氣候與環境研究
Climatic and Environmental Research

一种基于集合最优插值的排放源快速反演方法

吴煌坚 林伟 孔磊 唐晓 王威 王自发 陈松蹊

A Fast Emission Inversion Scheme Based on Ensemble Optimal Interpolation

WU Huangjian LIN Wei KONG Lei TANG Xiao WANG Wei WANG Zifa CHEN Songxi

在线阅读 View online: https://doi.org/10.3878/j.issn.1006-9585.2020.20043

您可能感兴趣的其他文章 Articles you may be interested in

沙尘预报客观检验中最优实况资料和圆插值半径选择研究

Research on the Selection of the Optimal Observations and Radius of Circular Interpolation during Objective Verification of Sand-Dust Forecasts

气候与环境研究. 2019, 24(6): 735 https://doi.org/10.3878/j.issn.1006-9585.2019.18082

基于西南地区台站降雨资料空间插值方法的比较

Comparison of Spatial Interpolation Methods Based on Monthly Precipitation Obsevation Data of Station in Southwest China

气候与环境研究. 2019, 24(1): 50 https://doi.org/10.3878/j.issn.1006-9585.2017.17072

不同人为源排放清单对大气污染物浓度数值模拟的影响：以浙江省为例

Effects of Different Anthropogenic Emission Inventories on Simulated Air Pollutants Concentrations: A Case Study in Zhejiang Province

气候与环境研究. 2017, 22(5): 519 https://doi.org/10.3878/j.issn.1006-9585.2017.16112

CMIP6情景中主要温室气体和气溶胶排放强度的时空分布特征分析

SpatialTemporal Characteristics of the Emission Intensities of Several Major Greenhouse Gases and Aerosols under CMIP6 Scenarios

气候与环境研究. 2020, 25(4): 366 https://doi.org/10.3878/j.issn.1006-9585.2020.20005

基于机器学习的华北气温多模式集合预报的订正方法

A Temperature Correction Method for Multi-model Ensemble Forecast in North China Based on Machine Learning

气候与环境研究. 2019, 24(1): 116 https://doi.org/10.3878/j.issn.1006-9585.2018.18049

基于能见度及AOD数据的北京市PM2.5浓度的反演

Inversion of PM2.5 Concentration in Beijing Based on Visibility and AOD Data

气候与环境研究. 2020, 25(5): 521 https://doi.org/10.3878/j.issn.1006-9585.2019.19114

关注微信公众号，获得更多资讯信息

吴煌坚, 林伟, 孔磊, 等. 2021. 一种基于集合最优插值的排放源快速反演方法 [J]. 气候与环境研究, 26(2): 191−201. WU Huangjian, LIN Wei, KONG Lei, et al. 2021. A Fast Emission Inversion Scheme Based on Ensemble Optimal Interpolation [J]. Climatic and Environmental Research (in Chinese), 26 (2): 191−201. doi:10.3878/j.issn.1006-9585.2020.20043

一种基于集合最优插值的排放源快速反演方法

吴煌坚[1] 林伟[2,3] 孔磊[4,5] 唐晓[4] 王威[6] 王自发[4,5,7] 陈松蹊[1,2,3]

1 北京大学光华管理学院，北京 100871
2 北京大学数学科学学院，北京 100871
3 北京大学统计科学中心，北京 100871
4 中国科学院大气物理研究所大气边界层物理和大气化学国家重点实验室，北京 100029
5 中国科学院大学，北京 100049
6 中国环境监测总站，北京 100012
7 中国科学院区域大气环境研究卓越创新中心，福建厦门 361021

摘 要 基于集合卡尔曼滤波的源反演方法是估计排放源、提高空气质量模拟和预报精度的有效方法。为构建排放源与污染物浓度之间的误差协方差矩阵，该方法通常需要运行几十次大气化学传输模式。庞大的计算量限制了该方法的应用，使其无法为实时预报系统快速更新排放源。本研究发展了一种基于集合最优插值的排放源反演方法。该方法使用历史集合数据构建误差协方差矩阵，仅需一次常规的空气质量模拟便可根据观测模拟差异反演排放源，从而显著降低计算量。本文使用该方法同化 2015 年 1 月全国 1107 个地面站点观测的 CO 小时浓度数据，结合 2014 年 1 月的历史集合数据集，估计 2015 年 1 月全国 15 km 分辨率的 CO 排放源。该方案反演的全国 CO 排放总量仅比使用 2015 年 1 月集合数据集的反演量高 1%，表明历史时段与反演时段的气象条件差异对月均 CO 排放的影响有限。使用历史集合数据集更新的排放源再次模拟可将全国 349 个独立验证站点的平均低估从 0.74 mg m^{-3} 降至 0.01 mg m^{-3}，均方根误差降低 18%，表明该方法可快速更新排放源并降低其不确定性。

关键词 排放源反演 集合最优插值 CO 排放源

文章编号 1006-9585(2021)02-0191-11 **中图分类号** P402 **文献标识码** A

doi:10.3878/j.issn.1006-9585.2020.20043

A Fast Emission Inversion Scheme Based on Ensemble Optimal Interpolation

WU Huangjian[1], LIN Wei[2,3], KONG Lei[4,5], TANG Xiao[4], WANG Wei[6], WANG Zifa[4,5,7], and CHEN Songxi[1,2,3]

1 *Guanghua School of Management*, *Peking University*, *Beijing* 100871
2 *School of Mathematical Sciences*, *Peking University*, *Beijing* 100871
3 *Center for Statistical Science*, *Peking University*, *Beijing* 100871
4 *State Key Laboratory of Atmospheric Boundary Layer Physics and Atmospheric Chemistry*, *Institute of Atmospheric Physics*, *Chinese Academy of Sciences*, *Beijing* 100029
5 *University of Chinese Academy of Sciences*, *Beijing* 100049

收稿日期 2020-04-10；**网络预出版日期** 2020-12-08

作者简介 吴煌坚，男，博士，1990 年出生，助理研究员，从事大气化学资料同化及源反演相关研究。E-mail: wuhuangjian@pku.edu.cn

资助项目 国家重点研发计划 2016YFC0207703、2016YFC0207701、2016YFC0208903，北京市自然科学基金 Z190001，国家自然科学基金 11671018、71532001、41875164，北京智源人工智能研究院项目

Funded by National Key Research and Development Program of China (Grants 2016YFC0207703, 2016YFC0207701, and 2016YFC0208903), Beijing Natural Science Foundation (Grant Z190001), National Natural Science Foundation of China (Grants 11671018, 71532001, and 41875164), Beijing Academy of Artificial Intelligence Program

6 *China National Environmental Monitoring Center*, *Beijing* 100012
7 *Center for Excellence in Regional Atmospheric Environment*, *Institute of Urban Environment*, *Chinese Academy of Sciences*, *Xiamen*, *Fujian Province* 361021

Abstract The emission inversion based on the ensemble Kalman filter (EnKF) is an effective method for estimating emissions and improving air quality modeling and forecasting. However, to construct the error covariance matrix between the emissions and pollutant concentrations, this method requires running the chemical transport model tens of times, which is computationally prohibitive and limits its application in updating the emissions for a real-time forecasting system. This study develops an emission inversion method based on the ensemble optimal interpolation (EnOI). The proposed method calculates the error covariance matrix from historical ensemble data and requires only a routine air quality simulation run for emission inversion from the contrasts between the observations and simulations, thereby greatly reducing the computational cost. The proposed method is applied to assimilating hourly surface observations of CO concentrations at 1107 sites over China in January 2015. During the experiment, CO emissions in January 2015 are estimated at a 15-km horizontal resolution using the historical ensemble dataset for January 2014. The total CO emission in China estimated by this scheme is only 1% higher than using an ensemble dataset for January 2015, indicating that the differences in meteorological conditions between the historical and estimated periods have a limited impact on the inversely estimated monthly CO emission. Simulations with the updated emissions reveal a decrease in the downward bias of average CO concentrations at 349 independent validation sites from 0.74 mg m^{-3} to 0.01 mg m^{-3} and a reduction of the root-mean-square error by 18%. The results suggest that the proposed method can be used as a fast emission updating scheme to lessen the uncertainties in the emission inventories.

