶再分析数据集。

鉴于此,本研究将在前期开发的沙尘同化系统基础上,集成地面 PM_{10} 质量浓度、卫星AOD观测非沙尘组分偏差校正技术,以及适用于沙尘气溶胶强度、位置误差协同校正的有效时刻偏移卡尔曼滤波同化算法(VTS-EnKF),建立10 a(2014—2023年)东亚地区春季(3—5月)逐3 h的沙尘气溶胶三维质量浓度再分析数据集,分辨率为 $0.25^\circ \times 0.25^\circ$ 。

1 数据与方法

1.1 地面观测与非沙尘组分偏差处理

截至目前,中国生态环境部大气环境监测总站已经建立了一个由1800多个地面站组成的空气质量监测网络,覆盖全国各地(图1b),提供包括 PM_{10} 在内的逐小时地面颗粒物质量浓度数据。这些 PM_{10} 观测数据因其高时空分辨率和较高的准确性,已成为沙尘测量的重要数据源。然而,由于 PM_{10} 质量浓度不仅包括沙尘气溶胶,还包括其他非沙尘来源粗颗粒物组分,所以不能直接量化沙尘质量浓度。这些非沙尘颗粒物来源包括交通、能源、工业等部门直接排放,以及二次无机有机气溶胶等的贡献。数据同化的一个基本假设是观测是无偏差的,然而在现实中观测中的偏差往往是不可避免的,以本研究

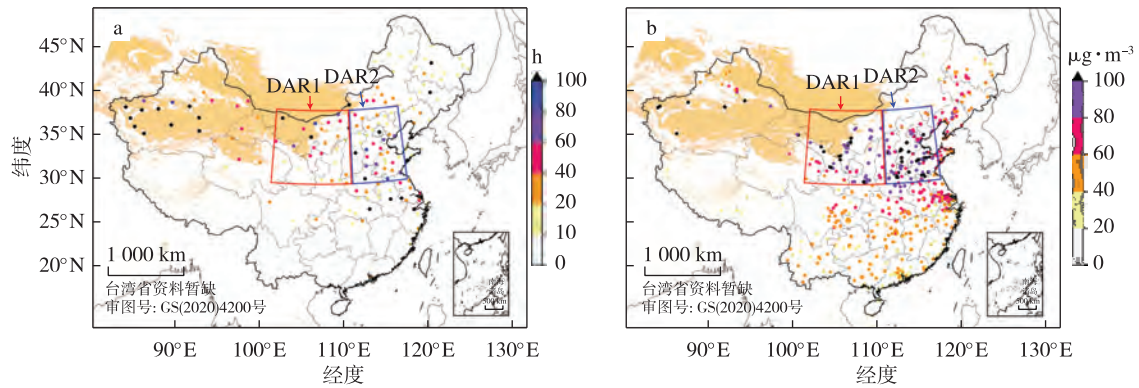


图1 地面气象观测站统计得到的2014—2023年春季总沙尘持续时长(a;单位:h);地面空气质量观测站统计得到的2014—2023年春季地面 PM_{10} 平均质量浓度(b;单位: $\mu g \cdot m^{-3}$)(红色框为1号沙尘影响区域(DAR1),蓝色框为2号沙尘影响区域(DAR2))

Fig.1 (a) Total spring dust duration (units:h) from 2014 to 2023 based on surface meteorological station observations;(b) mean surface PM_{10} mass concentration (units: $\mu g \cdot m^{-3}$) in spring from 2014 to 2023 based on surface air-quality monitoring stations.The red box indicates Dust-Affected Region 1 (DAR1),and the blue box indicates Dust-Affected Region 2 (DAR2)

的 PM_{10} 质量浓度观测为例,其是对所有空气动力学直径小于 $10 \mu m$ 的气溶胶质量浓度的测定,而非单一的沙尘颗粒物。如果将 PM_{10} 质量浓度直接同化到系统中,特别是当大部分 PM_{10} 来源是非沙尘颗粒时,沙尘同化系统将迅速发散(Jin et al, 2019b; 陈耀登等, 2021; 赵新宇等, 2023; 曹皓东等, 2024)。

本研究采用MLP(multi-layer perceptron)机器学习方法模拟了人为源粗颗粒物贡献,从而可以进一步实现在沙尘期间对 PM_{10} 质量浓度中沙尘气溶胶的定量解析。

$$y_{\text{non-dust}PM_{10}} = F(x_{SO_2}, x_{NO_x}, x_{CO}, x_{LE\text{-nondust-}PM_{2.5}}, x_{LE\text{-nondust-}PM_{10}}, W) \quad (1)$$

其中: $y_{\text{non-dust}PM_{10}}$ 为 PM_{10} 中非沙尘组分质量浓度; x_{SO_2} 、 x_{NO_x} 、 x_{CO} 为其他污染物质量浓度观测值; $x_{LE\text{-nondust-}PM_{2.5}}$ 、 $x_{LE\text{-nondust-}PM_{10}}$ 代表人为源污染物模式模拟值; W 为本地气象因子。

为了提高 $y_{\text{non-dust}PM_{10}}$ 模拟准确性,首先需要甄别出不同地区非沙尘时段,用于上述模型的训练。本研究综合使用美国国家气候数据中心(National Climatic Data Center, NCDC)提供的国家气象站的观测数据与中国地面空气质量监测站的地面 $PM_{2.5}$ 、 PM_{10} 等的相关数据,通过以下两个条件识别每个站内的沙尘时段表:

- 1) 当 PM_{10} 质量浓度观测值超过月均值的5倍,或者 $PM_{2.5}$ 与 PM_{10} 质量浓度的比值小于 $1/4$ 时;
- 2) 能见度小于3 km,相对湿度小于80%,风速大于 $3 m \cdot s^{-1}$,没有降水。

上述两个条件满足任一个条件即标注为“潜在”

的沙尘时段。在本研究中,2014—2023年3—5月所有非沙尘时段数据都被用于上述MLP机器学习模型的训练,从而获得每个站点的非沙尘组分动态模型,将其代入沙尘时段的 PM_{10} 质量浓度观测,从而得到地面站点去偏差后的沙尘质量浓度观测值。

$$y_{\text{dust}} = y_{PM_{10}} - y_{\text{non-dust}PM_{10}} \quad (2)$$

其中: y_{dust} 为去偏差后沙尘质量浓度; $y_{PM_{10}}$ 为 PM_{10} 质量浓度观测值。

在同化分析过程中,观测协方差矩阵同样是重要的输入。本研究假设观测不确定性包括 PM_{10} 数据和非沙尘 PM_{10} 偏差校正的不确定性,使用经验公式(2)来量化 y_{dust} 中的表征误差 σ_i^{PM} 。这里选择 $200 \mu g \cdot m^{-3}$ 作为 σ_i^{PM} 的最小阈值,主要是为了防止后验过于接近低值 PM_{10} 质量浓度观测数据,从而导致后验结果失真。

$$\sigma_i^{PM} = \begin{cases} 200, & y_{\text{dust}} \leq 200; \\ (y_{\text{dust}} - 200) \times 0.1 + 200, & y_{\text{dust}} > 200. \end{cases} \quad (3)$$

1.2 沙尘气溶胶光学厚度观测

气溶胶光学厚度常用于衡量大气中悬浮颗粒物(如气溶胶)对可见光的吸收和散射能力。AOD数据可以提供有关大气气溶胶的空间分布和浓度信息,这对研究气候变化、气象预测、空气质量监测以及评估地表辐射平衡等具有非常重要的作用。

本研究采用了中分辨率成像光谱辐射计(MODIS)与可见光/红外光影像辐射计(visible infrared imaging radiometer suite, VIIRS)的AOD数据。MODIS是美国国家航空航天局地球观测系统(Earth observation system, EOS)的基本仪器之一,

搭载在 Terra 和 Aqua 卫星上,能够测量多波长的反射率 (Xie et al., 2017)。VIIRS 主要用于观测和监测地球上的自然环境和气候变化。通过测量可见光和红外辐射,VIIRS 能够获取高分辨率的地表影像和大气数据。本研究使用了 MODIS 中 L2 级产品的 AOD 数据,具体参数如表 1 所示。

表 1 MODIS 与 VIIRS 卫星主要参数

Table 1 Key parameters of the MODIS and VIIRS satellite instruments

辐射计	卫星	空间分辨率	时间分辨率	过境时间 (当地时间)
MODIS	Terra	10 km×10 km	每日	10:30
	Aqua			13:30
VIIRS	Suomi NPP	6 km×6 km	每日	13:30

针对卫星数据,选取东亚地区 (70°~140° E, 15°~50° N) 的 MODIS 和 VIIRS 的数据产品,并对卫星观测点进行质量控制,剔除受天气条件、观测仪器误差等情况影响而产生的低质量的观测点,即质量标志小于 2 的像素点。

由于气溶胶光学厚度 τ_a 代表了粒子的光学厚度总和,所以需要通过以 Ångström 指数 (α) 为核心的经验函数,计算细颗粒组分比例 F_{sm} ,从而去除 AOD 中的非沙尘组分 (Anderson et al., 2005),获得沙尘气溶胶光学厚度 τ_d (dust optical depth, DOD)。

$$F_{sm} = -0.0512\alpha^2 + 0.5089\alpha + 0.02, \quad (4)$$

$$\tau_d = \tau_a \times (1 - F_{sm}). \quad (5)$$

Ångström 指数用于描述颗粒物的粒径大小分布情况,大于 2 表明为较小颗粒的气溶胶,小于 1 表明为大颗粒组分,如海盐和尘土。设置小于 1 的 Ångström 指数为阈值,剔除由细颗粒物主导的观测点,以进一步提高数据的准确性与代表性。

最后,定义经偏差校正后的沙尘质量浓度的不确定性,以用于随后的同化分析。与地面观测类似,此处的不确定性被假设为联合不确定性,包括 AOD 数据和非沙尘 AOD 偏差校正的不确定性,使用经验公式(5)来量化 DOD 中的表征误差 σ_i^{DOD} 。这里选择 0.1 作为 σ_i^{DOD} 的最小阈值,主要是为了防止后验过于接近低值 DOD 观测数据,从而导致同化系统失真。

$$\sigma_i^{DOD} = \begin{cases} 0.1, & \tau_d \leq 0.1; \\ (\tau_d - 0.1) \times 0.3 + 0.1, & \tau_d > 0.1. \end{cases} \quad (6)$$

