

Impact of data assimilation and aerosol radiation interaction on Lagrangian particle dispersion modelling

Mengwei Jia ^{a,b}, Xin Huang ^{a,b}, Ke Ding ^{a,b}, Qiang Liu ^{a,b}, Derong Zhou ^{a,b}, Aijun Ding ^{a,b,*}

^a School of Atmospheric Sciences, Nanjing University, Nanjing, 210023, China

^b Collaborative Innovation Center of Climate Change, Jiangsu Province, Nanjing, 210023, China

HIGHLIGHTS

- LPDMs driven by reanalysis datasets underestimate pollutant concentration.
- The underestimation can be minimized by assimilating meteorological data.
- Considering the feedbacks of aerosols to meteorology increases the accuracy of LPDMs.

ARTICLE INFO

Keywords:

Lagrangian particle dispersion modelling
WRF-FLEXPART
Eastern China
Data assimilation
Aerosols-PBL feedback

ABSTRACT

Lagrangian particle dispersion models (LPDMs) have been widely used in air pollution studies. However, substantial uncertainties still exist in LPDM simulations due to biased meteorological data, especially under stagnant and highly-polluted conditions. In this work, to better investigate the source contribution and formation of winter haze pollution in eastern China, we conduct a sensitivity study of WRF-FLEXPART by using different reanalysis data, applying observational meteorological nudging, and considering aerosols' radiative feedback on meteorology. We find that simulations driven by reanalysis datasets generally underestimate pollutant concentration, especially during periods with heavy haze pollution. The underestimation is directly caused by overestimated planetary boundary layer (PBL) height and lower PBL horizontal wind speeds. By assimilating meteorological data from surface and radiosonde observation, the WRF model can well represent the PBL dynamics and wind fields, especially those near the ground surface, which then substantially improves particle tracing in the LPDM. In addition, by including aerosols' radiative feedback in the WRF-Chem model, which significantly influences PBL evolution, the biases between LPDM modelling and observations are notably narrowed, particularly when the haze pollution is severe. Quantitatively, the accuracy increase of the simulations with aerosols' radiative effect accounted for 48% of the improvement produced by assimilating meteorological data. Overall, meteorological input is of great importance in LPDM modelling. In regions with intensive pollution like China and India, applying observational data assimilation or considering the feedbacks of aerosols to meteorology serve as an effective way to reduce the biases of LPDMs and better understand the source contributions as well as the formation and accumulation of pollution.

1. Introduction

Air pollution has become a great challenge facing many emission-intensive regions like China and India. Based on the Lagrangian formulation or the Eulerian equations, models that resolve atmospheric transport and diffusion processes have been developed to investigate and comprehensively understand the spatiotemporal variations and source contributions of pollutants (Byers, 1974; Dutton, 1986; Zhang

et al., 2016; Huang et al., 2016; Zhou et al., 2018). The Eulerian model is a grid-based model, in which coordinates are fixed in space and time. It can take not only transport and emissions into considerations, but also the atmospheric chemistry (Nguyen et al., 1997; Peters et al., 1995; Huang et al., 2016; Liu et al., 2016). Most of the mesoscale air quality models are Eulerian models, such as the Weather Research and Forecasting model coupled with Chemistry (WRF-Chem), the Community Multiscale Air Quality (CMAQ), and so on. However, the Eulerian

* Corresponding author. Collaborative Innovation Center of Climate Change, Jiangsu Province, Nanjing, 210023, China.
E-mail address: dingaj@nju.edu.cn (A. Ding).

approach is limited by large computational costs and the numerical diffusion limitations of the grid system. The Lagrangian approach uses individual air parcels to represent the mass of trace gas emissions as they move through time and space (Draxler and Hess, 1998; Lai and Chen, 2007; Stein et al., 2015; Stohl, 1998). It is capable of preserving tracer gradients at much smaller spatial scales, and hence the fine-scale structures of the flow can be well resolved. For long-term studies and assimilating observations to produce accurate analysis maps of chemical species (De Leeuw et al., 1990; Fisher and Lary, 1995; Stohl and Trickl et al., 1999; Stohl et al., 2003; Halse et al., 2013), Lagrangian particle dispersion models (LPDMs) pose smaller computational demands than Eulerian models (Stohl et al., 1998; Stein et al., 2015). The chemistry is linear in the LPDMs, so most aqueous phase chemistry reactions are not considered. For some very complex chemical reaction schemes, LPDMs were adopted for allowing separation of dynamics and chemistry, which considerably reduces the computational expense of the method (Derwent, 1990).

Owing to the advantages described above, LPDMs have been extensively utilized in the atmospheric sciences. By calculating the trajectories of thousands of released particles, LPDMs are able to simulate the transport and diffusion processes with considerations of advection and convection as well as turbulent motions (Ding et al., 2009; Zhang et al., 2009; Mahmud et al., 2012). At present, the FLEXible PARTicle dispersion model (FLEXPART) (Stohl et al., 2005) is one of the most widely used LPDMs (Stein et al., 2015) and has been more widely used for dispersion simulations (e.g. Forster et al., 2001; Spichtinger et al., 2001; Damoah et al., 2004; Ding et al., 2009; Decarlo et al., 2010; James et al., 2003; Stohl et al., 2002; Eckhardt et al., 2003; Stohl and James et al., 2004). Previous studies confirm the reliability of FLEXPART for the simulation of emission sources. Owing to their higher spatial resolution, the outputs of mesoscale meteorological models, such as the Weather Research and Forecasting Model (WRF) model, are often used as the meteorological driving fields for LPDMs (Fast and Easter, 2006; Ding et al., 2009; Brioude, 2011).

However, the model performance of LPDMs still needs to be evaluated in detail, especially the sensitivity to meteorological outputs under highly polluted conditions. The simulation biases can mainly be attributed to uncertainties in the meteorological driving fields and emission sources, rather than differences between LPDMs (Angevine et al., 2014; Hegarty et al., 2013). Studies have shown that horizontal wind and vertical mixing are the key factors to the propagation of uncertainty from meteorological fields to tracers fluxes (Lin and Gerbig, 2005; Gerbig, 2008), and are influenced by the Planetary Boundary Layer (PBL) scheme (Srinivas et al., 2016; Madala et al., 2015), the spatial resolution of meteorological fields (Hernández-Ceballos et al., 2014), and the driving datasets (Su et al., 2015; Daria et al., 2017; Harris et al., 2005). An et al. (2014) confirmed the similar underestimations of CO concentration in WRF-FLEXPART driven by either the European Centre for Medium-Range Weather Forecasts (ECMWF) or National Centers for Environmental Prediction (NCEP) Global Forecast System (GFS) data. Hoffmann et al. (2019) evaluated the impact of different ECMWF reanalysis datasets on LPDM simulations and found improvements of simulation results when using the Fifth Generation Atmospheric Reanalysis of the Global Climate (ERA5), compared with the older ERA-Interim reanalysis. Although ERA5 and NCEP FNL (Final) global analysis are the commonly used reanalysis datasets, there is no comparative study concerning the differences in WRF-FLEXPART simulations driven by them. Besides, the accuracy rate of model simulation using ERA5 or FNL is still a challenging issue in mesoscale numerical modelling, especially under stagnant weather conditions. Previous studies have shown that meteorological data assimilation can improve the accuracy of simulations even in stagnant weather conditions (Li et al., 2016; Ngan et al., 2012). Besides, several recent modelling and observational studies have revealed the effects of aerosols on the evolution of the PBL, especially under highly polluted conditions (Ding et al., 2016a; Liu et al., 2016; Wang et al., 2018; Huang et al., 2016).

Aerosols can also impact the large scale atmospheric circulation, which is also important to severe haze pollution (Yang et al., 2017a; Lou et al., 2019). Specifically, high concentrations of light-absorbing aerosols could greatly reduce incoming solar radiation, decrease the surface heat flux, heat the atmosphere, and stabilize the temperature stratification, thereby suppressing daytime PBL development and in turn enhancing haze pollution through decreased vertical dispersion (Liu et al., 2016; Wang et al., 2018; Huang et al., 2018, 2020; Zhou et al., 2018). These factors inducing notable model biases in LPDMs are particularly important in regions subject to severe haze pollution, such as eastern China. However, quantifying the influence from meteorology is a key issue which has not been fully addressed in this region.

Resulting from population explosion and rapid development, anthropogenic emissions, including industry, vehicle and residential activities, and power plants, have increased dramatically in eastern China (Liu et al., 2008; Ding et al., 2019). The Yangtze River Delta (YRD) region, one of the most economically developed regions in eastern China, has been suffering from severe air pollution in the past several decades (Ding et al., 2017, 2019; Wang et al., 2019; Huang et al., 2020). During the cold seasons, under the influence of the Siberian high pressure, stagnant weather frequently occurs in the YRD region. Intensive anthropogenic emissions and unfavorable dispersion conditions often lead to the formation and accumulation of haze pollution.