Keywords Emission inversion, Ensemble optimal interpolation, CO emission

1 引言

大气污染源排放清单是大气污染预报最为重要的基础数据之一，同时也是大气化学传输模式的重要不确定性来源。Moore and Londergan（2001）定量研究了空气质量模拟的不确定性来源，结果表明排放源不确定性最大，占预报不确定性的 25%～50%，高于气象、边界条件、初始浓度的影响。Beekmann and Derognat（2003）和 Hanna et al.（1998）的研究同样表明，排放源是模拟不确定性最重要的来源之一。传统自下而上的源清单构建方法通过排放因子和活动水平推算排放量（Hao et al., 2002）。该方法需要统计大量排放个体信息，包含每个点源的准确位置。但受有限的资源和实地信息制约，排放清单更新缓慢且包含较大不确定性（Zhang et al., 2009; 曹国良等, 2011; 魏巍等, 2011）。近年来发表的排放清单中（Zhang et al., 2009; 魏巍等, 2011; Kurokawa et al., 2013; Li et al., 2017），细颗粒物（PM2.5）、可吸入颗粒物（PM10）的排放不确定性在 100% 左右，一氧化碳（CO）和挥发性有机物（VOCs）的不确定性多在 50% 以上；部分清单的 SO_2 不确定性虽可低至 12%，但近年来脱硫设备和散煤管控等措施的实施使得 SO_2 排放大幅下降，更新缓慢的排放清单难以满足高精度预报预警和减排评估的需求。

随着观测技术的发展，越来越多研究者开始利用观测数据结合源反演方法估算排放清单（Streets et al., 2013）。作为先进的同化方法之一，集合卡尔曼滤波（EnKF）（Evensen, 2003）已经成为校验和优化清单的重要方法。朱江和汪萍（2006）使用理想试验探讨了利用 EnKF 开展排放源反演的可行性。Tang et al.（2013）使用该方法反演了北京及周边的 CO 排放，并对比反演前后的模拟结果，发现源反演可将验证站点偏差降低 48%。Miyazaki et al.（2017）通过同化多个卫星资料反演多年的地面 NO_x 排放，发现全球 NO_x 排放总量在 2005 年 [47.9 Tg(N) a^{-1}] 至 2014 年 [47.5 Tg(N) a^{-1}] 之间仅有少量变化。Peng et al.（2017）使用 EnKF 方法同时同化初始场和排放源，发现排放源反演可大幅提高长三角和珠三角的夜间 PM2.5 预报效果。

虽然基于 EnKF 的排放源反演方法可有效提高源清单精度，但该方法需多次运行大气化学传输模式，庞大的计算需求限制了其应用范围，使其无法为实时空气质量预报快速更新排放清单。为减少 EnKF 计算量，Evensen（2003）提出其次优方案，即集合最优插值（EnOI）。该方案利用不同时刻

的模拟结果构建集合状态，避免多次运行预报模式，从而大幅降低计算量。该方法在海洋资料同化中得到广泛应用（Counillon and Bertino, 2009; Oke et al., 2010; Kaurkin et al., 2016）。张金谱等（2014）将 EnOI 方法应用于珠三角的 PM10、SO_2 和 NO_2 浓度同化，使验证站点的均方根误差降低 32%～42%。然而常规的 EnOI 方法未扰动大气化学传输模式，难以建立排放强度与污染物浓度之间的误差协方差矩阵，故而无法直接应用于排放源反演。为建立该联系，Wang et al.（2016）对排放源进行扰动，使用扰动后模拟的集合数据计算排放与浓度之间的误差协方差矩阵，进而反演黑碳排放。该方案每次反演均需人工合成反演时段的平均气象场，并用其驱动 24 h 的集合模拟。对于长生命周期的污染物，该方法还需使用更多的计算资源延长集合模拟时段，从而使排放源扰动信息在模拟浓度中得到充分体现。

为降低源反演的计算量，本文发展了一种基于 EnOI 的排放源快速反演方法。该方法使用历史集合数据构建误差协方差矩阵，再根据常规的空气质量观测和模拟浓度反演排放源。相比基于 EnKF 的源反演方法和 Wang et al.（2016）的方法，本方法无需多次使用大气化学传输模式模拟反演时段，取得可重复使用的历史集合数据集后，仅需一次常规空气质量模拟即可快速更新排放清单。本文将该方法应用于主要大气污染物之一的 CO 排放源反演。虽然环境 CO 浓度很少超过世界卫生组织的警戒值（8 h 平均浓度不超过 10 mg m^{-3}）（https://apps.who.int/iris/handle/10665/260127[2020-04-10]），但它是对流层臭氧的重要前体物，同时是大气氧化性的主要控制因素（Dekker et al., 2017）。CO 的主要清除方式是被 OH 自由基氧化，其生命周期在几周至几个月之间（Holloway et al., 2000），长短适中的生命周期使其成为污染传输的常用示踪物（Naeher et al., 2001）。

2　方法和数据

2.1　基于集合最优插值（EnOI）的源反演方法

在 EnOI 的多数应用中，状态变量与观测变量相同，然而在源反演中，待更新的状态变量为排放强度，与观测的污染物浓度不同，故而无法将 EnOI 直接应用于排放源反演。针对于此，本文借鉴 EnKF 的参数估计方法，将浓度向量与排放源强度的组合作为状态变量：

$$\boldsymbol{X}_i = \begin{bmatrix} \boldsymbol{C}_i \\ \boldsymbol{E}_i \end{bmatrix}, i = 1, \cdots, m, \tag{1}$$

其中，$\boldsymbol{X}_i$ 为第 i 个状态变量，$\boldsymbol{C}_i$ 和 $\boldsymbol{E}_i$ 分别为第 i 个集合成员的浓度和排放强度，m 为集合成员个数。

EnOI 的状态更新方法与 EnKF 基本一致，其核心区别在于 EnKF 多次使用大气化学传输模式模拟同化时段，而 EnOI 使用历史的模拟样本构建集合。在以往研究中，此集合常为不同时刻的模拟结果。然而此方案未扰动排放源，无法建立排放与浓度之间的误差协方差矩阵。鉴于此，本文使用扰动排放源模拟得到的历史集合数据构建 EnOI 所需集合：

$$\boldsymbol{X}_i^{\mathrm{b}} = \begin{bmatrix} \boldsymbol{C}^{\mathrm{b}} + \boldsymbol{C}_i^{\mathrm{h}} - \overline{\boldsymbol{C}^{\mathrm{h}}} \\ \boldsymbol{E}^{\mathrm{b}} + \boldsymbol{E}_i^{\mathrm{h}} - \overline{\boldsymbol{E}^{\mathrm{h}}} \end{bmatrix}, i = 1, \cdots, m, \tag{2}$$

$$\boldsymbol{C}^{\mathrm{h}} = \left[\boldsymbol{C}_1^{\mathrm{h}}, \cdots, \boldsymbol{C}_m^{\mathrm{h}}\right], \tag{3}$$