1.3 LOTOS-EUROS 模式

本研究使用 LOTOS-EUROS 区域大气化学传

输模式来模拟东亚地区的沙尘气溶胶。LOTOS-EUROS 模式是由荷兰多个研究机构联合开发的,已经被大量应用于科学研究、政策制定和空气质量预报等 (Timmermans et al., 2017; Brasseur et al., 2019)。此外,它还被世界气象组织沙尘暴预警和评估系统采纳,提供北非-中东-欧洲 (NA-ME-E) 地区实时沙尘短期预报。

风蚀起沙是一个复杂的过程。它涉及到土壤颗粒 (大小分布、沙质/粘土/壤土比例)、地表状态 (植被覆盖、表面粗糙度、土壤湿度和地形) 以及气象条件 (摩擦速度) 的贡献。它是沙尘排放模型的重要组成部分,也是天然气溶胶的重要来源。模式中采用的沙尘起沙方案为 Zender03, 沙尘排放通量 f 主要受地面风速控制,通过摩擦速度 u_* 进行量化,并受地表、土壤特性的影响。只有当摩擦速度超过起沙阈值时,沙尘排放才会开始。土壤可侵蚀性通过尺寸相关的摩擦速度阈值来量化,该阈值被定义为启动土壤颗粒运动所需的最小摩擦速度。LOTOS-EUROS 使用了 5 种粒径来模拟不同大小的沙尘颗粒 (表 2)。

表 2 沙尘组分分类与粒径范围

Table 2 Classification of dust composition and corresponding particle-size ranges

沙尘分类	dust_ff	dust_f	dust_ccc	dust_cc	dust_c
粒径范围/ μm	(0.01, 1]	(1, 2.5]	(2.5, 4]	(4, 7]	(7, 10]

LOTOS-EUROS 沙尘模式中风蚀起沙参数方案等配置可详见 Jin et al. (2019a, 2019b)。

模式由欧洲中期天气预报中心 (European Centre for Medium-Range Weather Forecasts, ECMWF) 业务化运行的 3 h 间隔气象预报产品驱动,其原始空间分辨率约为 7 km。为适配本模式的网格配置,ECMWF 预报数据通过空间插值方法被重采样至模式所采用的 0.25°×0.25° 经纬度网格上,从而提供高时空一致性的初始和边界条件,保障模拟过程的准确性和稳定性。

1.4 同化再分析系统框架

研究使用的沙尘同化系统主要由两部分构成 (图 2), 其一是前文所述的 LOTOS-EUROS 模式,用于正向模拟。因为同化系统所用的基于集合方法的特性,使得同化算法可以独立于模式之外,所以模式只是连接在系统上,对模式只进行了最小程度的修改,任何其他模式都可以被替换上。另一部分是同化模块,包含了 EnKF、VTS-EnKF 等同化算法。

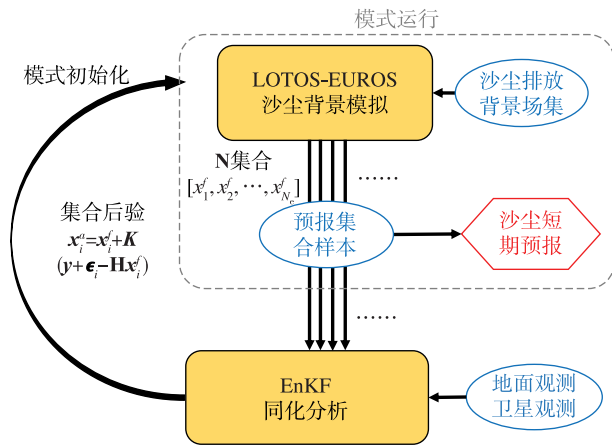


图2 沙尘同化系统框架图(引自 Pang et al.(2023))

Fig. 2 Schematic framework of the dust assimilation system(cited from Pang et al.(2023))

该模块使用 Python 编写,被嵌入同化工具 PyFilter 中,同化算法具有易于编写、修改的特点。整个同化系统由 PyFilter 统一控制流程,关键参数可以直接在配置文件中修改,它虽然无法获得内置同化方法系统那样的运行效率,但具有易于修改、可拓展性高等优点(Pang et al., 2023)。

整个同化系统的运行流程如下:一开始, N_c 个模式集合由被扰动的排放场驱动,进行初始预测,产生的预测结果由集合平均得到,当观测数据可以获得,就进入同化流程,输入观测矩阵与模式集合状态矩阵,进行同化分析,得到集合后验,再将集合后验输入回模式,重启模式进行预测,由此往复直到循环结束。

1.4.1 传统集合卡尔曼滤波(EnKF)

集合卡尔曼滤波(ensemble Kalman filter, EnKF)是最常见的同化方法,其利用有限的集合、近似高维的背景协方差(Bannister, 2017),为处理高维大气模式动力学提供了可行的解决方法。

EnKF 同化工作过程中,当前时刻沙尘质量浓度状态的集合个体 \mathbf{x}_i^f 由 LOTOS-EUROS 模拟得来,模式算子 \mathbf{M} 将上一时刻的状态量 $\mathbf{x}_{i,t-1}^a$ 积分至当前时刻, \mathbf{X}^f 代表当前所有集合的状态量矩阵:

$$\mathbf{x}_i^f = \mathbf{M}(\mathbf{x}_{i,t-1}^a), \quad (7)$$

$$\mathbf{X}^f = [\mathbf{x}_1^f, \mathbf{x}_2^f, \dots, \mathbf{x}_{N_c}^f], \quad (8)$$

$$\bar{\mathbf{x}}^f = \frac{1}{N_c} \sum_{i=1}^{N_c} \mathbf{x}_i^f. \quad (9)$$

其中: N_c 为集合数量。随后由集合状态量计算出集合平均值 $\bar{\mathbf{x}}^f$,将每个集合与 $\bar{\mathbf{x}}^f$ 相减可以得到集合扰动矩阵 \mathbf{X}^f ,由此可以得出集合近似情况下的背景协

方差矩阵 \mathbf{P}^f :

$$\mathbf{X}^f = [\mathbf{x}_1^f - \bar{\mathbf{x}}^f, \mathbf{x}_2^f - \bar{\mathbf{x}}^f, \dots, \mathbf{x}_{N_c}^f - \bar{\mathbf{x}}^f], \quad (10)$$

$$\mathbf{P}^f = \frac{1}{N_c - 1} \sum_{i=1}^{N_c} (\mathbf{x}_i^f - \bar{\mathbf{x}}^f) (\mathbf{x}_i^f - \bar{\mathbf{x}}^f)^T = \frac{1}{N_c - 1} \mathbf{X}^f \mathbf{X}^{fT}. \quad (11)$$

权重矩阵 \mathbf{K} 一般被称作卡尔曼增益,它负责调配模式与观测之间的权重, \mathbf{H} 是观测算子,它将模式状态量映射到观测空间中, \mathbf{y} 是观测向量,存储了所有的 PM_{10} 、AOD 等观测数据,经过 \mathbf{H} 的映射, \mathbf{y} 与 $\mathbf{H}\mathbf{x}_i^f$ 是一一对应的关系, \mathbf{O} 是观测误差矩阵。

$$\mathbf{K} = \mathbf{P}^f \mathbf{H}^T (\mathbf{H} \mathbf{P}^f \mathbf{H}^T + \mathbf{O})^{-1}. \quad (12)$$

后验集合成员 \mathbf{x}_i^a 可以通过下式更新:

$$\mathbf{x}_i^a = \mathbf{x}_i^f + \mathbf{K}(\mathbf{y} + \boldsymbol{\epsilon}_i - \mathbf{H}\mathbf{x}_i^f). \quad (13)$$

其中, $\boldsymbol{\epsilon}_i$ 为观测误差扰动项,服从均值为 0、方差为对应观测误差的正态分布,用于维持集合的扩散程度,防止集合收敛导致同化算法失效(Burgers et al., 1998)。

受到集合样本数量的限制,即便是相距甚远的观测也会对本地的状态有影响,即虚假的协相关性(Anderson, 2012),为了减弱该影响,需要引入局地化技术,详见 Pang et al.(2023)。

1.4.2 VTS 同化算法与再分析系统参数设计

在沙尘模拟中,一个大问题是经过远距离传输后,观测结果与模式在空间位置的出现不匹配。除了沙尘强度的差异以外,模式模拟的沙尘到达和离开的时间也与实际存在差异。沙尘强度和位置是评估沙尘预报的两个关键特征。前者代表了实际的沙尘质量浓度大小,而后者则揭示了在某一时刻何处受到沙尘的影响。后者误差在本研究中被称作“位置误差”,通常出现在沙尘的远距离传输后。引起沙尘位置误差的因素包括:模式中被忽略掉的某些物理过程、模式分辨率(Ravela et al., 2007)以及气象场和排放的不确定性。沙尘位置的不确定性很难量化,前期其他沙尘数据同化研究均只关注了沙尘质量浓度优化,尚无研究涉及位置误差造成的观测结果与模拟之间的不匹配问题。

如图 3a 所示,当存在位置误差时,传统 EnKF 将无法有效地同化沙尘观测,会使得一部分后验完全偏向先验模式结果,造成与实际的完全不符。本研究为更好地应对位置误差问题,将 VTS 与前述的传统 EnKF 耦合(VTS-EnKF)。在同化分析时,不仅使用扰动排放的原始集合成员计算背景协方差矩阵,而且同时使用相邻时间点 $t-k$ 与 $t+k$ 的集合成