To better understand the formation of haze pollution and emission attribution in this highly-polluted region, we aim to optimize LPDM modelling in this work on the biases of WRF(Chem)-FLEXPART simulations. More specifically, we attempt to improve the meteorological driving fields for simulations under stagnant weather conditions by assimilating observational data and considering aerosol radiative interaction, and then carry out quantitative estimations of the model improvements based on emission inventory and observations. As an indicator of direct emissions, carbon monoxide (CO) was used as the tracer species. The rest of this paper is structured as follows: Sect. 2 describes the observations, reanalysis data, and modelling methodology. We evaluate the effect on WRF-FLEXPART simulations of the choice of reanalysis data (Sect. 3.1), the influence of meteorological assimilation (Sect. 3.2), and the role of aerosol-PBL feedback (Sect. 3.3). Finally, conclusions are drawn in Sect. 4.

2. Methodology

To understand the formation of haze pollution and emission attribution in eastern China, the WRF-FLEXPART model was run with different reanalysis data, meteorological assimilation, and the aerosols' radiative feedback. Four sensitivity experiments were conducted (Table 1) and run in parallel. The WRF simulations were driven by ERA5 (EXP-ERA5) and FNL (EXP-FNL) reanalysis data. Other than nudging with NCEP ADP Global Upper Air Observational Weather Data (<https://rda.ucar.edu/datasets/ds351.0/#!>), the EXP-FNL* and the EXP-FNL experiments are identical. The aerosol-PBL interaction, which can be reflected by the differences between radiosonde observations and reanalysis data, has been shown to be significant on polluted days (Ding et al., 2013a, 2016a; Huang et al., 2018, 2020). To evaluate the effect of the aerosol-PBL radiative forcing interaction on atmospheric transport and diffusion, the feedback of aerosols to radiation schemes was turned

Table 1

Driving fields and their models in the sensitivity experiments.

Experiment	Driving data	Model	Nudging with observations	Aerosol coupling
EXP-ERA5	ERA5	WRF	Off	Off
EXP-FNL	FNL	WRF	Off	Off
EXP-FNL*	FNL	WRF	On	Off
EXP-CHEM	FNL	WRF/ Chem	Off	On

on in the WRF-Chem model in EXP-CHEM. All the other experimental settings were coincident between the four experiments.

2.1. Algorithm and statistics

Source attribution method has been implemented and used in many models (Stohl et al., 2006; Wang et al., 2014; Yang et al., 2017b). As described by Stohl et al. (2006), the outputs of the FLEXPART model were utilized to calculate the potential source contribution (PSC), which provide information on the concentration of a specified tracer over a region without consideration of loss processes or chemical transformation. CO was used as the tracer species for WRF (Chem)-FLEXPART. The primary sources of CO are direct industrial emission, transportation, forest burning and indoor heating. It is difficult to dissolve in water and to liquefy or solidify in the ambient atmosphere. Its main consumption in the atmosphere is to react with OH to produce ozone with a relatively stable reaction rate. For these reasons, although CO is not a completely inert tracer, it is always considered as the tracer of direct anthropogenic emission.

In this research, the virtual particles of CO were released at the height of 10 m above ground. The residence time for a thickness of 100 m above ground level for each receptor was calculated and considered as the “footprint” retroplume (Kljun et al., 2002; Stohl et al., 2007; Ding et al., 2009). It is on behalf of the probability distribution or residence time of a simulated air mass. The footprint retroplumes in WRF-FLEXPART were multiplied by the emission inventories to obtain the PSCs (Ding et al., 2013b). The spatial distribution of the PSCs was used to characterize the source-receptor relationship of air pollutants. The residence times of the specified tracers were aggregated over the geographic region shown as domain 2 (d02) in Fig. 1. However, the sum of 7-day backward PSC did not fully account for all contributions of emission transports. In fact, particles whose lifetimes are greater than seven days are not taken into account (Stohl et al., 2006; Ding et al., 2013b). The lifetime of CO in the troposphere is about 1–2 months, so the final simulated CO concentration should be added to a “baseline” (Ding et al., 2013b). The baseline adopted here was the lowest concentration in winter of the past three years, which approximatively represented the local retention of pollutants directly emitted in the winter atmosphere. In general, the PSC and final simulated CO concentration (SIM) are defined as:

$$PSC(i,j) = \text{footprint} \times \text{emission}/\text{height} \quad (1)$$

$$SIM = \text{baseline} + k \times \sum PSC \quad (2)$$

Here k is a flexible coefficient considering the existence of unit transformation for different emission inventories.

By comparing the spatiotemporal variation of the SIM with the observed value of pollutants, the accuracy of particle traceability can be assessed intuitively. The mean absolute error (MAE), root mean square error (RMSE), correlation coefficient (r), mathematical expectation value (μ), and accuracy rate (AR) were introduced to assess the predictions, where obs is the observed CO concentration.

- (1) The mean absolute error was computed, which is a widely used simple measure of error (MAE):

$$\text{MAE} = \frac{\sum_{i=1}^n |SIM_i - obs_i|}{\sum_{i=1}^n obs_i} \quad (3)$$

- (2) The root mean square error (RMSE) reflects the dispersion between the simulated values and the observed values:

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (SIM_i - obs_i)^2} \quad (4)$$

- (3) Pearson's correlation coefficient represents the linear correlation between variables:

$$r = \frac{n \sum_{i=1}^n SIM_i obs_i - \sum_{i=1}^n SIM_i \sum_{i=1}^n obs_i}{\sqrt{n \sum_{i=1}^n SIM_i^2 - \left(\sum_{i=1}^n SIM_i \right)^2} \sqrt{n \sum_{i=1}^n obs_i^2 - \left(\sum_{i=1}^n obs_i \right)^2}} \quad (5)$$

- (4) The mean value of errors (μ):

$$\mu = \frac{1}{n} \sum_{i=1}^n (SIM_i - obs_i) \quad (6)$$

- (5) The accuracy rate (AR) is defined as:

$$AR = \frac{1}{n} \sum_{i=1}^n \left(1 - \frac{|SIM_i - obs_i|}{obs_i} \right) \quad (7)$$

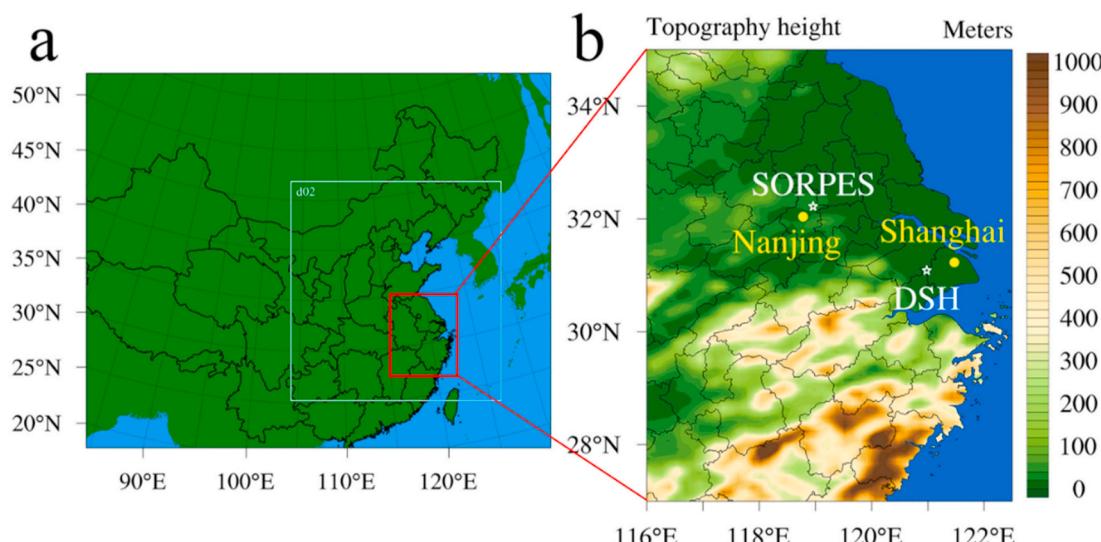


Fig. 1. WRF domain setting (a) and topographic field in the YRD region (b).