其中，$\boldsymbol{X}_i^{\mathrm{b}}$ 为构建的第 i 个背景状态变量，$\boldsymbol{C}^{\mathrm{b}}$ 和 $\boldsymbol{E}^{\mathrm{b}}$ 分别为反演时段的模拟污染物浓度和排放强度，$\boldsymbol{C}_i^{\mathrm{h}}$ 和 $\boldsymbol{E}_i^{\mathrm{h}}$ 分别为历史集合中第 i 个成员的模拟浓度和排放强度（历史集合数据集在 2.3 节详细介绍），$\overline{\boldsymbol{C}^{\mathrm{h}}}$和$\overline{\boldsymbol{E}^{\mathrm{h}}}$分别为历史集合的浓度和排放的平均值。

构建状态集合后，EnOI 可以使用与 EnKF 相同的状态更新方法。本文采用 Sakov and Oke（2016）提出的确定性 EnKF（DEnKF）方法进行排放源反演。相比 EnKF 方法，它无需扰动观测，避免在扰动过程中引入误差。同时还支持基于舒尔（Schur）积的局地化方法，且计算上更为高效。Sun et al.（2009）对比了几种常见的集合滤波算法，结果表明 DEnKF 最稳健，且在小集合的同化中精度最高。该方法的状态更新公式如下：

$$\overline{\boldsymbol{X}^{\mathrm{a}}} = \overline{\boldsymbol{X}^{\mathrm{b}}} + \boldsymbol{K}\left(\boldsymbol{y} - \boldsymbol{H}\overline{\boldsymbol{X}^{\mathrm{b}}}\right), \tag{4}$$

其中，$\overline{\boldsymbol{X}^{\mathrm{a}}}$和$\overline{\boldsymbol{X}^{\mathrm{b}}}$分别为分析状态和背景场的集合均值，$\boldsymbol{y}$ 为观测向量，$\boldsymbol{H}$ 为观测算子。增益矩阵 $\boldsymbol{K}$ 的计算公式如下：

$$\boldsymbol{K} = \boldsymbol{P}\boldsymbol{H}^{\mathrm{T}}\left(\boldsymbol{H}\boldsymbol{P}\boldsymbol{H}^{\mathrm{T}} + \boldsymbol{R}\right)^{-1}, \tag{5}$$

其中，$\boldsymbol{R}$ 为观测误差协方差矩阵，本文假设不同站点间的观测误差不相关，站点误差估计采用李飞等（2019）基于模拟分辨率的方案。$\boldsymbol{P}$ 为背景场的误差协方差矩阵，通过下式计算：

$$\boldsymbol{P} = \boldsymbol{\rho} \circ \left(\boldsymbol{A}^{\mathrm{b}}\boldsymbol{A}^{\mathrm{bT}}\right) / (m-1), \tag{6}$$

其中，∘代表舒尔积，$\boldsymbol{\rho}$为局地化矩阵，用于抑制长距离虚假相关，其元素个数与 $\boldsymbol{A}^{\mathrm{b}}\boldsymbol{A}^{\mathrm{bT}}$ 相同，各元素取值采用基于格点间距离的局地化函数（Gaspari and Cohn, 1999）。$\boldsymbol{A}^{\mathrm{b}}$ 为背景状态增量矩阵，表示集合成员与集合均值的差异：

$$\boldsymbol{A}^{\mathrm{b}} = \boldsymbol{X}^{\mathrm{b}}(\boldsymbol{I} - \boldsymbol{1}/m), \tag{7}$$

其中矩阵 **1** 是单位矩阵 $\boldsymbol{I}$ 的同型矩阵，且各元素皆为 1。

使用公式（4）可以更新集合状态的均值，而状态增量矩阵的更新公式如下：

$$\boldsymbol{A}^{\mathrm{a}} = \boldsymbol{A}^{\mathrm{b}} - \boldsymbol{KHA}^{\mathrm{b}}/2, \tag{8}$$

其中 $\boldsymbol{A}^{\mathrm{a}}$ 为更新后的状态增量矩阵。结合公式（4）和公式（7）可得各集合成员的分析状态：

$$\boldsymbol{X}_i^{\mathrm{a}} = \overline{\boldsymbol{X}^{\mathrm{a}}} + \boldsymbol{A}_i^{\mathrm{a}}, i = 1, \ldots, m, \tag{9}$$

其中 $\boldsymbol{X}_i^{\mathrm{a}}$ 为第 i 个集合成员的分析状态，包含更新后的排放源 E_i^{a}。

该方法的核心优势在于可重复利用历史集合数据，反演时段无需复杂的集合模拟和设置，仅需一次常规的空气质量模拟，即可反演模拟时段的排放源。故而该方法的计算量显著低于基于 EnKF 的排放源反演方法，有助于拓展排放源反演的应用范围，提高排放源的更新速度。

2.2 数值模式介绍

本文的空气质量模拟采用王自发等（2006）发展的嵌套网格空气质量预报模拟系统（NAQPMS）。该系统涉及的物理化学过程主要包括平流、扩散过程，干、湿沉降过程，气相、液相和非均相化学过程。其空间结构采用三维欧拉输送模型，以地形追随坐标为垂直坐标，可同时模拟区域和城市尺度沙尘、PM2.5、PM10、SO_2、NO_x、CO、O_3、NH_3 等多种污染物（王自发等, 2008）。其中气相化学模块采用 CBM-Z 机制（Zaveri and Peters, 1999），考虑了 71 种化学反应物种，133 个化学反应；液相化学模块基于 RADM 的反应机制，包含 22 种气体和气溶胶（Chang et al., 1987）。NAQPMS 的气溶胶包括硫酸盐、硝酸盐、铵盐和二次有机气溶胶。该系统使用 ISORROPIA（Nenes et al., 1998）计算气溶胶的无机化学组分在颗粒态与气态之间的分配及气溶胶的含水量。

本文中，NAQPMS 模拟所需气象场由中尺度天气预报模式（WRF）提供。气象模拟采用图 1 所示的两层嵌套设置。其中第一层嵌套（D1）的水平分辨率为 45 km×45 km，覆盖了亚洲大部分区域；第二层嵌套（D2）的分辨率为 15 km×15 km，覆盖了除南海南部以外的中国区域。本研究以 WRF 连续模拟 36 h、取后 24 小时的方式为 NAQPMS 模拟提供每日的气象场驱动，再将每日气象场拼接得到模拟时段的完整气象场。NAQPMS 模拟采用单层嵌套设置，模拟 D2 层 15 km 分辨率的各污染

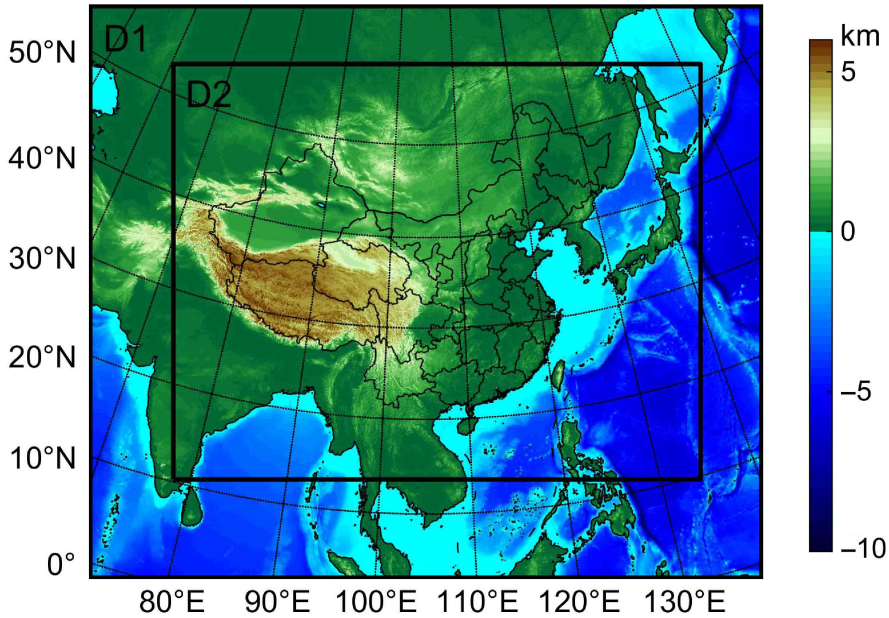

图 1 区域设置

Fig. 1 Domain configuration

物浓度。

本文使用清华大学提供的 2010 年 MIX 清单作为先验排放源（Li et al., 2017）。该清单集成了 MEIC（中国大陆）、REAS2（日本、台湾）、PKU-NH_3（中国氨排放清单）、CAPSS（韩国）等亚洲各地区的排放清单，涵盖 10 种大气污染物和温室气体，网格分辨率为 0.25°（纬度）×0.25°（经度）。该清单被用于东亚模式比较计划第三期（MICS-Asia III）的模拟研究，并且被联合国半球大气污染传输计划（HTAP）采用（http://www.meicmodel.org[2020-04-10]）。