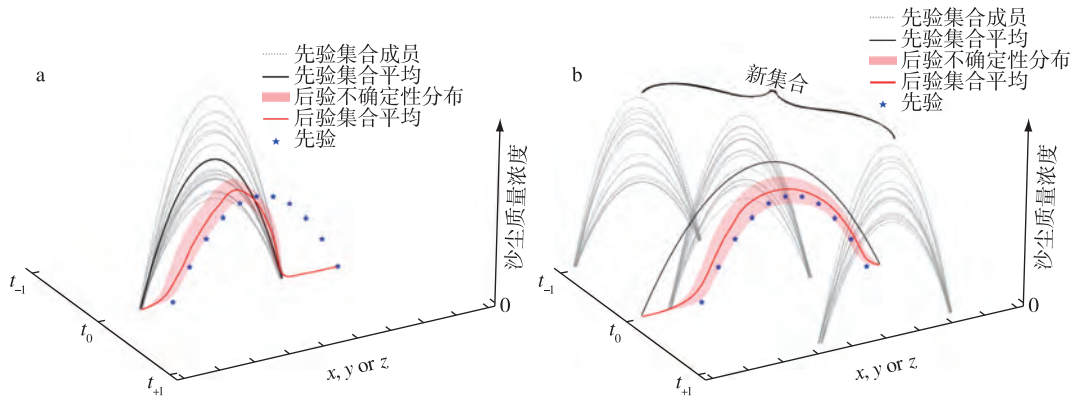


图3 EnKF(a)、VTS-EnKF(b)在有位置误差情况下的同化分析效果对比(引自 Pang et al.(2024))

Fig.3 Analysis strategy of (a) the EnKF and (b) the VTS-EnKF methods under position-error conditions(cited from Pang et al.(2024))

员,本研究中 $k=2$, 其对性能的影响在 Pang et al. (2024) 中已有详细探究,时间过小时会导致同化不稳定性加剧,过大的时间间隔会导致过多的沙尘团出现。这些新的集合代表了沙尘位置可能分布范围,从而表征了沙尘的位置误差。上述两种集合成员构建的协方差矩阵能同时量化沙尘强度和位置误差,无需对观测、气象场或其他物理参数进行额外处理(图3b)。形成的新的集合成员 $\mathbf{X}^{f,new}$ 用于估算沙尘的背景协方差矩阵:

$$\mathbf{X}^{f,new} = \begin{bmatrix} x_{t-k}^{f,1}, x_{t-k}^{f,2}, \dots, x_{t-k}^{f,N_e}, x_t^{f,1}, x_t^{f,2}, \dots, x_t^{f,N_e}, x_{t+k}^{f,1}, \\ x_{t+k}^{f,2}, \dots, x_{t+k}^{f,N_e} \end{bmatrix}. \quad (14)$$

VTS-EnKF 在沙尘同化应用中相较于传统 EnKF 同化算法的优势,详见 Pang et al.(2024)。

2 结果分析与讨论

2.1 沙尘年际趋势

通过 VTS-EnKF 同化技术结合 LOTOS-EUROS 模式与经预处理后的地面、卫星观测,本研究获得了2014—2023年春季(3—5月)的逐3h的东亚沙尘再分析数据,分辨率为 $0.25^\circ \times 0.25^\circ$ 。

图4、图5为2014—2023年的春季平均沙尘先验地面质量浓度场、平均沙尘后验地面质量浓度场和分析增量分布,反映了沙尘再分析的大致分布特征。从沙尘先验场(图4a、图5a)可以看到,模式模拟的沙尘主要集中于新疆塔里木盆地地区与中国戈壁沙漠区域,蒙古戈壁区域亦有零散的分布,沙尘向南最远延伸至 30°N 。新疆的沙尘质量浓度达到 $2000 \mu\text{g} \cdot \text{m}^{-3}$,中国戈壁的沙尘质量浓度同样极高,达到 $1000 \mu\text{g} \cdot \text{m}^{-3}$,覆盖范围在2020年最小,在

2017与2018年最大。通过同化分析后,从后验沙尘质量浓度分布(图4b、图5b)可以看到,整体质量浓度场被降低。对照增量(图4c、图5c)可以看到广泛的负增量分布(约 $-300 \mu\text{g} \cdot \text{m}^{-3}$),表明先验模式模拟对沙尘质量浓度是高估的。尤其是中国戈壁沙漠地区,从增量图中可以发现明显的负增量,2017—2018年增量达到了 $-1500 \mu\text{g} \cdot \text{m}^{-3}$,说明这两年的先验沙尘场出现了显著的高估。对于新疆地区的沙尘,后验场的修正范围并不明显,除了2016、2019年可以看到明显的负增量,其余年份为 $-500 \sim -300 \mu\text{g} \cdot \text{m}^{-3}$,这主要是由于新疆及其周围地区地面观测站点相较于东部更加稀疏,故同化的作用不明显。

表3给出了2014—2023年年均、每月、每年的沙尘先验、后验平均质量浓度及其增量。由表3可以看出同化分析对沙尘先验场的修正程度。先验场的整体沙尘平均质量浓度为 $65.24 \mu\text{g} \cdot \text{m}^{-3}$,经过同化地面、卫星观测,整体的平均质量浓度降至 $39.99 \mu\text{g} \cdot \text{m}^{-3}$,增量达到 $-25 \mu\text{g} \cdot \text{m}^{-3}$,增幅为 -38.9% ,说明同化算法有效地修正了模式模拟对沙尘的高估。从各年先验场来看,2014年的年均质量浓度最高,达到了 $83.69 \mu\text{g} \cdot \text{m}^{-3}$,2018年的年均质量浓度与2014年相当,最低为2020年的 $48.43 \mu\text{g} \cdot \text{m}^{-3}$ 。后验场年均质量浓度同样显示2014年的年均质量浓度最高,达到 $54.45 \mu\text{g} \cdot \text{m}^{-3}$,降幅达到 $-29.23 \mu\text{g} \cdot \text{m}^{-3}$,与先验场不同的是后验场中的年均质量浓度最低的年份为2017与2019年,仅为约 $30 \mu\text{g} \cdot \text{m}^{-3}$ 。分析增量在2016、2017、2018年达到最大,约为 $-36 \mu\text{g} \cdot \text{m}^{-3}$,表明先验场显著高估,而通过同化算法,高估的趋势被有效遏制。

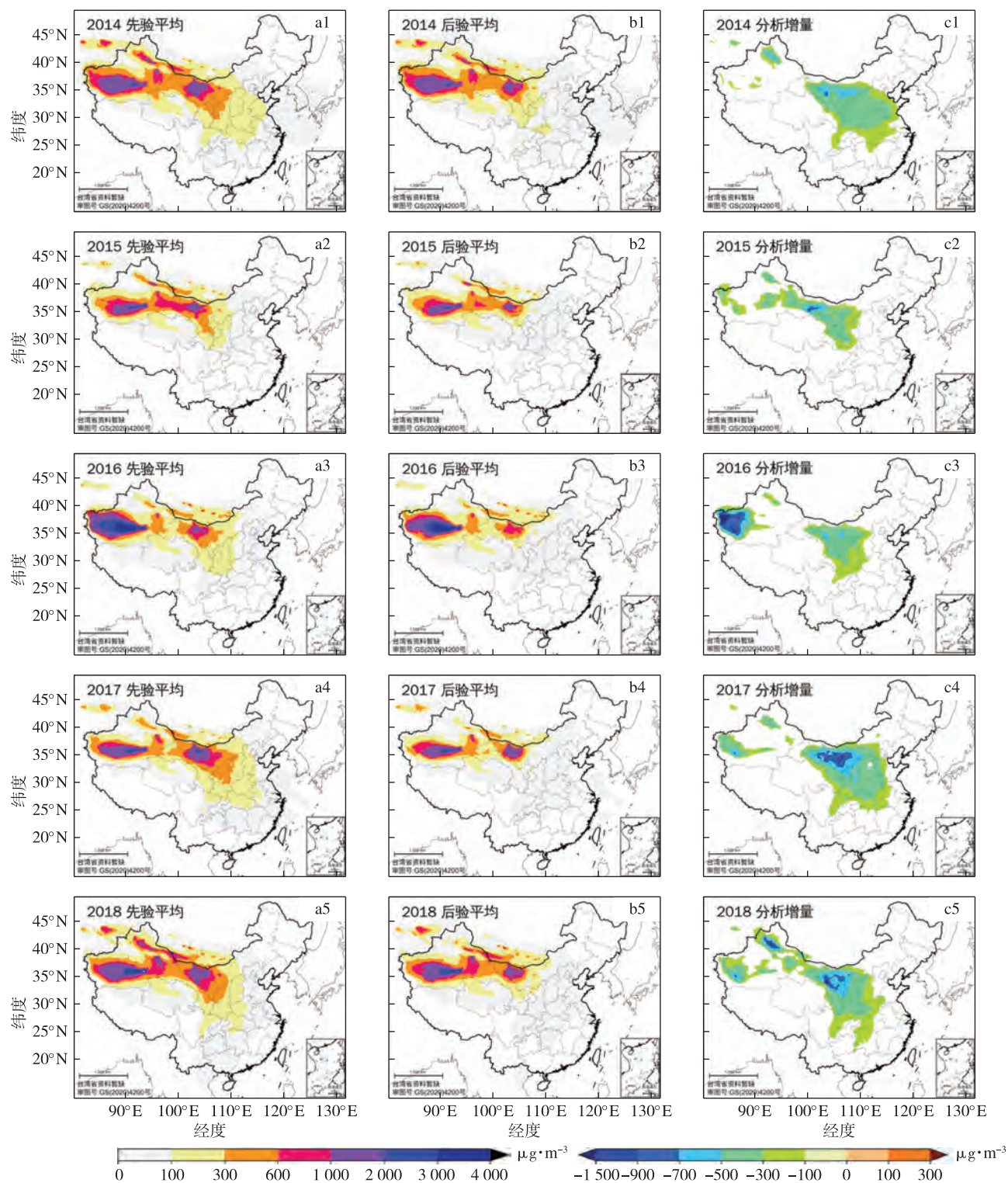


图4 2014—2018年的平均沙尘先验质量浓度场(a)、后验质量浓度场(b)、沙尘后验场减去对应先验场的增量分布(c) (单位: $\mu\text{g} \cdot \text{m}^{-3}$)

Fig.4 Mean dust mass concentration fields for 2014—2018: (a) prior mass concentration, (b) posterior mass concentration, and (c) analysis increment (posterior minus prior) (units: $\mu\text{g} \cdot \text{m}^{-3}$)