2.2. Model configurations and evaluation

Nanjing and Shanghai, two of the largest cities in the YRD region, were selected as the sites of interest (Fig. 1b). Both the Station for Observing Regional Processes of the Earth System (SORPES) at Nanjing University (Ding et al., 2016b, 2019) and the Shanghai Dianshan Lake Station (DSH) (Li et al., 2019) are located in the suburbs, which is more suitable for model sensitivity research than the urban sites with complex emission and underlying properties. We ran seven-day backward simulations for each hour from December 5, 2017 to January 4, 2018. WRF-FLEXPART version 3.3 was set up with releases of 3000 computational particles under the footprint level (100 m). The simulation domain was from 105°E to 126°E and 25.5°N to 43.5°N at a 0.1° × 0.1° spatial resolution. The 2014 Peking University (PKU) global emission inventories (Zhong et al., 2017, <http://inventory.pku.edu.cn/home.html>) with a spatial resolution of 0.1° × 0.1° were used as emission in formula (1). A summary of the WRF-FLEXPART configurations and parameter values is given in Table 2.

The WRF version 3.9.1 simulations were driven by ERA5 and FNL reanalysis data every 6 h to generate the hourly meteorological fields for the FLEXPART model. The time steps of the WRF module and the FLEXPART module were 60 s and 600 s separately. The WRF simulations were performed in two domains (Fig. 1): the first domain had a 30-km horizontal resolution with 169 × 136 grid cells and covered most of China; and the second domain had 10 km horizontal resolution with 229 × 238 grid cells covering eastern China. Thirty-three vertical layers for all the grid meshes were used from the ground level to the top pressure of 100 hPa. The detailed domain settings and configuration options are exhibited in Table 3. The second domain with a horizontal resolution of 10 km was configured to be close to the WRF-FLEXPART grid area.

In the EXP-FNL*, the observation nudging data assimilation was used to gradually nudge the WRF model simulations towards meteorological observations every 6 h. NCEP ADP Global Upper Air Observational Weather Data is composed of global upper air weather reports operationally collected by the National Centers for Environmental Prediction (NCEP) (<https://doi.org/10.5065/39C5-Z211>, Satellite Services Division et al., 2004, last access: 14 October 2020). These include radiosondes, pibals, and aircraft reports from the Global Telecommunications System (GTS), and satellite data from the National Environmental Satellite Data and Information Service (NESDIS). They contain up to 20 mandatory levels data, including atmospheric pressure measurements, geopotential height, upper air temperature, dew point temperature, upper level winds and surface winds from platform observations. The differences between the model and observations were used to create an innovation that is multiplied by various factors and added to model tendency equations to gradually nudge the model towards observations (Stauffer and Seaman 1994; Deng et al., 2009) (available here: <http://www2.mmm.ucar.edu/wrf/users/wrfv3.1/nudging.html>, last access: 14 October 2020). It can be defined as the equation:

Table 3
WRF-Chem configuration options and settings.

Domain settings	1	2
Horizontal grid	169 × 136	229 × 238
Grid spacing	30 km	10 km
Vertical layers in total	33	
Vertical layers from surface to 800 hPa	22	
Configuration options		
Boundary layer	Yonsei University	
Cumulus parameterization	Grell-Freitas	
Microphysics	Lin et al.	
Longwave radiation	RRTMG	
Shortwave radiation	RRTMG	
Surface layer	Revised MM5 Monin-Obukhov	
Land surface	Noah	
Gas-phase chemistry	CBMZ	
Aerosol chemistry	MOSAIC	

$$\frac{\partial q_p}{\partial t} = \frac{\sum_{i=1}^N W_q^2(i, x, y, z, t)[q_o(i) - q_m(x_i, y_i, z_i, t)]}{\sum_{i=1}^N W_q(i, x, y, z, t)} \quad (8)$$

where q is the quantity being nudged, p is the dry hydrostatic pressure, F_q is the physical trend terms of q , G_q is the nudging strength for q , N is the total number of observations, i is the index of the current observation, W_q is the spatiotemporal weighting function based on the temporal and spatial interval between the observation and the current model location, q_o is the observed value of q , and $q_m(x_i, y_i, z_i, t)$ is the model value of q interpolated to the observation location. The quantity $q_o - q_m$ is the innovation; as the model value (q_m) develops, the innovation associated with a given observation will develop over time (before and after the observation time). Thus, as the model value approaches the observed value, the nudging tendency term decreases.

Most of these experiments used identical configuration options in the WRF model (Table 3). A total of thirty-three vertical layers were set up in the WRF model with finer layers near ground. Two-thirds layers (twenty-two layers) were under 800 hPa, and eight layers were between 800 and 950 hPa. The physical process parameterization schemes included the Yonsei University (YSU) PBL scheme, the Grell-Freitas cumulus parameterization scheme, the Lin microphysics scheme, the RRTMG longwave and shortwave radiation scheme, the revised Monin-Obukhov surface-layer scheme and the Noah land-surface scheme. The WRF-Chem model was used to further simulate the aerosols' radiative interaction with the CBMZ gas-phase chemistry scheme and the MOSAIC aerosol chemistry scheme. The spin-up time of the WRF-Chem model was 72 h. The MIX emission inventory developed by Li et al. (2017) was adopted to represent sources of anthropogenic emissions, including contributions from agriculture, industry, power, residential, and transportation sources. All times referred to here are Beijing local time.

Table 4 shows the statistical results of the comparison for the DSH station in Shanghai and the SORPES station in Nanjing. In all four experiments, the correlation coefficient between WRF simulated and observed values was above 0.8 for air temperature and above 0.52 for wind speed. Error values are lowest for EXP-FNL*. We conclude that the meteorological fields from EXP-FNL* are superior to those from EXP-FNL, EXP-ERA5, and EXP-CHEM. EXP-CHEM presented improvements for both RMSE and MAE in Nanjing and Shanghai, which indicated the reduction in error as a result of including the aerosol radiative forcing in the model.

Fig. 2 presents hourly time series of observed and simulated PM_{2.5} concentrations from 25 November to 12 December 2017 in four cities in eastern China (Shanghai, Nanjing, Ningbo, and Xiamen). The WRF-Chem simulations capture the spatiotemporal variations of PM_{2.5}

Table 2
Summary of LPDM configurations and key parameters.

Items	Configurations/Parameters
Model	WRF-FLEXPART v3.3
Number of particles	3000 particles released at 10 m above ground level
Backward period	7 days
Baseline	Lowest concentrations in winter
Footprint level	100 m above ground level
Emission data	PKU global emission inventories (0.1° × 0.1°)
Simulation domain	21° in longitude by 18° in latitude

Table 4

A comparison of observed and simulated 2-m air temperature and 10-m wind speed in Shanghai and Nanjing.

Category		EXP-ERA5	EXP-FNL	EXP-FNL*	EXP-CHEM
T _{SH} (°C)	MAE	0.82	0.77	0.94	0.72
	RMSE	2.16	2.67	1.88	2.31
	r	0.83	0.85	0.94	0.89
WS _{SH} (m/s)	MAE	0.94	0.92	0.83	0.91
	RMSE	2.64	2.71	1.26	1.61
	r	0.71	0.71	0.82	0.65
T _{NJ} (°C)	MAE	0.49	0.46	0.39	0.42
	RMSE	2.44	2.32	1.90	2.19
	r	0.84	0.87	0.90	0.88
WS _{NJ} (m/s)	MAE	0.78	0.80	0.75	0.77
	RMSE	2.10	2.15	1.28	2.06
	r	0.52	0.52	0.75	0.60

concentrations well. The concentrations of PM_{2.5} in Shanghai during this period ranged from 12.4 to 231.6 µg/m³, and those of Nanjing were from 12.1 to 228.4 µg/m³. Ningbo and Xiamen are located in the south of Shanghai and Nanjing, and away from the North China Plain, a highly polluted area in China. The concentrations of PM_{2.5} in these two cities were significantly lower than those of Shanghai and Nanjing, which only ranged from 7.3 to 117.3 µg/m³ (Ningbo) and 8.2–66.0 µg/m³ (Xiamen). The WRF-Chem simulations reasonably reproduced such kind of difference among these cities.

3. Results and discussions

In this section we investigate the impact of the driving meteorological fields on WRF-FLEXPART model simulations. This is achieved through a comparative analysis of the effect on the statistical metrics due to (i) different reanalysis data (EXP-FNL vs. EXP-ERA5); (ii) including observational meteorological nudging (EXP-FNL vs. EXP-FNL*); and (iii) considering model representation of aerosols' radiative interaction on meteorology (EXP-FNL vs. EXP-CHEM). The statistical analysis results for each experiment in Shanghai and Nanjing are shown in Fig. 3. The numbers in Fig. 3 characterize the simulated sites, and the colors represent different sensitivity experiments. The deviation in WRF simulations initialized with different reanalysis data was negligible in Shanghai. Compared with EXP-FNL, the experiment with meteorological data assimilation (EXP-FNL*) and the experiment including the aerosol-PBL radiative feedback (EXP-CHEM) show statistically significant increases in the correlation coefficients (*r*) and RMSE reductions at both stations. The EXP-FNL* experiment was more skillful than the other experiments for both stations. In Nanjing, an inland city, *r* increased from 0.31 in EXP-FNL to 0.62 in EXP-FNL*. In the following

discussions, we attempt to uncover the drivers behind the differences between these WRF-FLEXPART simulation experiments.