2.3　历史集合数据集

本文通过 NAQPMS 模式模拟的历史集合数据集建立浓度与排放源之间的联系，进而使用观测浓度反演排放源。为使该数据集反映排放误差对浓度的影响，需对排放源进行扰动。若直接对每个网格的排放施加扰动，远小于格点数的扰动样本将包含过大的采样误差。鉴于此，本文采用 Evensen（2003）的伪随机扰动方法生成 500 组均值为 0、标准差为 1 的伪随机扰动样本。该样本在每个空间格点上服从标准正态分布，每个样本在空间上均较为平滑，且格点间样本的相关性随距离递减。当格点间距离达到解相关距离时，样本的理想相关系数为 e^{-1}。若解相关距离过大，反演算法难以差异化调整临近排放源；若解相关系数过小，有限的样本将产生显著的远距离虚假相关。本文使用试错法在不产生显著虚假相关的前提下减小解相关距离，最终将解相关距离设为 150 km。为减少 NAQPMS 模拟次数，本文采用 Evensen（2004）提出的算法将 500 组扰动样本压缩为 50 组。以上先过量采样再压缩的方法，可以在不增加样本量的前提下减小抽样误差，规避小概率样本，使最终生成的样本在每个格点上更接近标准正态分布。根据先验排放源的不确定性（Li et al., 2017），本文将压缩后的扰动样本均值调整为 1，将标准差按表 1 调整。使用调整后的样本扰动先验排放源，得到集合模拟所需的扰动排放清单，即：

$$\boldsymbol{E}_i^{\mathrm{h}} = \boldsymbol{E}^{\mathrm{o}} \circ \boldsymbol{\beta}_i^{\mathrm{h}}, i = 1, \cdots, 50, \qquad (10)$$

其中，$\boldsymbol{E}^{\mathrm{o}}$ 为先验排放源，$\boldsymbol{\beta}_i^{\mathrm{h}}$ 和 $\boldsymbol{E}_i^{\mathrm{h}}$ 分别为第 i 个集合成员的扰动样本和扰动排放源。将扰动后的 50 组排放源分别使用 NAQPMS 模拟，得到对应的 50 组模拟浓度，取其时间平均，即得 $\boldsymbol{C}_i^{\mathrm{h}}$（$i$=1, ⋯, 50）。

表 1　排放源扰动系数标准差

Table 1　Standard deviations of the emission perturbation coefficients

污染物	标准差
PM2.5	1.3
PM10	1.32
SO_2	0.12
NO_x	0.31
CO	0.7
NMVOCs	0.68

在历史集合数据集的模拟中，本研究采用与 2.2 节相同的模式和区域设置。模拟时段为 2013 年 12 月 17 日至 2014 年 1 月 31 日和 2014 年 12 月 17 日至 2015 年 1 月 31 日。从两时段中分别取 2014 年 1 月和 2015 年 1 月的模拟月均浓度和扰动排放源构建 EnOI 所需的背景状态集合。值得指出的是，该数据可重复应用于 EnOI 源反演，反演时段所采用的模式和区域设置均无需与历史数据集相同。

2.4　试验设置

基于 EnOI 的源反演方法根据模拟浓度误差调整排放源。为反演 2015 年 1 月全国的 CO 排放，本文首先使用先验排放源模拟 2014 年 12 月 25 日至 2015 年 1 月 31 日的全国空气质量，再使用两组集合数据集分别进行两组反演试验。反演试验 1 使用 2014 年 1 月历史集合数据集构建浓度与排放之间的误差协方差矩阵，而反演试验 2 使用 2015 年 1 月的集合数据集。两组试验均根据 2015 年 1 月同化站点的 CO 观测模拟浓度差异，反演 2015 年 1 月全国 15 km 分辨率的 CO 排放。反演中，观测数据来自中国环境监测总站，站点分布与人口分布基本一致，主要集中在胡焕庸线以东地区。为保障数据质量，本研究先采用一种基于概率分布的自适应离群值检测方法剔除异常数据（Wu et al., 2018），再排除有效数据较少的站点，得到全国共 1456 个站点的观测数据集。为客观评估反演效果，本研究在包含两个以上观测站点的城市中随机选取一个站点作为独立验证站点，最终从全国 1456 个站点中选取 349 个站点作为验证站点，其余 1107 个站点作为同化站点。两类站点的分布如图 2 所示。

为验证源反演效果，本文使用反演的排放源再次模拟 2014 年 12 月 25 日至 2015 年 1 月 31 日的

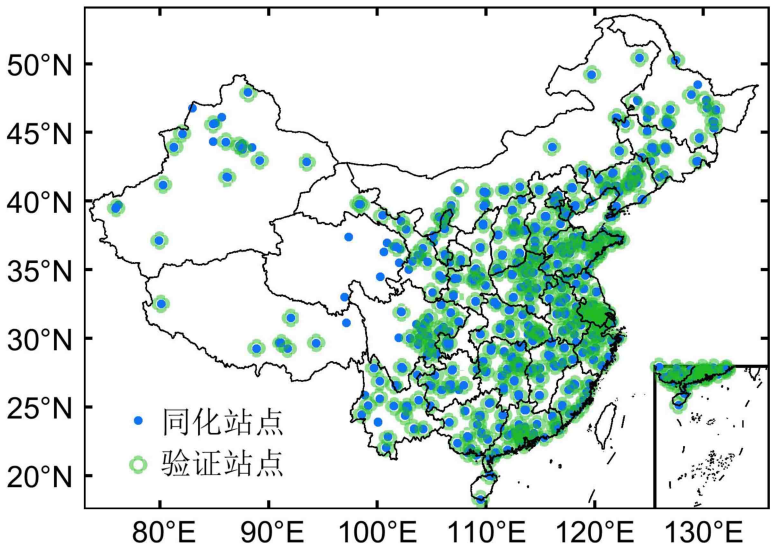

图 2 空气质量监测站点分布

Fig. 2 Distribution of air quality monitoring stations

全国空气质量。通过对比两套排放源在独立验证站点的 CO 模拟结果，定量评估基于 EnOI 的源反演对 CO 模拟的改进效果。

3 结果

本文首先使用先验排放源模拟 2015 年 1 月全国空气质量。图 3a 对比了地面月均 CO 模拟和观测浓度的空间分布。由图可见，采用先验排放源模拟的 CO 浓度在全国大部分地区显著低于观测值，特别是在山西、吉林、辽宁等地和珠三角周边的中国南部地区。全国 349 个验证站点的月均浓度偏差达−0.74 mg m^{-3}。图 4 使用箱型图对比了全国验证站点每日的观测和模拟浓度。由图可见，模拟的 CO 浓度变化趋势与观测较为一致，峰值均出现在 4 日、10 日和 25 日，低值均出现在 6 日和 27 日。但模拟浓度的中位数和变化范围普遍低于观测浓度。图 5 给出了北京、太原、武汉、广州 4 个城市验证站点的进一步分析（图 6a 展示了 2010 年 1 月的先验 CO 排放强度）。4 个城市中，北京市的 CO 模拟相关系数最高，达 0.81，日均浓度偏差为−0.35 mg m^{-3}；太原市的模拟浓度存在显著低估，偏差达−2.16 mg m^{-3}；广州市也有明显低估，偏差为−0.69 mg m^{-3}；武汉市的模拟结果则有明显高估，偏差为 1.20 mg m^{-3}。