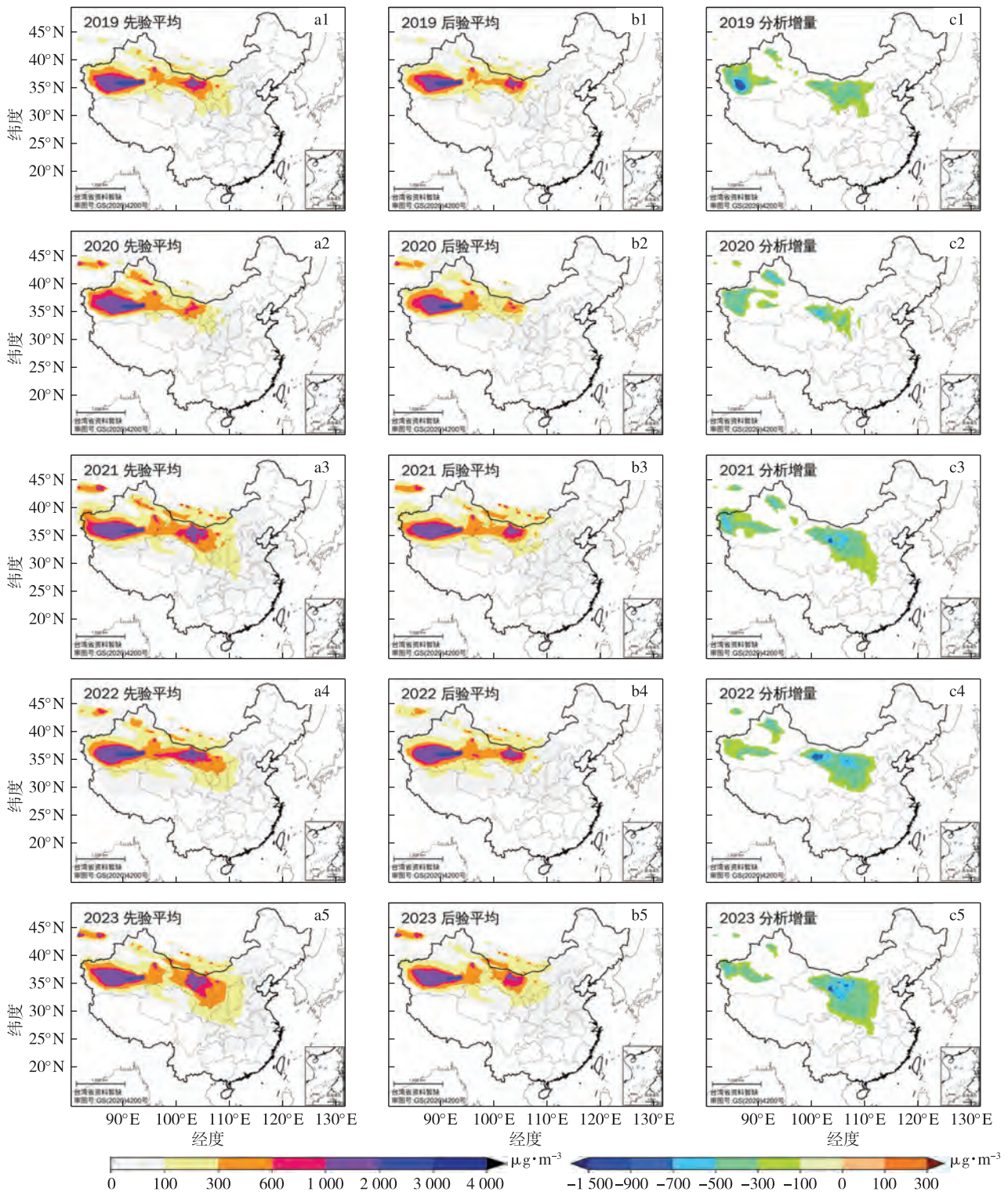


图5 2019—2023年的平均沙尘先验质量浓度场(a)、后验质量浓度场(b)、沙尘后验场减去对应先验场的增量分布(c) (单位: $\mu\text{g}\cdot\text{m}^{-3}$)

Fig.5 Mean dust mass concentration fields for 2019—2023: (a) prior mass concentration, (b) posterior mass concentration, and (c) analysis increment (posterior minus prior) (units: $\mu\text{g}\cdot\text{m}^{-3}$)

表 3 2014—2023 年先验、后验沙尘平均质量浓度及其分析增量

Table 3 Average dust prior and posterior mass concentrations and corresponding analysis increments for the entire study period, by month, and by year from 2014 to 2023

时期	$\mu\text{g} \cdot \text{m}^{-3}$		
	先验平均质量浓度	后验平均质量浓度	分析增量
2014—2023 年	65.24	39.99	-25.25
3 月	35.31	24.15	-11.15
4 月	68.96	42.96	-26.00
5 月	91.50	52.86	-38.64
2014 年	83.69	54.45	-29.23
2015 年	50.44	32.88	-17.56
2016 年	77.37	46.28	-36.34
2017 年	66.96	30.62	-36.98
2018 年	83.67	46.69	-36.98
2019 年	49.54	30.03	-19.51
2020 年	48.43	32.93	-15.50
2021 年	65.55	40.48	-25.07
2022 年	62.86	40.78	-22.09
2023 年	63.94	44.76	-19.18

前述基于所有时间步长的整个东亚空间平均值来评估沙尘质量浓度数据,这虽然提供了一个大致的误差评价框架,但可能忽略了那些持续时间短暂但影响深远的沙尘天气事件。因此,为了进一步提高评估的细致度和针对性,本研究也将重点放在具体区域内(DAR1和DAR2)的沙尘质量浓度的变化趋势上,以更加精确地评估数据集的可靠性。

图6为两个主要的受沙尘影响的区域(DAR1和DAR2)中沙尘先验场与后验场的平均质量浓度变化。由图6可见,先验沙尘场有效地再现了各次沙尘天气的暴发和消散过程,但是在质量浓度与持续时间的估计上,存在显著的高估现象,平均峰值质量浓度更是达到了 $2\,000\ \mu\text{g} \cdot \text{m}^{-3}$ 。通过同化过程,沙尘质量浓度的高估得到了修正,最突出的例证出现在2014年5月,当时模型在先验场中模拟了多次强度大、持续时间长的沙尘暴,而实际观测显示沙尘事件的强度较模拟预测更低,且持续时间较短(约1d)。此外,这种同化过程的改善在2018、2021和2023年也有良好的表现。

DAR1区域由于更接近沙尘的源地,受沙尘天气的影响较大,这导致该区域内的总体质量浓度明显高于DAR2区域。与此相对,DAR2区域受沙尘影响的程度更多地依赖于大气环流的输送作用。大部分时间内,DAR2区域的平均沙尘质量浓度较低,并仅在2017、2021和2023年的特定时期出现了质量浓度的短暂升高,这些时期正好对应多起强沙尘天气事件的发生,这些事件甚至还造成了遥远地区的影响(Jin et al., 2019a, 2019b, 2022; Chen et al., 2023)。这表明本沙尘数据集不仅在再现沙尘天气事件的强度与动态上具有可靠性,而且在对沙尘事件发生时间的捕捉方面也有着优异的表现。

表4为2014—2023年同化过程对沙尘模拟先验场误差的修正效果,通过均方根误差(root mean square error, RMSE)的对比,能够清晰地展示数据同化算法在降低误差方面的效果。

总体而言,先验场的整体RMSE高达 $256.67\ \mu\text{g} \cdot \text{m}^{-3}$,而通过同化步骤处理后,后验场质量浓度的RMSE降低至 $59.33\ \mu\text{g} \cdot \text{m}^{-3}$,降幅高达76.9%。巨大的降幅表明本研究使用的数据同化算法在减小沙尘模拟先验场误差方面是有效的。

深入观察各年份的RMSE可知,2017与2018年的RMSE值最高,分别达到了 393.32 与 $392.49\ \mu\text{g} \cdot \text{m}^{-3}$,表明该年份模式预测误差相对较大;2014年的预测误差最小,RMSE仅为 $108.98\ \mu\text{g} \cdot \text{m}^{-3}$ 。对照后验场的RMSE数据可知,每一年份的误差都有明显降低,所有年份的RMSE均低于 $82\ \mu\text{g} \cdot \text{m}^{-3}$,其中2014年的后验RMSE最为优异,降至 $15.69\ \mu\text{g} \cdot \text{m}^{-3}$ 。值得注意的是,最高RMSE发生在2017年,其分析增量达到了 $-334.32\ \mu\text{g} \cdot \text{m}^{-3}$,验证了数据同化算法对于先验场偏差的矫正能力,为沙尘天气预报及环境监测领域提供了更精确的工具和方法。

图7为2014—2023年沙尘先验场与后验场相对于地面 PM_{10} 观测的RMSE的动态变化情况。先验场的RMSE变化波动较大,通过折线图可以观察到其变化范围较大,峰值在历年中一度剧烈上升,超过了 $2\,000\ \mu\text{g} \cdot \text{m}^{-3}$ 。这一观察结果揭示了模型在模拟沙尘天气过程时的不足,特别是在估算沙尘的具体质量浓度值时,模型先验场表现出不准确性。而在整合了地面 PM_{10} 、卫星DOD观测数据之后,误差降低,RMSE数值在多数时间内稳定于 $200\ \mu\text{g} \cdot \text{m}^{-3}$ 以下。