3.1. Comparative analysis of simulation results driven by different reanalysis data

The CO concentrations at the two stations obtained through the inverse calculation of SIM in EXP-ERA5 and EXP-FNL were compared with observational data (OBS) to evaluate the influence on the simulation accuracy of using different reanalysis data. The RMSEs of EXP-ERA5 and EXP-FNL are relatively similar; the average RMSEs for the two stations are 0.58 and 0.60 mg/m³ (Fig. 3). A scatter plot of simulated and observed CO concentrations in Nanjing is shown in Fig. 4a. The slopes of linear fits to the data were 0.14 for EXP-FNL and 0.17 for EXP-ERA5, which indicate the tracer, CO, was greatly underestimated in both experiments. We further subdivided the data into polluted and normal periods, based on whether the CO concentration was greater or lower than 1.5 mg/m³. Fig. 4b shows the accuracy rates (AR) of EXP-FNL and

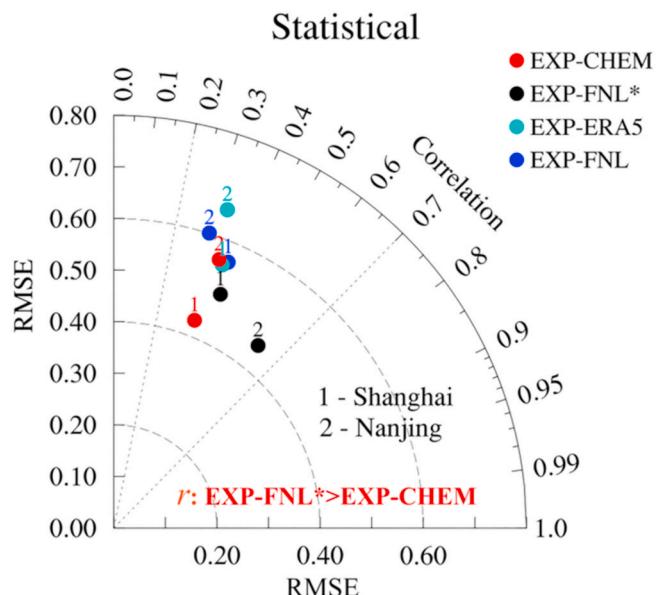


Fig. 3. Correlation coefficient (*r*) and RMSE between simulated and observed values of CO concentrations at two different stations over the full experimental period in the four experiments. The radius indicates RMSE; the azimuthal angle is the Pearson correlation coefficient.

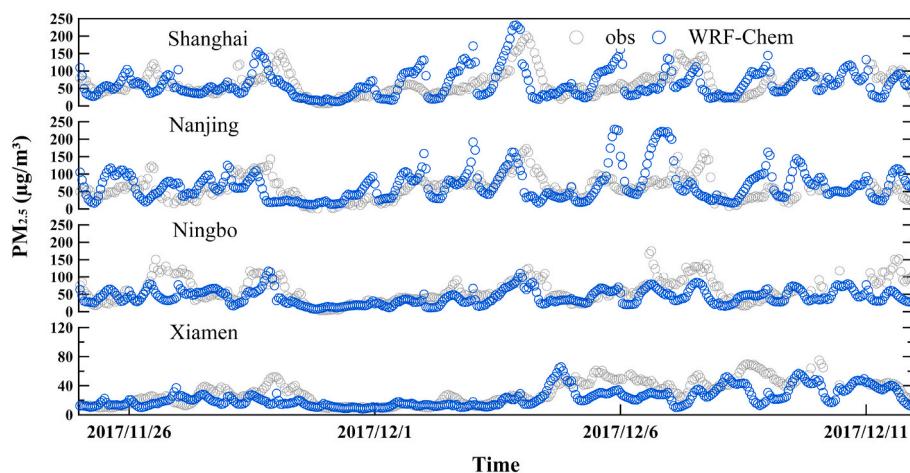


Fig. 2. Comparison of PM_{2.5} time series between WRF-Chem simulations and observations in Shanghai, Nanjing, Ningbo and Xiamen.

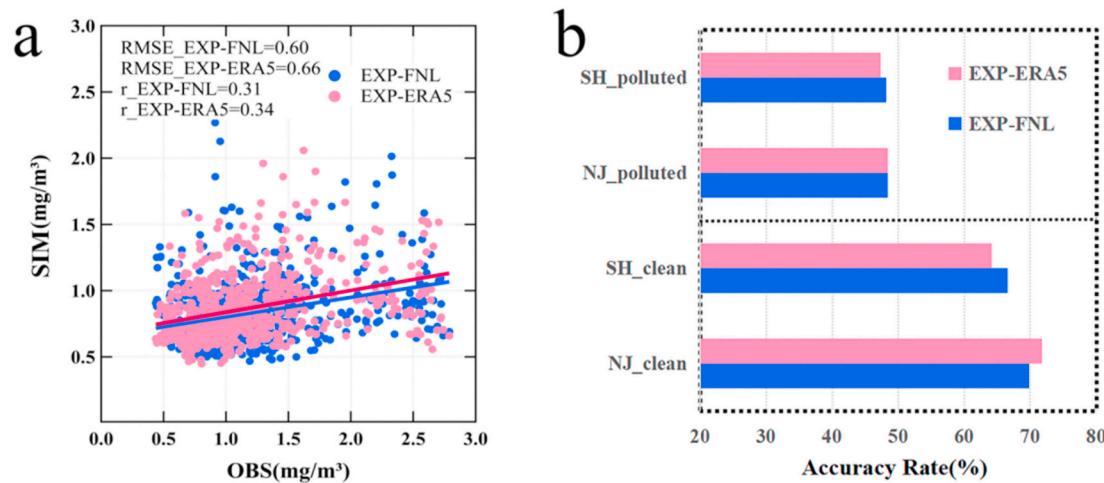


Fig. 4. (a) Scatterplot of observed (OBS) and simulated CO concentrations (SIM) at the Nanjing site and their liner fitting; (b) Accuracy Rates for the Nanjing (NJ) and Shanghai (SH) sites in normal and the polluted periods for EXP-FNL and EXP-ERA5.

EXP-ERA during normal and polluted periods. In both cases, the ARs for EXP-FNL and EXP-ERA5 are similar. However, there is a significant decrease in ARs during polluted times compared with normal periods; the average ARs for the two experiments, based on the two sites, decreases from 68.0% to 48.1%.

Fig. 5a–b show the average vertical profiles of horizontal wind speed (WS) based on sounding data (black circle) and WRF models simulations under different pressure (P) (EXP-ERA5, red line; EXP-FNL, blue line). The radiosonde stations are located at Shanghai Hongqiao Airport and Nanjing Lukou Airport. The coordinates of these stations were used to extract the meteorological fields from the WRF model simulations. Above 950 hPa, the WRF-simulated wind speed fields driven by the two reanalysis datasets are both consistent with the sounding observations. However, below 950 hPa, the horizontal wind speeds are significantly higher than the sounding observations. On the one hand, the higher simulated horizontal wind speed near the ground surface leads to a decrease in the residence time of particles released in the WRF-FLEXPART simulation. It follows that under conditions of a constant

emission source, the SIM is underestimated. On the other hand, the stagnant air condition associated with low wind speed is not conducive to the transport and dispersion of pollutants, so the difference between the SIM and observations is greater during the polluted periods.