模拟浓度低于观测反映出先验排放强度可能低于真实排放，反之亦然。反演试验 1 根据 2015 年 1 月全国 1115 个站点的模拟误差，结合 2014 年 1 月的集合数据集，使用 EnOI 方法反演得到 2015 年 1 月的排放源调整系数（如图 6b）。由于观测站点未覆盖所有区域，故而台湾、新疆、西藏、青海和内蒙古的部分区域未进行反演调整。但除台湾外，大部分高排放地区均得到有效反演，未反演区域的先验排放强度普遍较低，对反演误差影响较小。反演后大部分地区的排放强度均有所增加，全国 2015 年 1 月的 CO 排放量为 46.6 Tg，比先验排放量（21.2 Tg）高 120%。在 CO 浓度被显著低估的山西和新疆等地，反演的排放强度可比先验排放源高 3 倍以上。

为评估源反演效果，本文使用反演的排放源再次模拟全国 CO 浓度。由图 3b 可见，源反演可显著提高 CO 模拟浓度，从而将偏差幅度由 0.74 mg m^{-3} 降至 0.01 mg m^{-3}。其中，山西、甘肃、两广地区和东北三省的 CO 模拟偏差均有显著降低。与此同时，源反演可将全国验证站点的平均均方根误差降低 18%，相关系数提升 0.03。在图 4 的箱形图对比中，反演后模拟浓度的中位数和变化范围均显著增大，与观测高度一致。在 4 城市验证站点的浓度对比（图 5）中，源反演将 CO 模拟的相关系数提高 0.02～0.20。在太原市验证站点中，源反演将 CO 浓度提高 1.97 mg m^{-3}，偏差幅度降低 91%。与此同时，广州市验证站点的偏差幅度也大幅降低，降幅达 88%。在 CO 浓度被高估的武汉，源反演可降低其排放强度和模拟浓度，同时大幅抑制 9～11

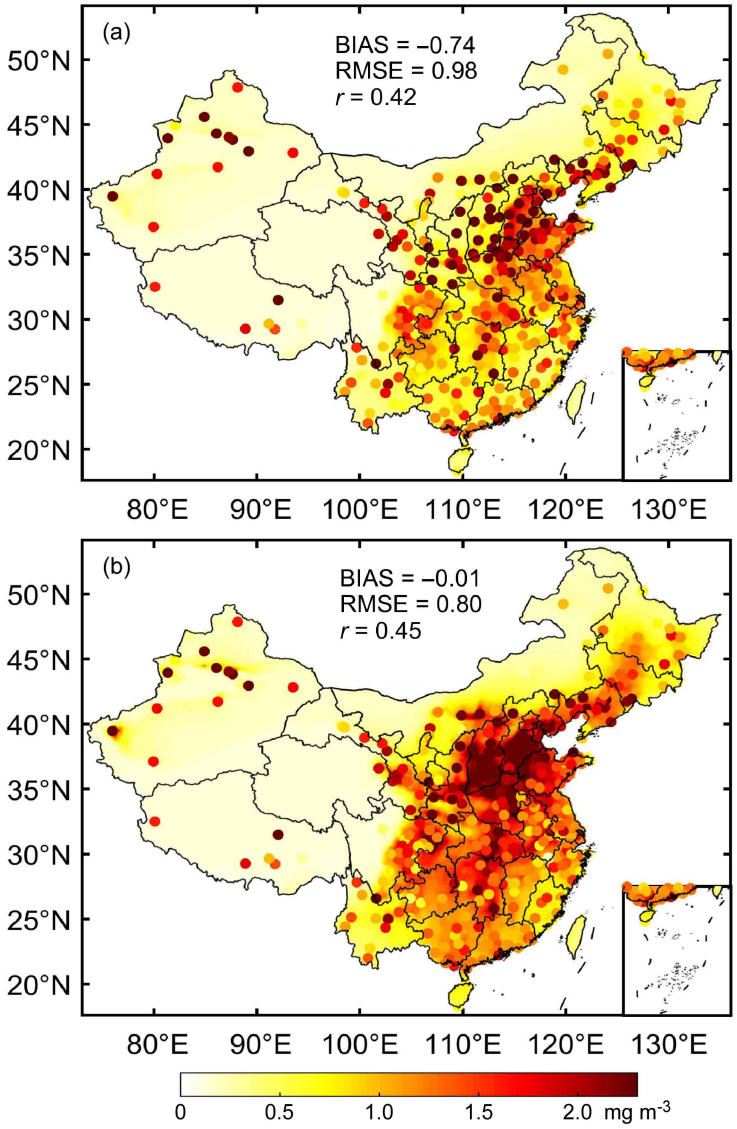

图 3　（a）反演前和（b）反演试验 1 得到的 2015 年 1 月 CO 模拟浓度分布。反演试验 1 使用 2014 年 1 月的集合数据集，结合反演前 2015 年 1 月的模拟浓度误差，估计 2015 年 1 月的 CO 排放。圆点颜色表示验证站点的观测月均浓度。验证站点的平均偏差（BIAS）、均方根误差（RMSE）和标准差（*r*）列于图中

Fig. 3　Simulated CO concentrations (a) before and (b) after inversion test 1 in Jan 2015. Inversion test 1 estimates the CO emission in Jan 2015 with the ensemble dataset of Jan 2014 and the simulated error of Jan 2015 before inversion. Colors of circles indicate the observed monthly mean concentrations at validation sites. Biases (BIAS), root-mean-square errors (RMSE), and correlations (*r*) over all validation sites are also shown

日的虚假峰值。

为减小计算量，本文的源反演方法忽略了历史时段与反演时段的误差协方差差异，使用历史集合数据集构建表征排放与浓度切线性关系的误差协方差矩阵。历史时段与反演时段的气象条件差异可影响误差协方差矩阵，进而影响反演结果。然而，反演中使用的污染物平均浓度仍取自反演时段的观测和模拟结果，因此反演结果可以体现反演时段的气象和污染特征，从而限制两个时段气象条件差异对结果的影响。为定量评估该影响，本文在反演试验 2 中使用 2015 年 1 月的集合数据集再次构建误差协方差矩阵并反演排放源。结果发现，使用 2015 年 1 月集合数据集反演的全国 CO 排放量为 46.0 Tg，仅比使用 2014 年 1 月集合数据集反演的结果小 1%。在太原市和广州市的验证站点中（图 5），两套反演源的模拟均方根误差差异均小于 0.01 mg m^{-3}。在全国 349 个验证站点中，使用反演时段集合数据集的平均均方根误差为 0.78 mg m^{-3}，仅比