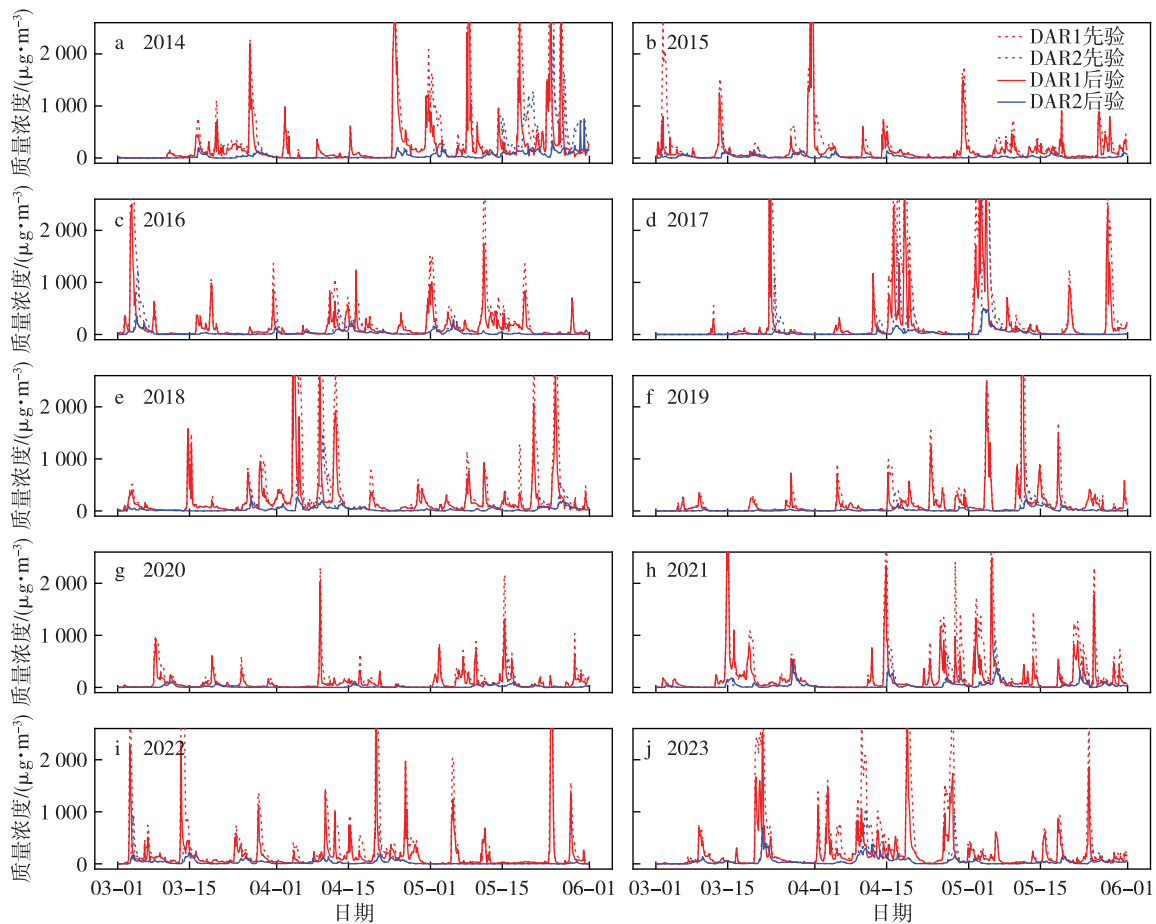


图6 2014—2023年(a—j)DAR1与DAR2区域内沙尘先验场与后验场的平均质量浓度变化曲线(单位: $\mu\text{g}\cdot\text{m}^{-3}$)

Fig.6 Time series of mean dust prior and posterior mass concentrations (units: $\mu\text{g}\cdot\text{m}^{-3}$) in DAR1 and DAR2 from (a—j) 2014 to 2023

表4 2014—2023年沙尘先验、后验平均质量浓度的RMSE及其分析增量

Table 4 RMSEs of dust prior and posterior mass concentrations, and their analysis increments, for the entire period, for March-April, and for each year from 2014 to 2023 $\mu\text{g}\cdot\text{m}^{-3}$

时期	先验平均质量浓度的RMSE	后验平均质量浓度的RMSE	分析增量
2014—2023年	256.67	59.33	-197.34
3月	176.49	66.63	-109.86
4月	271.63	59.36	-212.27
5月	316.63	52.33	-264.29
2014年	108.98	15.69	-93.29
2015年	234.85	66.92	-167.93
2016年	295.03	66.63	-228.40
2017年	393.32	59.01	-334.32
2018年	392.49	81.37	-311.12
2019年	182.34	55.24	-127.10
2020年	194.80	52.77	-142.03
2021年	249.24	72.23	-177.01
2022年	252.07	52.75	-199.33
2023年	263.58	70.71	-192.87

值得注意的是,在某些极端沙尘天气事件中,模型先验场的预测误差迅速蹿升,超过 $2\,000\ \mu\text{g}\cdot\text{m}^{-3}$,表明模型在这些时期对沙尘质量浓度的预估存在巨大的偏差。然而,在结合了地面和卫星观测数据的数据同化后,即使在这些极端天气情况下, RMSE 也被有效地控制在很低的水平,这表明数据同化方法在提升分析场准确性方面具有显著的优势。

2.2 沙尘月度趋势

从表3可以观察到,沙尘质量浓度在春季几个月中表现出随时间递增的趋势,这可能与春季沙尘活动的季节性增强有关。先验场和后验场均显示了这一特点,先验场的平均质量浓度从3月的 $35.31\ \mu\text{g}\cdot\text{m}^{-3}$ 增加至4月的 $68.96\ \mu\text{g}\cdot\text{m}^{-3}$,再到5月的 $91.50\ \mu\text{g}\cdot\text{m}^{-3}$ 。后验场的趋势与先验场相似,但数值较低,分别是 24.15 、 42.96 和 $52.86\ \mu\text{g}\cdot\text{m}^{-3}$ 。分析误差的逐月减少也反映了数据同化对沙尘模拟的积极影响。通过结合地面、卫星观测资料来更新先验场,后验场更接近实际观测情况,并在一定程度

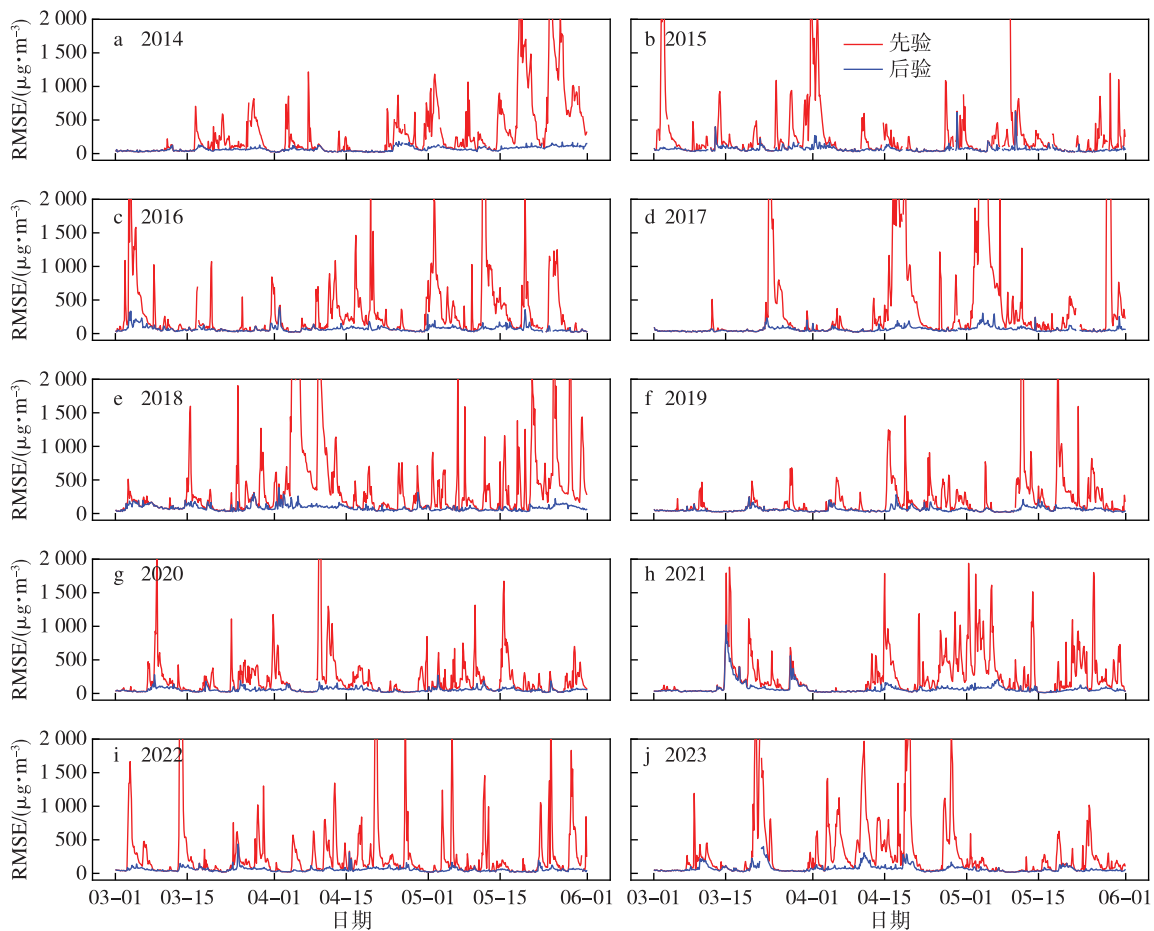


图7 2014—2023年(a—j)沙尘先验场、后验场与地面 PM_{10} 观测之间的质量浓度的均方根误差(单位: $\mu\text{g} \cdot \text{m}^{-3}$)

Fig.7 Time series of mass concentration RMSEs (units: $\mu\text{g} \cdot \text{m}^{-3}$) between surface PM_{10} observations and dust prior (red) and posterior (blue) estimates from (a—j) 2014 to 2023

上能够减少预测中的不确定性。表4中的RMSE变化也提供了有关沙尘质量浓度预测准确率的信息。准确率随月份增加呈现降低趋势,这可能是因为沙尘天气的不断增强导致了预测难度的提升。其中,5月的平均RMSE最高,达到 $316.63 \mu\text{g} \cdot \text{m}^{-3}$,表明先验场与实际之间的差异相对较大。然而经过数据同化后,后验场的RMSE显著降低,稳定在 $67 \mu\text{g} \cdot \text{m}^{-3}$ 以下。这表明数据同化明显提高了预测准确性。

表5所示的多年数据反映了在重点区域内沙尘质量浓度的月变化趋势,以及这种趋势可能在某些年份被打破的特殊情况。由前述可知,沙尘整体质量浓度在3月、4月和5月呈递增趋势,这与许多地区春季沙尘活动增强的季节因素有关。而在重点关注区域DAR1、DAR2内,可以发现不同的趋势。在DAR1内,2014、2015、2017与2019年显示了这种随时间递增的模式,但在某些年份如2018和2023年则表现出了不同的变化。其中特别引人关注的是

2023年,4月的沙尘质量浓度远高于3月和5月。这种异常的质量浓度峰值可能是受到了特定天气事件、沙尘源活动异常、气候波动,甚至可能是人为因素的影响,其具体原因需要更详细的分析。2020年各月的沙尘质量浓度同比最低,变化最为平稳,后验场维持在 $78 \mu\text{g} \cdot \text{m}^{-3}$ 以下。DAR2区域内的平均沙尘质量浓度远低于DAR1,且月度的变化趋势并不明显,这与季风活动导致的沙尘长距离输送有关。