3.2. Influence of observational meteorological nudging

From December 2017 to the beginning of January 2018, several heavy air pollution episodes occurred in the YRD region. In order to improve the traceability of stagnant air conditions in the WRF-FLEXPART model, NCEP ADP Global Upper Air Observational Weather Data was nudged into the meteorological driving field. The detailed descriptions of nudging data are in section 2.2. Compared to EXP-FNL, the results of EXP-FNL* present a significant improvement in the polluted periods. Fig. 6 presents scatterplots of observed and simulated CO concentrations (SIM) in EXP-FNL* and EXP-FNL at the Nanjing station during the normal episodes ($\text{CO} \leq 1.5 \text{ mg/m}^3$) and high polluted episodes ($\text{CO} > 1.5 \text{ mg/m}^3$). EXP-FNL* shows significant improvement

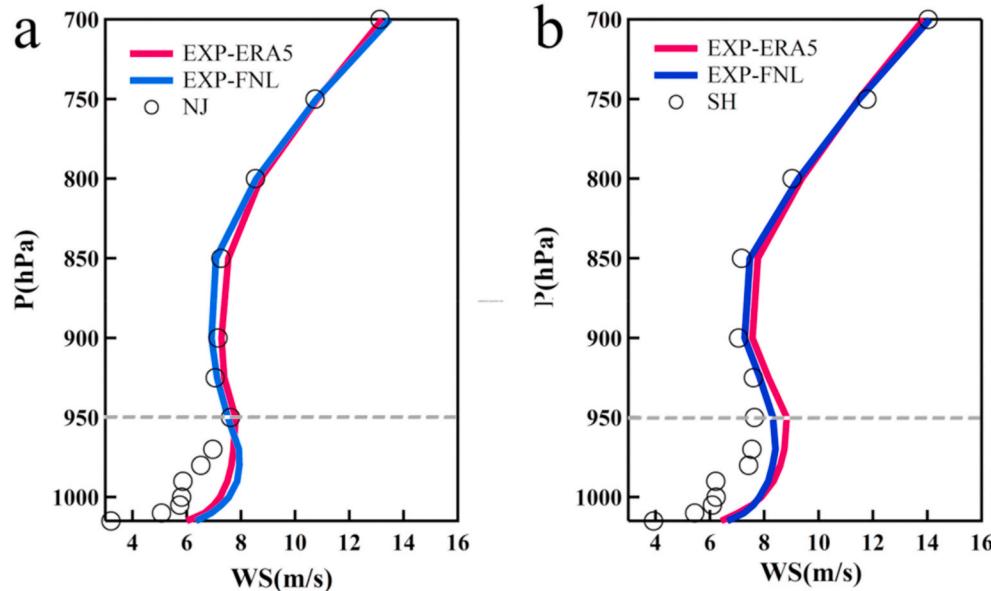


Fig. 5. Average vertical profiles of horizontal wind speed (WS) at the (a) Nanjing (NJ) and (b) Shanghai (SH) stations based on observations (black circles) and results from the EXP-ERA5 (red lines) and EXP-FNL (blue lines) experiments. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

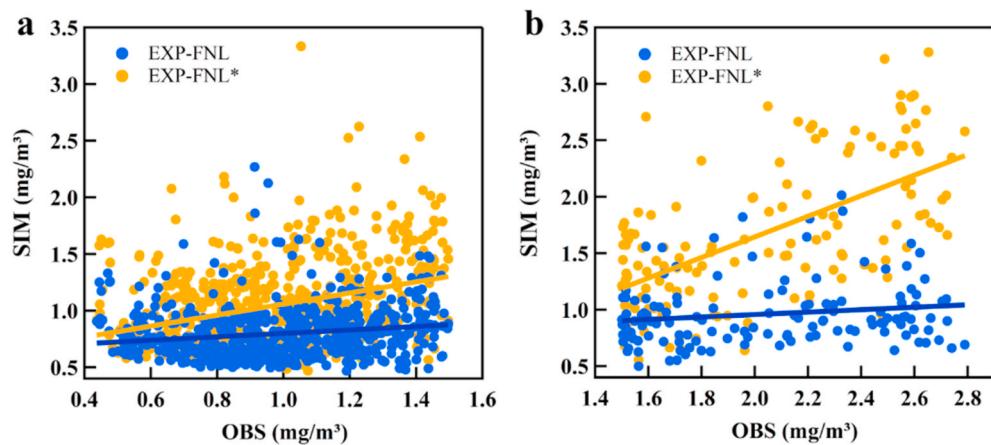


Fig. 6. Scatterplot of simulated (SIM) and observed (OBS) CO concentrations at the Nanjing site during (a) normal periods ($\text{CO} \leq 1.5 \text{ mg/m}^3$) and (b) heavy polluted periods ($\text{CO} > 1.5 \text{ mg/m}^3$) and the resulting lines of best fit for EXP-FNL* (orange markers) and EXP-FNL (blue markers) simulations. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

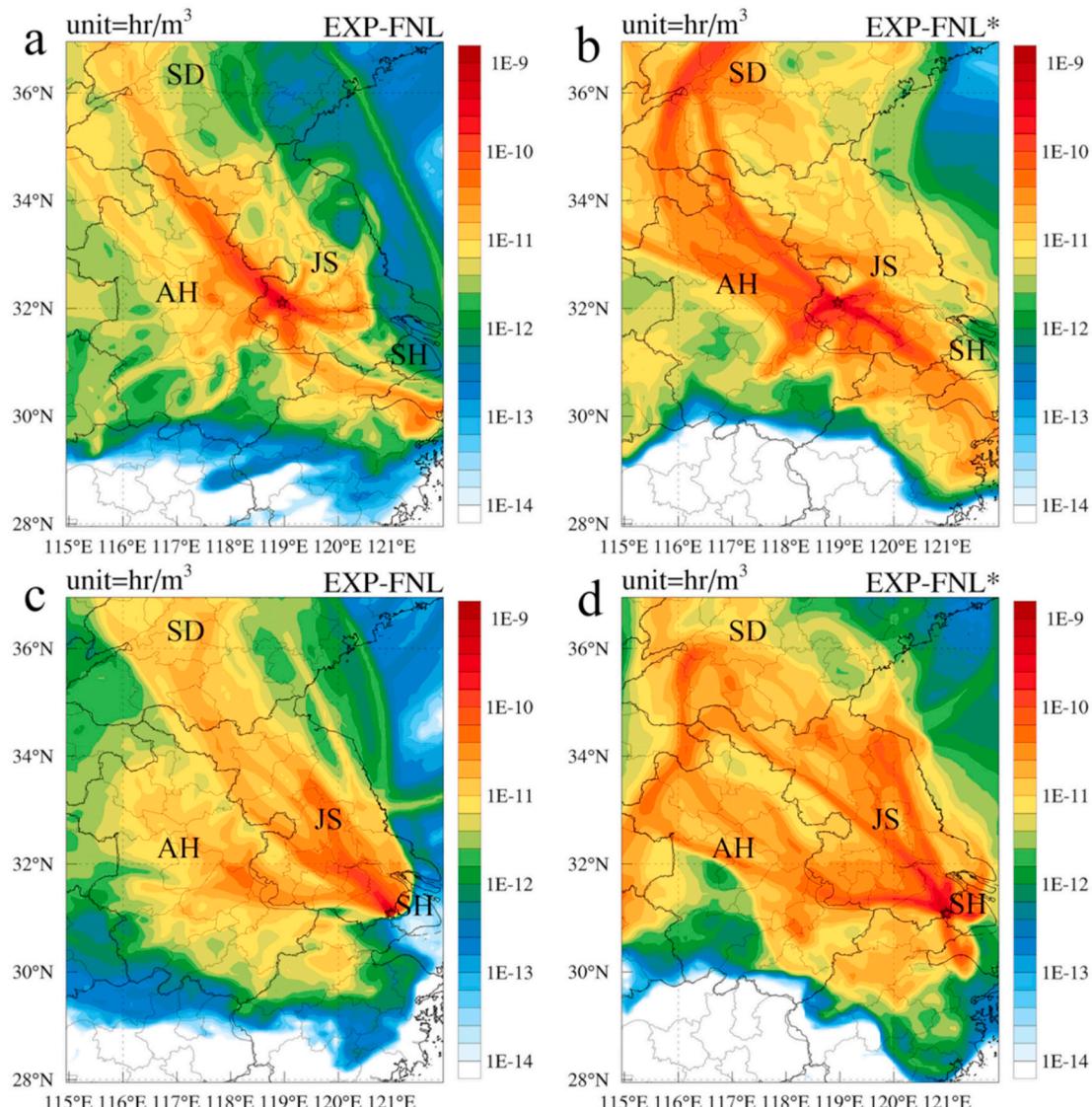


Fig. 7. Average Footprint of the polluted periods: (a) EXP-FNL in Nanjing; (b) EXP-FNL* in Nanjing; (c) EXP-FNL in Shanghai; (d) EXP-FNL* in Shanghai. SD, AH, JS, SH are Shandong Province, Anhui Province, Jiangsu Province, and the city of Shanghai, respectively.

over EXP-FNL. Based on linear fitting, the slope of the fit to the EXP-FNL simulation results is only 0.15 (normal) and 0.11 (polluted), while the slopes have been improved to 0.48 (normal) and even 0.91 (polluted) for EXP-FNL*. The uncertainties of WRF-FLEXPART triggered by the meteorological driving field can be reduced through the assimilation of meteorological data.