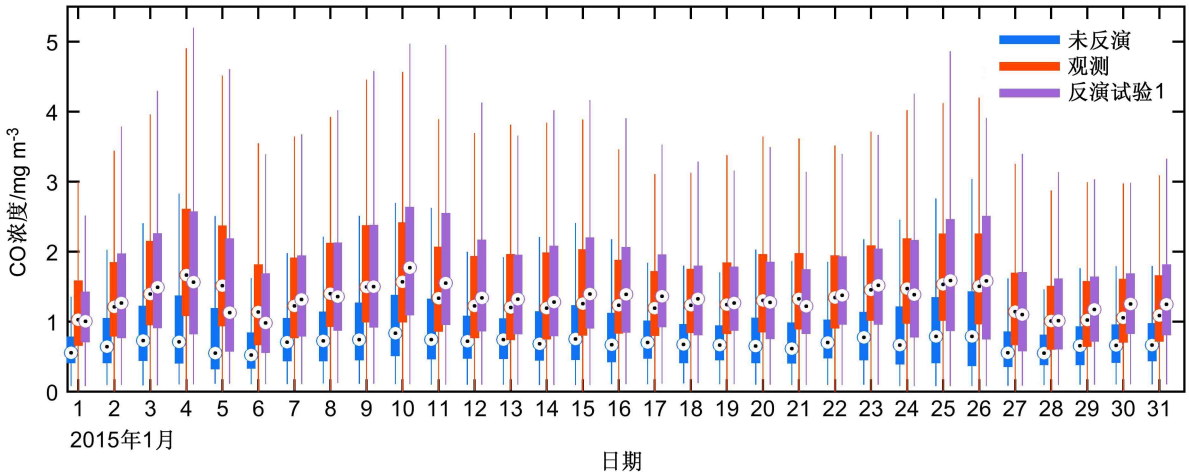

图 4 反演前后 2015 年 1 月全国验证站点的每日 CO 浓度箱型图（圆点表示中位数，箱形的顶部和底部分别表示 75 和 25 百分位数）

Fig. 4 Boxplot of daily CO concentrations before and after emission inversion over all validation sites in Jan 2015 (central mark indicates the median, and the top and bottom edges of the box indicate the 25th and 75th percentiles, respectively)

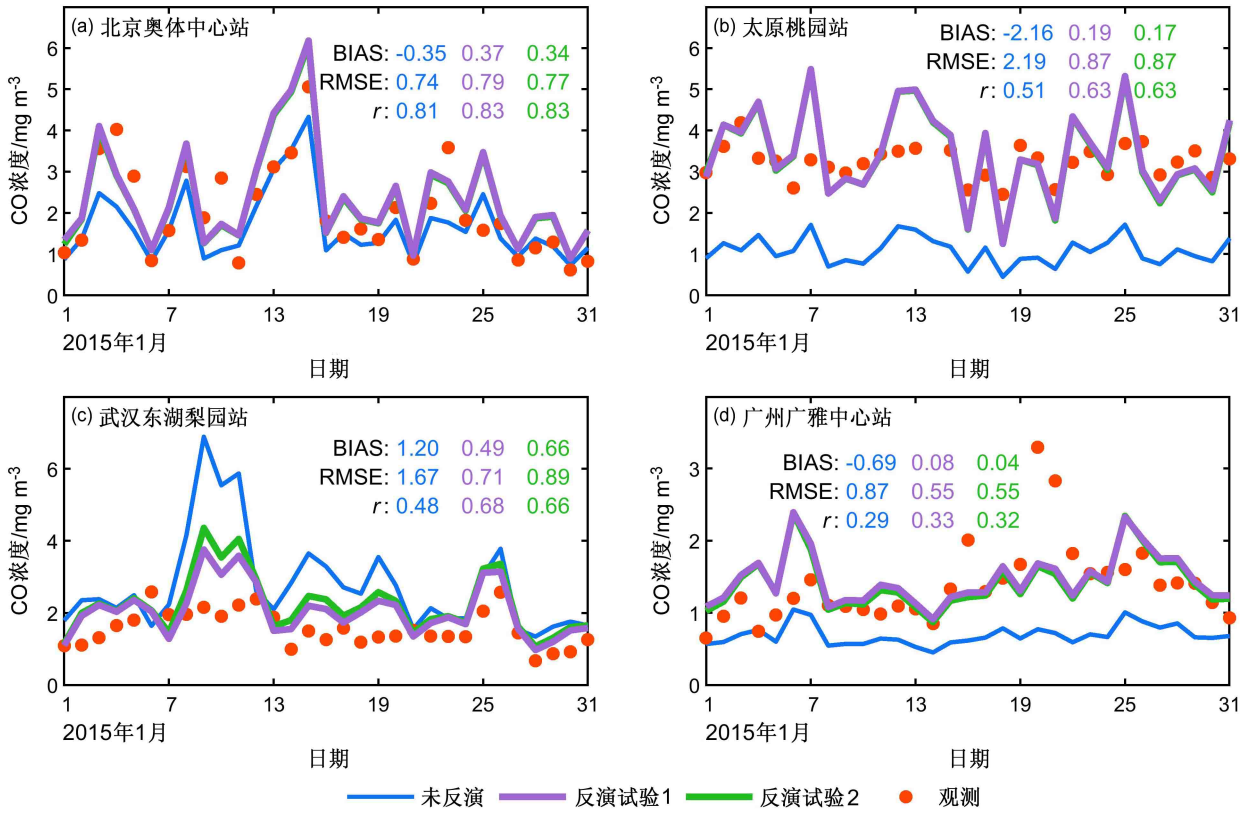

图 5 反演前后站点的观测和模拟时间序列。反演试验 1 和反演试验 2 分别表示使用 2014 年 1 月和 2015 年 1 月集合数据集进行反演，得到的 2015 年 1 月模拟结果。各站点的偏差（BIAS）、均方根误差（RMSE）和标准差（r）列于图中

Fig. 5 Time series of simulated and observed CO concentrations before and after emission inversion with corresponding biases (BIAS), root-mean-square errors (RMSE), and correlations (r). Inversion test 1 and inversion test 2 represent the concentrations in Jan 2015 by the inversions with ensemble dataset of Jan 2014 and Jan 2015, respectively

使用历史集合数据集低 0.02 mg m^{-3}，占反演前均方根误差的 2%。以上结果表明，2014 年 1 月与 2015 年 1 月的气象差异对反演结果影响有限，使用历史集合数据集代替反演时段集合数据集可有效反演排放源。