2.3 独立数据集对比

首先对整体区域内的沙尘平均质量浓度进行比较。图8显示了研究区域内($70^{\circ} \sim 140^{\circ} \text{E}$, $15^{\circ} \sim 50^{\circ} \text{N}$)的沙尘后验场与MERRA-2沙尘数据集在所有网格点的平均质量浓度的对比散点图。从图中的分布可以清楚地观察到,在较低的质量浓度范围内,尤其是在 $75 \mu\text{g} \cdot \text{m}^{-3}$ 以下,MERRA-2数据集与后验场之间表现出较为显著的相关性。这表明在该质量浓度区间内,两个数据集之间的沙尘质量浓度估算是相对一致的。然而,散点图同时揭示了一个重要

表5 2014—2023年每年3、4、5月的DAR1与DAR2区域内的平均沙尘先验、后验质量浓度与分析增量

Table 5 Average dust prior and posterior mass concentrations and analysis increments in DAR1 and DAR2 for March, April, and May from 2014 to 2023 $\mu\text{g} \cdot \text{m}^{-3}$

年份	月份	先验平均质量浓度		后验平均质量浓度		分析增量	
		DAR1	DAR2	DAR1	DAR2	DAR1	DAR2
2014年	3	183.66	30.00	78.87	15.94	-104.79	-14.06
	4	245.48	19.02	169.62	13.31	-75.86	-5.71
	5	719.87	311.05	192.48	44.81	-527.39	-266.23
2015年	3	207.66	47.28	60.65	26.10	-147.01	-21.18
	4	167.76	22.19	90.88	12.77	-76.88	-9.42
	5	190.38	21.74	101.47	15.62	-88.91	-6.12
2016年	3	255.81	69.05	84.65	19.15	-171.16	-49.91
	4	205.58	62.78	79.83	23.72	-125.75	-39.07
	5	272.49	34.57	92.72	14.44	-179.76	-20.13
2017年	3	151.14	33.72	42.30	4.58	-108.83	-29.13
	4	375.59	97.50	76.66	13.61	-298.93	-83.88
	5	716.12	226.98	134.48	29.41	-581.63	-197.57
2018年	3	167.74	22.72	98.91	18.00	-68.83	-4.72
	4	676.60	120.92	152.42	32.17	-524.18	-88.75
	5	354.24	62.61	116.35	23.27	-237.90	-39.35
2019年	3	56.87	5.80	33.00	4.08	-23.86	-1.72
	4	173.27	29.96	80.65	11.44	-92.62	-18.52
	5	373.99	61.66	133.40	20.71	-240.59	-40.95
2020年	3	106.47	17.43	55.46	10.80	-51.01	-6.63
	4	87.54	7.07	42.52	4.75	-45.01	-2.33
	5	186.52	28.48	77.58	14.77	-108.94	-13.71
2021年	3	207.12	31.90	164.88	35.35	-42.24	3.44
	4	254.61	68.92	80.47	21.09	-174.13	-47.83
	5	387.94	85.13	112.64	33.83	-275.31	-51.29
2022年	3	300.13	67.12	113.61	31.96	-186.52	-35.16
	4	343.83	46.46	122.93	22.78	-220.89	-23.67
	5	196.84	3.84	81.23	2.38	-115.61	-1.45
2023年	3	254.94	48.06	114.44	39.22	-140.49	-8.85
	4	601.63	123.71	191.89	51.84	-409.74	-71.87
	5	184.13	41.61	89.56	19.88	-94.58	-21.74

的现象:随着质量浓度的增加,在更高的质量浓度区域,这种一致性开始显著下降。尤其在沙尘质量浓度较高的情况下,相比于后验场,MERRA-2对沙尘质量浓度的估算存在明显的低估倾向。这种系统性的低估在2014—2023年的长期研究中一直是一个持续的特点,表明MERRA-2数据整体低估了区域内的沙尘质量浓度。该发现对于改善沙尘天气的数

值预报和相关影响评估具有一定意义,表明本数据集在处理高质量浓度沙尘事件时,相较于主流的MERRA-2更具优势。

除了研究整体的特征,本文进一步对研究区内沙尘天气影响的区域性特征加以深入分析。图9和图10分别为DAR1区域和DAR2区域内的沙尘后验场与MERRA-2数据的平均质量浓度的对比。这

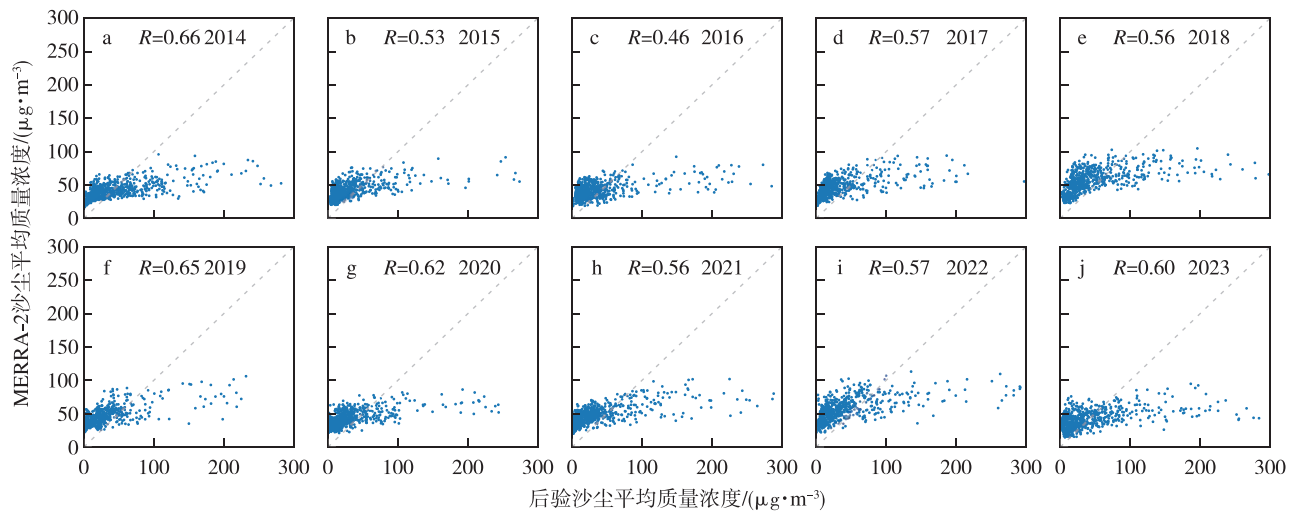


图 8 2014—2023 年(a—j) 沙尘后验场与 MERRA-2 数据在研究区域内所有格点的平均质量浓度对比散点图(单位: $\mu\text{g} \cdot \text{m}^{-3}$)

Fig.8 Scatter plots of domain-averaged dust mass concentrations comparing posterior estimates with MERRA-2 dust products for (a—j) 2014—2023 (units: $\mu\text{g} \cdot \text{m}^{-3}$)

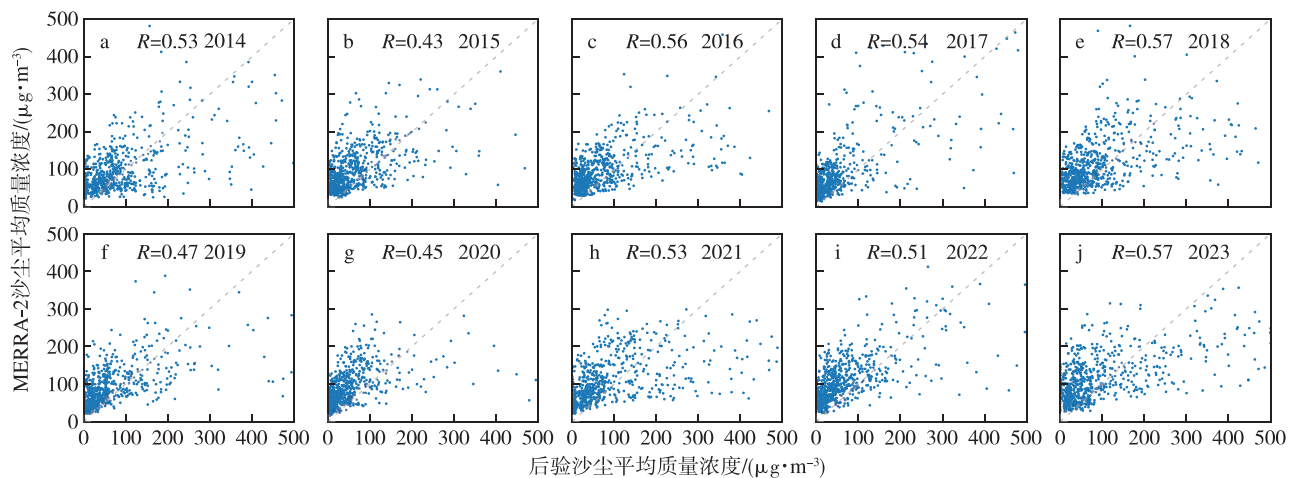


图 9 2014—2023 年(a—j)在 DAR1 区域内沙尘后验场与 MERRA-2 数据的平均质量浓度对比散点图(单位: $\mu\text{g} \cdot \text{m}^{-3}$)

Fig.9 Scatter plots of averaged dust mass concentrations over DAR1 comparing posterior estimates with MERRA-2 dust products for (a—j) 2014—2023 (units: $\mu\text{g} \cdot \text{m}^{-3}$)

两个子区域为研究范围内沙尘暴事件发生最为频繁、强度最大的区域,因此数值上要远大于整体区域的值。由图可见,DAR1 区域内的质量浓度主要集中在 $100 \mu\text{g} \cdot \text{m}^{-3}$ 附近,DAR2 区域的沙尘质量浓度则更多集中在 $50 \mu\text{g} \cdot \text{m}^{-3}$ 以下的较低值。DAR1 区域观测到的沙尘平均质量浓度普遍高于 DAR2。这一分布差异表明,相对于 DAR1 区域,DAR2 区域更易受沙尘天气的影响。