The local transport in this paper referred to the air mass transmission originating from local cities and neighboring areas. In contrast, the regional transport referred to the air masses coming from non-adjacent areas. By comparing the average footprint of the inland city, Nanjing, during the heavy polluted period in the EXP-FNL and EXP-FNL* experiments (Fig. 7a and b), the difference in transport can be understood visually. Averaged over all high polluted periods, the EXP-FNL simulations (Fig. 7a) reveal a weak airflow from the northwest to the region, but significant local transport. In contrast, the results from EXP-FNL* (Fig. 7b) not only depicted the transport of air masses from Shandong Province to Jiangsu Province, but also simulate the transport channel in central Anhui Province. Spatial distributions of the average footprint during high polluted events in the coastal city of Shanghai are shown in Fig. 7c (EXP-FNL) and Fig. 7d (EXP-FNL*). The multiple paths of air mass transport in the southwestern part of Shandong Province and central Anhui Province are the main channels affecting the air pollution in the YRD region of eastern China during these periods. Airflow from Jiangsu Province in eastern China is the main source of air masses affecting Shanghai. Similar to the result for Nanjing, EXP-FNL* (Fig. 7d) is able to characterize the impact of regional transport on local conditions more accurately. Compared with EXP-FNL*, the areas in EXP-FNL where the average footprint is underestimated are mainly located in western Shandong, central Anhui, and Jiangsu. The contribution of local atmospheric transport has a greater role in EXP-FNL, whereas EXP-FNL* can more accurately simulate the impact of both local and regional transport at the release point.

The horizontal wind field is the key meteorological element in the propagation of uncertainty from the meteorological fields to LPDMs simulation results (Lin and Gerbig, 2005). Fig. 8 presents the average vertical profiles of wind speed for the normal and the polluted periods at the Shanghai site based on observations and the outputs from the EXP-FNL* and EXP-FNL experiments. The vertical profile of horizontal wind speed in EXP-FNL* is basically consistent with the observed sounding data, and is closer to the observed profile than EXP-FNL during

both normal (Fig. 8a) and polluted periods below 950 hPa (Fig. 8b). With assimilation of observational data, the WRF model is able to restore the local stagnant meteorological conditions during the polluted periods, which is an important guarantee for the accuracy of particle tracing in FLEXPART. At the same time, although the simulation results of the WRF model nudged by observational meteorological data are similar to the real atmosphere, inevitable errors due to model errors still existed, such as the integral instability and underlying surface properties.

The average 950 hPa horizontal wind fields in EXP-FNL* during polluted periods and normal periods are shown in Fig. 9a-d. During the polluted periods, the wind direction (indicated by the black arrows) in the eastern part of China is mainly northwesterly, and the wind speed at 950 hPa is significantly reduced in Jiangsu and Anhui Province. The northwesterly airflow carries a large amount of pollutants to the YRD region. Local accumulation of these pollutants is consistent with the surface CO observations during the same period (Fig. 9c). Fig. 9b shows the spatial distribution of the difference in the horizontal 950 hPa wind speed field between EXP-FNL* and EXP-FNL during the polluted periods. Regions where the wind speed is significantly lower in EXP-FNL* compared with EXP-FNL are mainly concentrated north of Nanjing and Shanghai to the west of Shandong Province, and in most of the northern and central parts of Anhui Province. During the normal periods, the relatively clean eastern air mass at sea is transported to Nanjing and Shanghai (Fig. 9d), accompanied by lower CO concentration than the north region (Fig. 9f). The wind speed field of EXP-FNL in the horizontal 950 hPa is slightly higher than EXP-FNL* in eastern China during the normal periods (Fig. 9e). These results are consistent with the difference in wind profiles between EXP-FNL and EXP-FNL* shown in Fig. 8. By assimilating meteorological data from surface and radiosonde observation, the WRF model is able to successfully reveal the wind field, especially those near the ground surface, which is then substantially improve the particle tracing in the LPDM.

3.3. Contribution from aerosols' radiative feedback

In consideration of the aerosol's effect on the evolution of the PBL, the feedback from the aerosols to the radiation schemes was turned on in EXP-CHEM. Fig. 10 shows the probability distribution of biases in simulated CO concentrations at the Nanjing site from EXP-FNL, EXP-

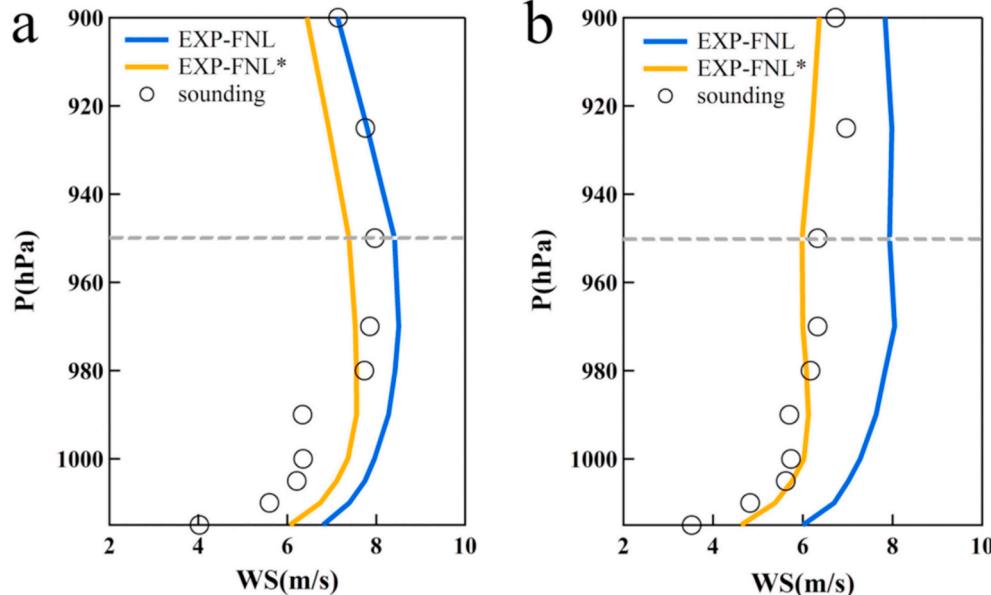


Fig. 8. Average vertical profiles of horizontal wind speed (WS) based on sounding observations in Shanghai (black circles) and results from EXP-FNL* and EXP-FNL during (a) normal and (b) polluted periods.

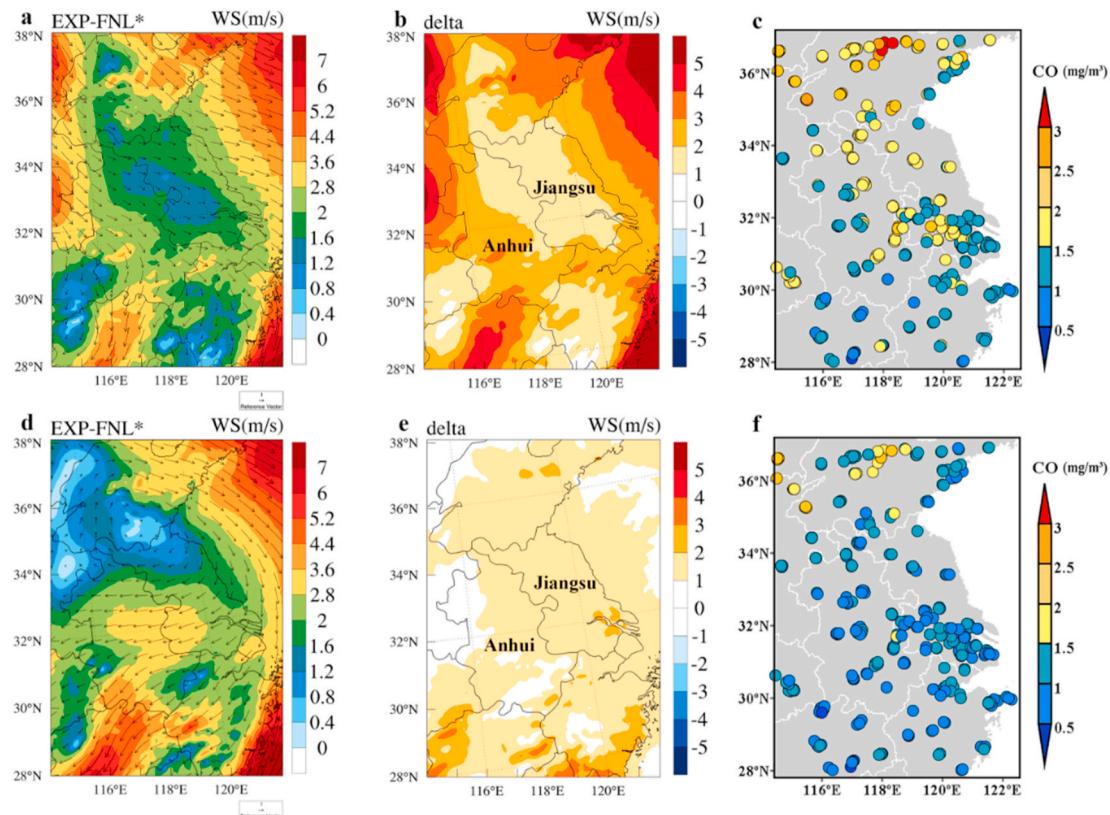


Fig. 9. Average 950 hPa horizontal wind field in (a,d) EXP-FNL* and (b,e) EXP-FNL minus EXP-FNL*: (a,b) during polluted periods; (c,d) during normal periods. In (a) and (d), the black vectors indicate the wind direction. Observed surface CO concentrations during (c) polluted periods and (f) normal periods.