综上所述，使用先验排放源模拟的 2015 年 1

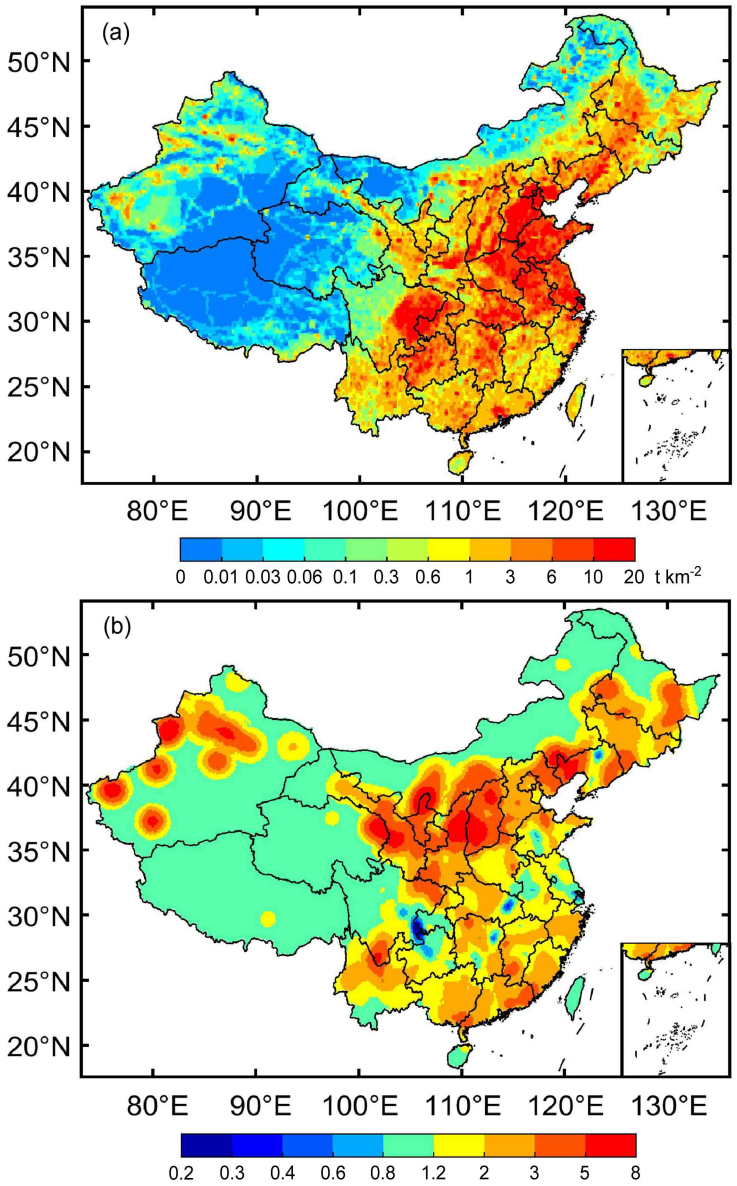

图 6　（a）CO 先验排放强度与（b）反演调整系数

Fig. 6　(a) A priori CO emissions and (b) adjustment factors from inversion

月 CO 浓度存在显著偏差。本文使用基于 EnOI 的源反演方法同化站点观测数据，结合 2014 年 1 月的历史集合数据集，快速反演全国 15 km 分辨率的 CO 排放。该方案反演的全国 CO 排放仅比使用 2015 年 1 月集合数据集反演的结果高 1%，表明历史时段与反演时段的气象条件差异只能有限地影响月均 CO 排放的反演结果。使用反演的排放源再次模拟，可大幅降低 CO 模拟偏差，同时显著降低均方根误差，提高相关系数，表明基于 EnOI 的快速源反演方法可有效降低排放源的不确定性，提高空气质量的模拟和预报精度。

4　结论

本文发展了一种基于 EnOI 的大气排放源反演方法。该方法根据模拟的浓度误差调整排放强度，可降低排放清单的不确定性，为空气质量预报快速更新排放源。相比基于 EnKF 的源反演方法，本文方法无需多次使用大气化学传输模式模拟反演时段，仅需一次未扰动排放源的常规模拟即可快速更新排放清单，将排放源反演计算量降低至少一个数量级。

本文将该方法应用于 2015 年 1 月全国 15 km

分辨率的 CO 排放源反演。相比基于 2010 年 1 月的先验 CO 排放，反演的全国 2015 年 1 月 CO 排放比先验排放源高 120%，且排放强度在大部分地区均有所增强。在 CO 模拟浓度被显著低估的山西、新疆等地，反演的排放强度可达先验源的 3 倍以上。使用 2014 年 1 月集合数据集反演的全国 CO 排放仅比使用 2015 年 1 月集合数据集的反演结果高 1%，表明历史时段与反演时段的气象条件差异只能有限地影响月均 CO 排放的反演结果。为评估排放源反演效果，本文对比了反演前后全国 349 个独立验证站点的 CO 浓度，发现使用历史数据集反演的排放源可显著降低模拟偏差，将全国 349 个独立验证站点的平均低估从 0.74 mg m^{-3} 降至 0.01 mg m^{-3}，同时将均方根误差降低 18%。

本文将基于 EnOI 的污染源反演方法成功应用于 CO 排放源反演，后续研究可尝试将其应用于黑碳、NO_2 和一次颗粒物等已使用 EnKF 成功反演的物种（Peng et al., 2017; Miyazaki et al., 2017; Houweling et al., 2017）。然而作为 EnKF 的次优方案，EnOI 未使用观测数据在线更新排放源。排放源的误差可影响线性化的误差协方差估计，故而使用本文方法反演强非线性的污染源可能存在较大的不确定性（如使用臭氧浓度反演挥发性有机物排放）。此外，本文方法使用历史集合数据集估计反演时段的误差协方差矩阵，两时段的气象条件差异可影响反演结果，使其难以适用于极端天气或高时间分辨率（如一天）的污染源反演。尽管难以适用于上述极端情形，本文方法可将污染源反演的计算量降低一个数量级并且适用于多种污染物的月均排放源反演。此外，为进一步提高反演精度，可使用更新的排放源再次模拟，并根据新的模拟浓度误差再次使用基于 EnOI 的源反演方法，通过反复迭代以应对源反演中普遍存在的非线性和偏差问题。后续研究可进一步对比相同计算量下，基于 EnKF 的源反演方法与基于 EnOI 的迭代源反演的反演精度，以做出更优选择。

参考文献（References）

Beekmann M, Derognat C. 2003. Monte Carlo uncertainty analysis of a regional-scale transport chemistry model constrained by measurements from the atmospheric pollution over the Paris area (ESQUIF) campaign [J]. J. Geophys. Res. Atmos., 108(D17): 8559. doi:10.1029/2003JD003391

曹国良, 张小曳, 龚山陵. 2011. 中国区域主要颗粒物及污染气体的排放源清单 [J]. 科学通报, 56(3): 261−268. Cao Guoliang, Zhang Xiaoye, Gong Shanling. 2011. Emission inventories of primary particles and pollutant gases for China [J]. Chinese Science Bulletin (in Chinese), 56(3): 261−268. doi:10.1007/s11434-010-4373-7

Chang J S, Brost R A, Isaksen I S A, et al. 1987. A three‐dimensional Eulerian acid deposition model: Physical concepts and formulation [J]. J. Geophys. Res. Atmos., 92(D12): 14681−14700. doi:10.1029/JD092iD12p14681

Counillon F, Bertino L. 2009. Ensemble Optimal Interpolation: Multivariate properties in the Gulf of Mexico [J]. Tellus A: Dynamic Meteorology and Oceanography, 61(2): 296−308. doi:10.1111/j.1600-0870.2008.00383.x

Dekker I N, Houweling S, Aben I, et al. 2017. Quantification of CO emissions from the city of Madrid using MOPITT satellite retrievals and WRF simulations [J]. Atmospheric Chemistry and Physics, 17(23): 14675−14694. doi:10.5194/acp-17-14675-2017

Evensen G. 2003. The ensemble Kalman filter: Theoretical formulation and practical implementation [J]. Ocean Dynamics, 53(4): 343−367. doi:10.1007/s10236-003-0036-9

Evensen G. 2004. Sampling strategies and square root analysis schemes for the EnKF [J]. Ocean Dynamics, 54(6): 539−560. doi:10.1007/s10236-004-0099-2

Gaspari G, Cohn S E. 1999. Construction of correlation functions in two and three dimensions [J]. Quart. J. Roy. Meteor. Soc., 125(554): 723−757. doi:10.1002/qj.49712555417

Hanna S R, Chang J C, Fernau M E. 1998. Monte Carlo estimates of uncertainties in predictions by a photochemical grid model (UAM-IV) due to uncertainties in input variables [J]. Atmos. Environ., 32(21): 3619−3628. doi:10.1016/S1352-2310(97)00419-6

Hao M J, Tian H Z, Lu Y Q. 2002. Emission inventories of NO_x from commercial energy consumption in China, 1995−1998 [J]. Environ. Sci. Technol., 36(4): 552−560. doi:10.1021/es015601k

Holloway T, Levy II H, Kasibhatla P. 2000. Global distribution of carbon monoxide [J]. J. Geophys. Res. Atmos., 105(D10): 12123−12147. doi:10.1029/1999JD901173

Houweling S, Bergamaschi P, Chevallier F, et al. 2017. Global inverse modeling of CH_4 sources and sinks: An overview of methods [J]. Atmospheric Chemistry and Physics, 17(1): 235−256. doi:10.5194/acp-17-235-2017

Kaurkin M N, Ibrayev R A, Belyaev K P. 2016. ARGO data assimilation into the ocean dynamics model with high spatial resolution using Ensemble Optimal Interpolation (EnOI) [J]. Oceanology, 56(6): 774−781. doi:10.1134/S0001437016060059

Kurokawa J, Ohara T, Morikawa T, et al. 2013. Emissions of air pollutants and greenhouse gases over Asian regions during 2000−2008: Regional Emission inventory in ASia (REAS) version 2 [J]. Atmospheric Chemistry and Physics, 13(21): 11019−11058. doi:10.5194/acp-13-11019-2013

Li M, Zhang Q, Kurokawa J I, et al. 2017. MIX: A mosaic Asian anthropogenic emission inventory under the international collaboration framework of the MICS-Asia and HTAP [J].