对沙尘后验场与 MERRA-2 数据的比较可知,两个区域内数据间的相关性表现出了明显的差异。在 DAR1 区域内,相关系数(R)大致稳定在 0.5。这表明尽管存在一些差距,但仍能观察到一定程度的

正向关联。在 DAR2 区域内,这种相关性的波动幅度更大,最低为 2014 年的 0.38,最高为 2017 年的 0.84。这种波动不仅凸显出在沙尘天气评估上的不确定性,而且表明 MERRA-2 数据在这些关键区域可能存在不可忽视的误差。

3 结论与讨论

沙尘是中国北方地区典型的灾害天气。构建长时间尺度东亚地区高分辨率沙尘气溶胶再分析数据集,对深化理解中国沙尘天气发生机理和提升沙尘天气预报水平具有重要意义。受到风蚀起沙过程参数化方案、长距离输送误差等限制,当前沙尘模拟结

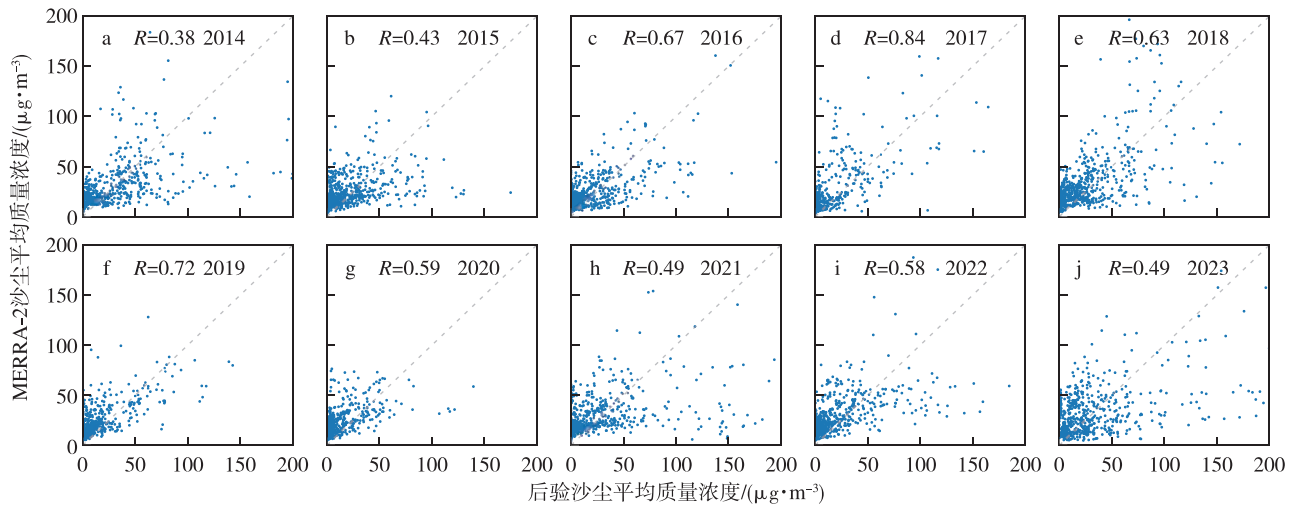


图 10 2014—2023 年(a—j)在 DAR2 区域内沙尘后验场与 MERRA-2 数据的平均质量浓度对比散点图(单位: $\mu\text{g} \cdot \text{m}^{-3}$)
 Fig.10 Scatter plots of averaged dust mass concentrations over DAR2 comparing posterior estimates with MERRA-2 dust products for (a—j) 2014—2023 (units: $\mu\text{g} \cdot \text{m}^{-3}$)

果存在较大不确定性。

本研究在前期开发的沙尘同化系统基础上,集成地面 PM_{10} 质量浓度、卫星 AOD 观测非沙尘组分偏差校正技术,以及适用于沙尘气溶胶强度、位置误差协同校正的有效时刻偏移卡尔曼滤波同化算法(VTS-EnKF),通过同化地面 PM_{10} 、MODIS/VIRRS AOD 卫星观测数据与 LOTOS-EUROS 沙尘模式,建立了 10 a(2014—2023 年)东亚地区春季(3—5 月)逐 3 h 的沙尘气溶胶三维质量浓度再分析数据集,分辨率为 $0.25^\circ \times 0.25^\circ$ 。

本研究分析了该再分析数据集相较于 MERRA-2 沙尘再分析产品的优势。在 $75 \mu\text{g} \cdot \text{m}^{-3}$ 以下,MERRA-2 数据集与再分析场之间的沙尘质量浓度估算是相对一致的。而在沙尘质量浓度较高的情况下,MERRA-2 相比于再分析场对沙尘质量浓度的估算存在明显的低估倾向,并在受沙尘天气

影响的区域表现出较大的不确定性。

本研究同时讨论了 2014—2023 年(共 10 a)春季沙尘月、年际变化趋势,揭示了沙尘在年际变化中的变化趋势,其中主要的沙尘活动出现在新疆的塔里木盆地以及中国的戈壁沙漠,同时蒙古戈壁区域也偶有沙尘的分布。通过与观测比较知道,先验场对沙尘的预估质量浓度通常偏高,通过同化后,可见广泛的负增量分布,平均质量浓度从 $65.24 \mu\text{g} \cdot \text{m}^{-3}$ 降至 $39.99 \mu\text{g} \cdot \text{m}^{-3}$ 。特别是对于人口稠密的区域 DAR1 和 DAR2,数据集能够准确再现沙尘天气事件的强度和动态,并能够捕捉到沙尘事件的发生时间。在数据同化前后,均方根误差从 $256.67 \mu\text{g} \cdot \text{m}^{-3}$ 降低至 $59.33 \mu\text{g} \cdot \text{m}^{-3}$,降幅达到 76.9%。在春季,整体上沙尘质量浓度随着月份的增加而增加,而在个别年份中,沙尘的月度趋势会因特定的沙尘暴事件产生变化。

参考文献(References)

Anderson J L, 2012. Localization and sampling error correction in ensemble Kalman filter data assimilation[J]. *Mon Wea Rev*, 140(7): 2359-2371. DOI: 10.1175/mwr-d-11-00013.1.

Anderson T L, Wu Y H, Chu D A, et al., 2005. Testing the MODIS satellite retrieval of aerosol fine-mode fraction[J]. *J Geophys Res: Atmos*, 110(D18): 2005JD005978. DOI: 10.1029/2005JD005978.

Bannister R N, 2017. A review of operational methods of variational and ensemble-variational data assimilation[J]. *Quart J Roy Meteor Soc*, 143(703): 607-633. DOI: 10.1002/qj.2982.

Brasseur G P, Xie Y, Petersen A K, et al., 2019. Ensemble forecasts of air quality in eastern China-Part 1: model description and implementation of the MarcoPolo-Panda prediction system, version 1[J]. *Geosci Model Dev*, 12(1): 33-67. DOI: 10.5194/gmd-12-33-2019.

Burgers G, van Leeuwen P J, Evensen G, 1998. Analysis scheme in the ensemble Kalman filter[J]. *Mon Wea Rev*, 126(6): 1719-1724. DOI: 10.1175/1520-0493(1998)126<1719: asitek>2.0.co;2.

曹皓东, 张鹏, 廖蜜, 等, 2024. 风云三号微波温度订正序列再定标历史数据集质量评估[J]. *大气科学学报*, 47(4): 669-680. Cao H D, Zhang P, Liao M, et al., 2024. Quality assessment of the long-term recalibration history dataset of the FY-3 microwave temperature sounder[J]. *Trans Atmos*