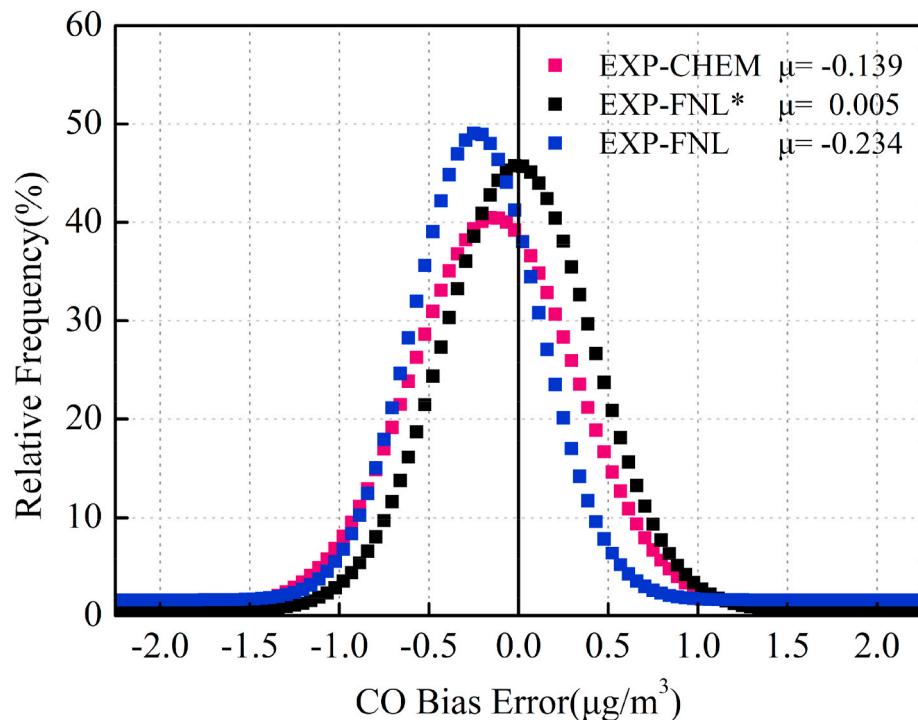


Fig. 10. Probability distribution and mean value (μ) of the bias errors in simulated CO concentrations in EXP-CHEM (pink markers), EXP-FNL (blue markers) and EXP-FNL* (black markers). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

FNL*, and EXP-CHEM. The mean value (μ) of the distribution is a measure of the accuracy of the simulated results. The closer μ is to zero, the more accurate the simulation results are. The value of μ is -0.23 for EXP-FNL, which indicates model underestimation of CO concentrations. The value of μ is just 0.005 for EXP-FNL*, indicative of good accuracy emerging in this experiment. For EXP-CHEM, μ is -0.14 , which is between the values for EXP-FNL and EXP-FNL*. On the one hand, consideration of aerosols' radiative forcing is able to promote the simulation performance of the FLEXPART model, while on the other hand, these results indicate the diversity of factors affecting the simulation of atmospheric transport in the FLEXPART model.

Accuracy rates based on EXP-FNL, EXP-CHEM, and EXP-FNL* simulation results for the Nanjing and Shanghai sites during polluted periods are shown in Fig. 11. Compared with EXP-FNL, the accuracy rates of EXP-FNL* increase by 25.6% and 19.1% for the Nanjing and Shanghai sites, respectively, so that the total accuracy rates are 74.9% and 67.3% . Compared with EXP-FNL, the accuracy rate for the EXP-CHEM experiments with aerosols' radiative effect increases by 11.6% (Nanjing) and 9.8% (Shanghai). The accuracy increase in EXP-CHEM was 48% of the difference in accuracy between EXP-FNL and EXP-FNL*, which implying that almost half part (48%) of the improvement produced by assimilating meteorological data is due to the comprehensive consideration of radiative forcing. The biases between EXP-CHEM and EXP-FNL* may be attributable to many reasons, such as the unavailability of immediate emission inventory and surface characteristics like soil moisture and terrain (Koster et al., 2010; Kumar et al., 2006; LeMone et al., 2008; Brioude et al., 2012). Although EXP-CHEM cannot fully explain the improvement of EXP-FNL*, the gap between the LPDM modelling and observations is notably narrowed particularly when the pollution is severe.

Besides horizontal winds, vertical mixing is another key factor for the propagation of uncertainty from meteorological fields to LPDMs output (Gerbig et al., 2008), especially in stagnant weather conditions. PBL height is a rational indicator for vertical mixing, as the plume is mixed homogeneously throughout the PBL. Fig. 12a compares the hourly PBL height (PBLH) between EXP-FNL* and observations during daytime (10:00–16:00). The vertical profile of aerosol extinction coefficient at 532 nm was observed by a Lidar (Model AGHJ-I-LIDAR) to estimate the top of the PBL, during 1–26 December 2017 at the East China University of Science and Technology, Fengxian Campus (30.8°N , 121.5°W). The height of the PBL is extracted using the gradient method, according to the rate of the backscattered signal decays with height (Comerón et al., 2013). It characterizes the boundary layer's height at the minimum

value of the first or the second derivative of the backscattered signal. The gradient method is convenient but is easily disturbed by environmental noise. As EXP-FNL* includes the assimilation of observations, it is able to successfully simulate PBLH and its diurnal variation. Therefore, the PBLH estimates from the EXP-FNL* experiments are selected as reference values so that a comparison can be made between PBLH values for eastern China in EXP-FNL and EXP-CHEM. Fig. 12b and c show the spatial distributions of the difference in PBLH based on daytime values during polluted periods. EXP-FNL overestimates the PBLH in eastern China by $40\text{--}120\text{ m}$ (Fig. 12b), while the difference in PBLH between EXP-CHEM and EXP-FNL* is generally smaller than 40 m in Jiangsu, Shandong, Zhejiang, and Anhui Province (Fig. 12c). We conclude that by considering the aerosols' radiative forcing, which is suppressing the development of PBL, the WRF-FLEXPART model can depict daytime atmospheric transport and diffusion accurately. In contrast, the overestimation of PBL development in EXP-FNL amplifies the vertical diffusion, which prevents air pollutants from accumulating into a lower height.

4. Conclusions

LPDMs have been extensively employed to investigate the source attribution and formation processes of haze pollution. To increase the model performance in polluted regions like eastern China, we improve the simulations of WRF-FLEXPART by (i) using different reanalysis data, (ii) including observational meteorological nudging, and (iii) considering model representation of aerosols' radiative interaction to meteorology. The potential source contribution method was applied to evaluate the model performance with respect to regional air pollution transport and boundary layer diffusion. We attempted to identify the distributions of parameters that influence winter haze pollutant with regards to characterizing regional air pollution transport, and a more feasible meteorological input and model configurations for LPDM simulations in eastern China is proposed.