Atmospheric Chemistry and Physics, 17(2): 935−963. doi:10.5194/acp-17-935-2017

李飞, 唐晓, 王自发, 等. 2019. 基于京津冀高密度地面观测网络的大气污染物浓度地面观测代表性误差估计 [J]. 大气科学, 43(2): 277−284. Li Fei, Tang Xiao, Wang Zifa, et al. 2019. Estimation of representative errors of surface observations of air pollutant concentrations based on high-density observation network over Beijing–Tianjin–Hebei region [J]. Chinese Journal of Atmospheric Sciences (in Chinese), 43(2): 277−284. doi:10.3878/j.issn.1006-9895.1804.17267

Miyazaki K, Eskes H, Sudo K, et al. 2017. Decadal changes in global surface NO_x emissions from multi-constituent satellite data assimilation [J]. Atmospheric Chemistry and Physics, 17(2): 807−837. doi:10.5194/acp-17-807-2017

Moore G E, Londergan R J. 2001. Sampled Monte Carlo uncertainty analysis for photochemical grid models [J]. Atmos. Environ., 35(28): 4863−4876. doi:10.1016/S1352-2310(01)00260-6

Naeher L P, Smith K R, Leaderer B P, et al. 2001. Carbon monoxide as a tracer for assessing exposures to particulate matter in wood and gas cookstove households of highland Guatemala [J]. Environ. Sci. Technol., 35(3): 575−581. doi:10.1021/es991225g

Nenes A, Pandis S N, Pilinis C. 1998. ISORROPIA: A new thermodynamic equilibrium model for multiphase multicomponent inorganic aerosols [J]. Aquatic Geochemistry, 4: 123−152. doi:10.1023/A:1009604003981

Oke P R, Brassington G B, Griffin D A, et al. 2010. Ocean data assimilation: A case for ensemble optimal interpolation [J]. Australian Meteorological and Oceanographic Journal, 59: 67−76. doi:10.22499/2.5901.008

Peng Z, Liu Z Q, Chen D, et al. 2017. Improving PM2.5 forecast over China by the joint adjustment of initial conditions and source emissions with an ensemble Kalman filter [J]. Atmospheric Chemistry and Physics, 17(7): 4837−4855. doi:10.5194/acp-17-4837-2017

Sakov P, Oke P R. 2016. A deterministic formulation of the ensemble Kalman filter: An alternative to ensemble square root filters [J]. Tellus A: Dynamic Meteorology and Oceanography, 60(2): 361−371. doi:10.1111/j.1600-0870.2007.00299.x

Streets D G, Canty T, Carmichael G R, et al. 2013. Emissions estimation from satellite retrievals: A review of current capability [J]. Atmos. Environ., 77: 1011−1042. doi:10.1016/j.atmosenv.2013.05.051

Sun A Y, Morris A, Mohanty S. 2009. Comparison of deterministic ensemble Kalman filters for assimilating hydrogeological data [J]. Advances in Water Resources, 32(2): 280−292. doi:10.1016/j.advwatres.2008.11.006

Tang X, Zhu J, Wang Z F, et al. 2013. Inversion of CO emissions over Beijing and its surrounding areas with ensemble Kalman filter [J]. Atmos. Environ., 81: 676−686. doi:10.1016/j.atmosenv.2013.08.051

Wang P, Wang H, Wang Y Q, et al. 2016. Inverse modeling of black carbon emissions over China using ensemble data assimilation [J]. Atmospheric Chemistry and Physics, 16(2): 989−1002. doi:10.5194/acp-16-989-2016

王自发, 谢付莹, 王喜全, 等. 2006. 嵌套网格空气质量预报模式系统的发展与应用 [J]. 大气科学, 30(5): 778−790. Wang Zifa, Xie Fuying, Wang Xiquan, et al. 2006. Development and application of nested air quality prediction modeling system [J]. Chinese Journal of Atmospheric Sciences (in Chinese), 30(5): 778−790. doi:10.3878/j.issn.1006-9895.2006.05.07

王自发, 庞成明, 朱江, 等. 2008. 大气环境数值模拟研究新进展 [J]. 大气科学, 32(4): 987−995. Wang Zifa, Pang Chengming, Zhu Jiang, et al. 2008. IAP progress in atmospheric environment modeling research [J]. Chinese Journal of Atmospheric Sciences (in Chinese), 32(4): 987−995. doi:10.3878/j.issn.1006-9895.2008.04.23

魏巍, 王书肖, 郝吉明. 2011. 中国人为源 VOC 排放清单不确定性研究 [J]. 环境科学, 32(2): 305−312. Wei Wei, Wang Shuxiao, Hao Jiming. 2011. Uncertainty analysis of emission inventory for volatile organic compounds from anthropogenic sources in China [J]. Environmental Science (in Chinese), 32(2): 305−312. doi:10.13227/j.hjkx.2011.02.021

Wu H J, Tang X, Wang Z F, et al. 2018. Probabilistic automatic outlier detection for surface air quality measurements from the China national environmental monitoring network [J]. Advances in Atmospheric Sciences, 35(12): 1522−1532. doi:10.1007/s00376-018-8067-9

Zaveri R A, Peters L K. 1999. A new lumped structure photochemical mechanism for large-scale applications [J]. J. Geophys. Res. Atmos., 104(D23): 30387−30415. doi:10.1029/1999JD900876

张金谱, 胡嘉镗, 王雪梅. 2014. 集合最优插值同化方法在珠三角空气质量模拟中的初步应用 [J]. 环境科学学报, 34(3): 558−566. Zhang Jinpu, Hu Jiatang, Wang Xuemei. 2014. Preliminary application of ensemble optimal interpolation data assimilation method on air quality numerical modeling in the Pearl River Delta [J]. Acta Scientiae Circumstantiae (in Chinese), 34(3): 558−566. doi:10.13671/j.hjkxxb.2014.0103

Zhang Q, Streets D, Carmichael G, et al. 2009. Asian emissions in 2006 for the NASA INTEX-B mission [J]. Atmospheric Chemistry and Physics, 9: 5131−5153. doi:10.5194/acp-9-5131-2009

朱江, 汪萍. 2006. 集合卡尔曼平滑和集合卡尔曼滤波在污染源反演中的应用 [J]. 大气科学, 30(5): 871−882. Zhu Jiang, Wang Ping. 2006. Ensemble Kalman smoother and ensemble Kalman filter approaches to the joint air quality state and emission estimation problem [J]. Chinese Journal of Atmospheric Sciences (in Chinese), 30(5): 871−882. doi:10.3878/j.issn.1006-9895.2006.05.16