- Sci,47(4):669-680.DOI:10.13878/j.cnki.dqkxxb.20231121001.(in Chinese).
- Chen S Y,Zhao D,Huang J P,et al.,2023.Mongolia contributed more than 42% of the dust concentrations in northern China in March and April 2023 [J].*Adv Atmos Sci*,40(9):1549-1557.DOI:10.1007/s00376-023-3062-1.
- 陈耀登,沈洁,范水勇,等,2021.FY-4A 卫星云导风观测误差优化及同化效果影响研究[J].*大气科学学报*,44(3):418-427. Chen Y D,Shen J, Fan S Y,et al.,2021.A study of the observational error statistics and assimilation applications of the FY-4A satellite atmospheric motion vector [J].*Trans Atmos Sci*,44(3):418-427.DOI:10.13878/j.cnki.dqkxxb.20190921001.(in Chinese).
- Di Tomaso E,Schutgens N A J,Jorba O,et al.,2017.Assimilation of MODIS Dark Target and Deep Blue observations in the dust aerosol component of NMMB-MONARCH version 1.0[J].*Geosci Model Dev*,10(3):1107-1129.DOI:10.5194/gmd-10-1107-2017.
- Escribano J,Boucher O,Chevallier F,et al.,2016.Subregional inversion of North African dust sources[J].*J Geophys Res: Atmos*,121(14):8549-8566. DOI:10.1002/2016jd025020.
- Gong S L,Zhang X Y,2008.CUACE/Dust-an integrated system of observation and modeling systems for operational dust forecasting in Asia[J].*Atmos Chem Phys*,8(9):2333-2340.DOI:10.5194/acp-8-2333-2008.
- Huneeus N,Schulz M,Balkanski Y,et al.,2011.Global dust model intercomparison in AeroCom phase I[J].*Atmos Chem Phys*,11(15):7781-7816. DOI:10.5194/acp-11-7781-2011.
- Jin J B,Lin H X,Heemink A,et al.,2018.Spatially varying parameter estimation for dust emissions using reduced-tangent-linearization 4DVar[J].*Atmos Environ*,187:358-373.DOI:10.1016/j.atmosenv.2018.05.060.
- Jin J B,Lin H X,Segers A,et al.,2019a.Machine learning for observation bias correction with application to dust storm data assimilation[J].*Atmos Chem Phys*,19(15):10009-10026.DOI:10.5194/acp-19-10009-2019.
- Jin J B,Segers A,Heemink A,et al.,2019b.Dust emission inversion using himawari-8 AODs over East Asia; an extreme dust event in May 2017[J].*J Adv Model Earth Syst*,11(2):446-467.DOI:10.1029/2018ms001491.
- Jin J B,Segers A,Liao H,et al.,2020.Source backtracking for dust storm emission inversion using an adjoint method; case study of Northeast China [J].*Atmos Chem Phys*,20(23):15207-15225.DOI:10.5194/acp-20-15207-2020.
- Jin J B,Segers A,Lin H X,et al.,2021.Position correction in dust storm forecasting using LOTOS-EUROS v2.1: grid-distorted data assimilation v1.0 [J].*Geosci Model Dev*,14(9):5607-5622.DOI:10.5194/gmd-14-5607-2021.
- Jin J B,Pang M J,Segers A,et al.,2022.Inverse modeling of the 2021 spring super dust storms in East Asia[J].*Atmos Chem Phys*,22(10):6393-6410.DOI:10.5194/acp-22-6393-2022.
- Lin C,Wang Z,Zhu J,2008a.An ensemble Kalman filter for severe dust storm data assimilation over China[J].*Atmos Chem Phys*,8(11):2975-2983. DOI:10.5194/acp-8-2975-2008.
- Lin C Y,Zhu J,Wang Z F,2008b.Model bias correction for dust storm forecast using ensemble Kalman filter[J].*J Geophys Res: Atmos*,113(D14):2007JD009498.DOI:10.1029/2007JD009498.
- Niu T,Gong S L,Zhu G F,et al.,2008.Data assimilation of dust aerosol observations for the CUACE/dust forecasting system[J].*Atmos Chem Phys*,8(13):3473-3482.DOI:10.5194/acp-8-3473-2008.
- Pang M J,Jin J B,Segers A,et al.,2023.Dust storm forecasting through coupling LOTOS-EUROS with localized ensemble Kalman filter[J].*Atmos Environ*,306:119831.DOI:10.1016/j.atmosenv.2023.119831.
- Pang M J,Jin J B,Segers A,et al.,2024.Valid time shifting ensemble Kalman filter (VTS-EnKF) for dust storm forecasting[J].*Geosci Model Dev*,17(22):8223-8242.DOI:10.5194/gmd-17-8223-2024.
- Pang M,Jin J B,Yang T,et al.,2025.The sensitivity of aerosol data assimilation to vertical profiles; case study of dust storm assimilation with LOTOS-EUROS v2.2[J].*Geoscientific Model Development*,18(12):1-25.DOI:10.5194/gmd-18-3781-2025.
- Ravela S,Emanuel K,McLaughlin D,2007.Data assimilation by field alignment[J].*Phys D: Nonlinear Phenom*,230(1/2):127-145.DOI:10.1016/j.physd.2006.09.035.
- Tian R,Ma X Y,Zhao J Q,2021.A revised mineral dust emission scheme in GEOS-Chem; improvements in dust simulations over China[J].*Atmos Chem Phys*,21(6):4319-4337.DOI:10.5194/acp-21-4319-2021.
- Timmermans R,Kranenburg R,Manders A,et al.,2017.Source apportionment of PM_{2.5} across China using LOTOS-EUROS[J].*Atmos Environ*,164:370-386.DOI:10.1016/j.atmosenv.2017.06.003.
- Wang Y Q,Zhang X Y,Gong S L,et al.,2008.Surface observation of sand and dust storm in East Asia and its application in CUACE/Dust[J].*Atmos Chem Phys*,8(3):545-553.DOI:10.5194/acp-8-545-2008.
- Xie Y,Zhang W H,Qu J,2017.Detection of Asian dust storm using MODIS measurements[J].*Remote Sensing*,9(8):869.DOI:10.3390/rs9080869.
- Yin Z C,Wan Y,Zhang Y J,et al.,2022.Why super sandstorm 2021 in North China? [J].*Natl Sci Rev*,9(3):nwab165.DOI:10.1093/nsr/nwab165.
- 尹志聪,霍芊伊,麻晓晴,等,2023.触发2023年春季中国北方沙尘暴的沙源累积和天气扰动机制[J].*大气科学学报*,46(3):321-331. Yin Z C, Huo Q Y, Ma X Q, et al., 2023. Mechanisms of dust source accumulation and synoptic disturbance triggering the 2023 spring sandstorm in northern China[J].*Trans Atmos Sci*,46(3):321-331.DOI:10.13878/j.cnki.dqkxxb.20230501007.(in Chinese).
- Yuan Z H,Bao G,Dorjsuren A,et al.,2024.Climatic constraints of spring phenology and its variability on the Mongolian Plateau from 1982 to 2021

- [J]. *J Geophys Res: Biogeosciences*, 129(2): e2023JG007689. DOI: 10.1029/2023jg007689.
- Yumimoto K, Uno I, Sugimoto N, et al., 2008. Adjoint inversion modeling of Asian dust emission using lidar observations[J]. *Atmos Chem Phys*, 8(11): 2869-2884. DOI: 10.5194/acp-8-2869-2008.
- 张鹏, 王春姣, 陈林, 等, 2018. 沙尘气溶胶卫星遥感现状与需要关注的若干问题[J]. *气象*, 44(6): 725-736. Zhang P, Wang C J, Chen L, et al., 2018. Current status of satellite-based dust aerosol remote sensing and some issues to be concerned[J]. *Meteor Mon*, 44(6): 725-736. DOI: 10.7519/j.issn.1000-0526.2018.06.001. (in Chinese).
- 张小曳, 2001. 亚洲粉尘的源区分布、释放、输送、沉降与黄土堆积[J]. *第四纪研究*, 21(1): 29-40. Zhang X Y, 2001. Source distributions, emission, transport, deposition of Asian dust and loess accumulation[J]. *Quat Sci*, 21(1): 29-40. (in Chinese).
- 赵新宇, 闵锦忠, 朱利剑, 等, 2023. 风廓线雷达资料的应用: 质量评估[J]. *大气科学学报*, 46(3): 453-465. Zhao X Y, Min J Z, Zhu L J, et al., 2023. Application of wind profiler radar data: quality assessment[J]. *Trans Atmos Sci*, 46(3): 453-465. DOI: 10.13878/j.cnki.dqkxxb.20220503002. (in Chinese).
- Zhou C H, Gong S L, Zhang X Y, et al., 2008. Development and evaluation of an operational SDS forecasting system for East Asia: CUACE/Dust[J]. *Atmos Chem Phys*, 8(4): 787-798. DOI: 10.5194/acp-8-787-2008.

· ARTICLE ·

A decadal dust aerosol mass concentration reanalysis over East Asia during 2014—2023

JIN Jianbing¹, LI Dehao¹, PANG Mijie², CHENG Zheqi¹, XU Canjie¹, LIAO Hong¹

¹*School of Environmental Science and Engineering/State Key Laboratory of Climate System Prediction and Risk Management (CPRM)/Jiangsu Key Laboratory of Atmospheric Environment Monitoring and Pollution Control/Jiangsu Collaborative Innovation Center of Atmospheric Environment and Equipment Technology, Nanjing University of Information Science and Technology, Nanjing 210044, China;*

²*Institute of Applied Mathematics, Delft University of Technology, Delft 2628 CD, Netherlands*

Abstract Dust storms are among the most severe hazardous weather phenomena affecting northern China and adjacent regions. The primary dust source areas—including the Alxa-Hexi Corridor, the Tengger Desert, and the southern Mongolian Gobi Desert—emit more than 800 Mt of dust annually. During spring, the interaction between the Siberian high and Mongolian cyclones generates strong near-surface winds and enhanced vertical convection, forming a three-dimensional “uplift-suspension-transport” structure that promotes dust storm development. Under ongoing global warming, declining spring precipitation over the Mongolian Plateau and extensive desertification—currently affecting over 75% of Mongolia—are expected to further intensify transboundary dust transport into China, with severe consequences for public health, agriculture, and transportation. These challenges underscore the urgent need for long-term, high-quality dust datasets to improve understanding of dust emission mechanisms and forecasting capabilities.

Atmospheric models are essential tools for simulating dust emission, transport, and deposition, as well as for assessing impacts on climate, ecosystems, and human health. However, large uncertainties in emission parameterizations and long-range transport processes persist, often resulting in substantial biases in simulated dust concentrations, in some cases differing from observations by up to two orders of magnitude. Recent advances in atmospheric observation systems provide valuable constraints, including China’s nationwide hourly PM₁₀ monitoring network and satellite remote sensing products with broad spatial coverage and multi-dimensional aerosol information, such as MODIS aerosol optical depth (AOD). In this context, data assimilation methods grounded in Bayesian theory offer an effective framework for integrating observational data with model simulations to generate spatially continuous and more accurate dust reanalysis datasets. Despite progress, existing studies have primarily focused on individual dust events, and long-term dust reanalysis efforts remain limited due to observation biases, sparse data coverage over source regions, transport errors, and the strong spatiotemporal variability of dust emissions.

Building upon a self-developed dust storm assimilation system, this study integrates ground-based PM_{10} observations, bias-corrected satellite AOD data, and an effective valid time shift ensemble Kalman filter (VTS-EnKF) designed to jointly correct dust intensity and transport position errors. Using this framework, we construct a high-resolution ($0.25^\circ \times 0.25^\circ$, 3-hourly) three-dimensional dust aerosol mass concentration reanalysis dataset for East Asia during spring (March-May) over the period 2014—2023. This dataset provides a robust basis for investigating long-term dust variability, transboundary transport processes, and associated impacts on climate, the environment, and public health.

Comparisons with MERRA-2 dust reanalysis demonstrate clear advantages of the newly developed dataset. While MERRA-2 exhibits reasonable agreement at low dust concentrations ($<75 \mu\text{g} \cdot \text{m}^{-3}$), it substantially underestimates dust levels and exhibits larger uncertainties under moderate to severe dust conditions, particularly in dust-affected regions. Analysis of springtime dust variability from 2014 to 2023 reveals pronounced interannual and spatial heterogeneity, with dominant dust activity over the Tarim Basin and the Gobi Desert in China and episodic contributions from the Mongolian Gobi. Relative to observations, prior simulations tend to overestimate dust concentrations, whereas data assimilation introduces widespread negative analysis increments, reducing the regional mean concentration from 65.24 to 39.99 $\mu\text{g} \cdot \text{m}^{-3}$. Notably, the reanalysis accurately captures both the intensity and timing of dust events in densely populated areas. Overall, the assimilation framework substantially improves dust representation, reducing RMSE by 76.9% and yielding a more reliable depiction of monthly and interannual dust variability.

Keywords dust storm; reanalysis data; data assimilation

DOI: 10.13878/j.cnki.dqkxxb.20251102008

(责任编辑:刘菲)