The main conclusions from this study are: (1) Simulations driven by two commonly used reanalysis datasets, FNL and ERA5, generally underestimate pollutant concentration in the LPDM, with similar errors, especially during periods with heavy pollution. The decreases of the residence time of particles in the LPDM are directly caused by overestimation of PBLH and PBL horizontal wind speeds. Stagnant air conditions featuring low wind speeds are not conducive to the transport and dispersion of near-surface pollutants, so significant underestimation exists during the polluted periods. (2) By assimilating meteorological

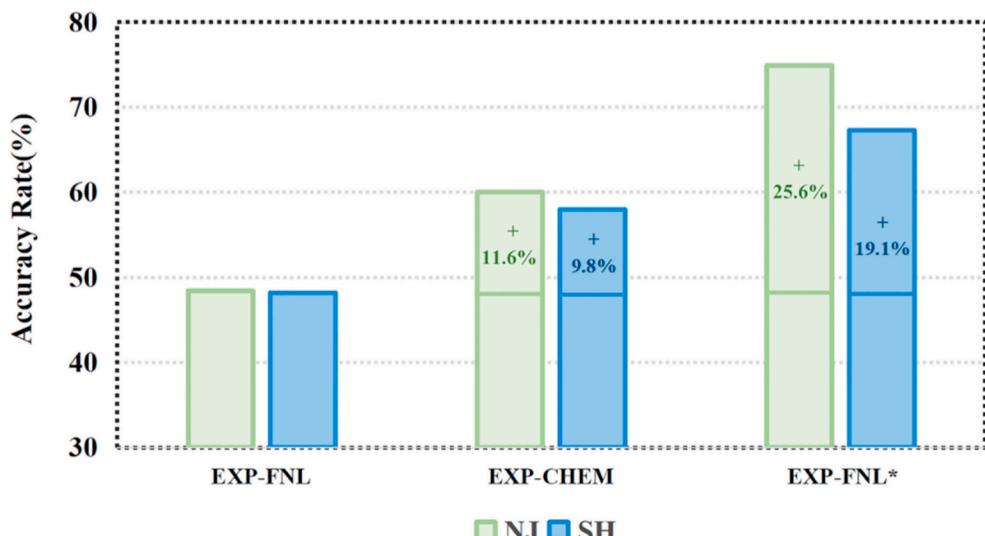


Fig. 11. Accuracy rate (%) of the EXP-FNL, EXP-CHEM, and EXP-FNL* experiments for polluted periods at the Nanjing (NJ) and Shanghai (SH) sites.

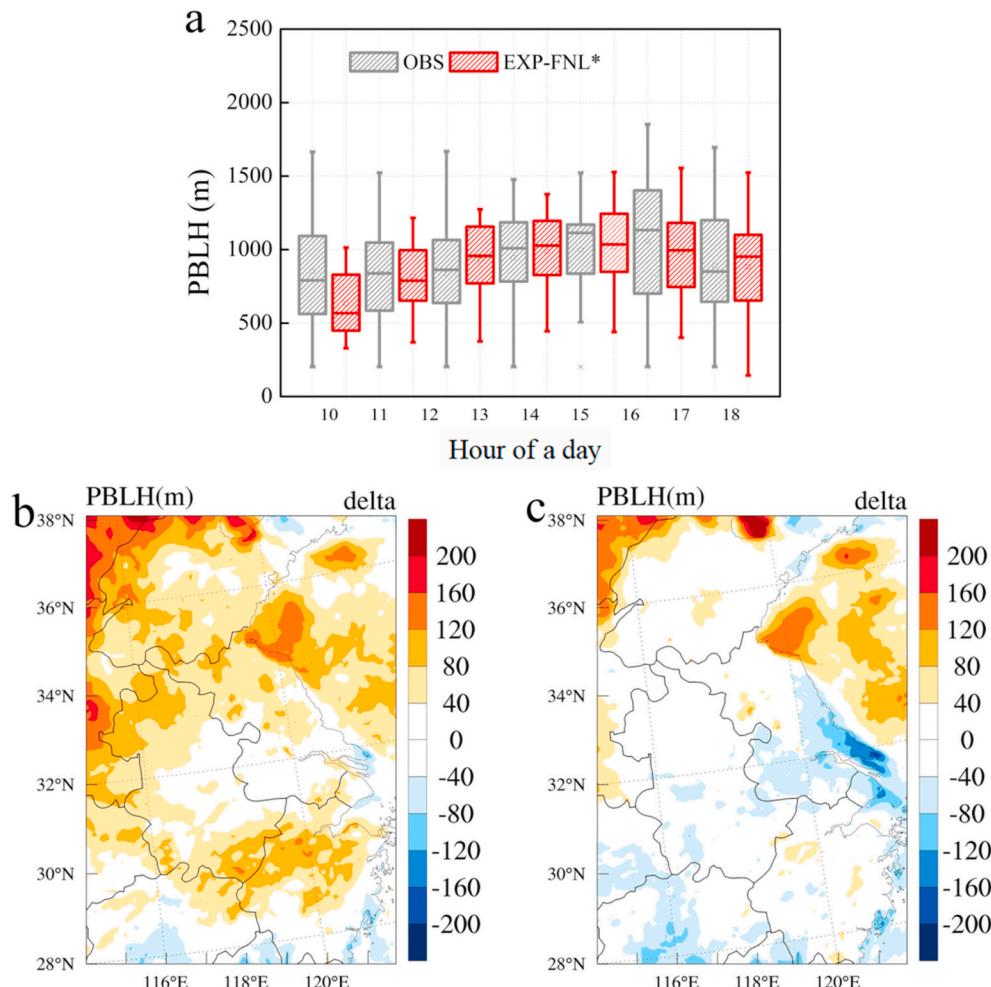


Fig. 12. (a) Diurnal variation of planetary boundary layer height (PBLH) (m) in EXP-FNL* and in observations. (b,c) Spatial distribution of the difference in PBLH (m) during polluted daytime periods using EXP-FNL* values as reference; (b) EXP-FNL minus EXP-FNL* and (c) EXP-CHEM minus EXP-FNL*.

data from surface weather stations and radiosonde observations (EXP-FNL*), the WRF model is able to well represent the wind field, especially near the ground surface, which then substantially improves particle tracing in the LPDM. The accurate simulation of the air velocity during the local stagnant meteorological conditions is the key to improving the traceability in the model. The data assimilation method is convenient and practical towards improving simulation accuracy. However, one limitation of this method is the availability of meteorological observations in applications. (3) By including aerosols' radiative interaction, which influences PBL development significantly, the gap between the LPDM modelling and observations is notably narrowed, particularly when air pollution gets severe. Quantitatively, the accuracy increase of the simulations with aerosols' radiative effect accounted for 48% of the improvement produced by assimilating meteorological data. Near-surface meteorological fields are strongly affected by the aerosols' radiative interaction, and this factor should be investigated further in the future.

Overall, the meteorological input is of great importance in LPDMs, especially during stagnant weather conditions. In regions with intensive pollution like China and India, applying observational data assimilation to meteorology can serve as an effective way to reduce biases in LPDMs and better understand the source regions as well as the accumulation of pollution. For air quality forecasting or regions with few observational data, considering aerosols' radiative interaction on meteorology would better simulate air pollution formation and its source contribution.

Data availability

Meteorological datasets used in this work can be acquired from <https://rda.ucar.edu/datasets/ds633.0/> (the fifth generation of ECMWF atmospheric reanalyses of the global climate), <https://rda.ucar.edu/datasets/ds083.2/> (NCEP FNL (Final) Operational Global Analysis data), and <https://rda.ucar.edu/datasets/ds351.0/#/> (NCEP ADP Global Upper Air Observational Weather Data). Emission inventories used in this work are available from <http://inventory.pku.edu.cn/> and <http://www.micmodel.org/dataset-meic.html>. Observational air pollutant datasets in the YRD region are public data from the Ministry of Environmental Protection (MEP) of People's Republic of China.

CRediT authorship contribution statement

Mengwei Jia: Data curation, modelling, Writing - original draft preparation. **Xin Huang:** Supervision, Writing - review & editing. **Ke Ding:** Methodology, Software, Validation. **Qiang Liu:** Software, Visualization. **Derong Zhou:** Methodology, Writing - review & editing. **Aijun Ding:** Conceptualization, Supervision, Writing - review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was funded by the Ministry of Science and Technology of the People's Republic of China (2016YFC0200500; 2018YFC0213800; 2016YFC0202000), and the National Natural Science Foundation of China (91544231, 41725020, 91744311, and 41422504). The numerical modelling was conducted on computing facilities at the High-Performance Computing Centering (HPCC) at Nanjing University.

References

- An, X.Q., Yao, B., Li, Y., Li, N., Zhou, L.X., 2014. Tracking source area of Shangdianzi station using Lagrangian particle dispersion model of FLEXPART. *Meteorol. Appl.* 21 (3), 466–473.
- Angevine, W.M., Brioude, J., McKeen, S., Holloway, J.S., 2014. Uncertainty in Lagrangian pollutant transport simulations due to meteorological uncertainty from a mesoscale WRF ensemble. *Geosci. Model Dev. (GMD)* 7 (6), 4603–4643.
- Brioude, J., Angevine, W., McKeen, S., Hsie, E.Y., 2012. Numerical uncertainty at mesoscale in a Lagrangian model in complex terrain. *Geosci. Model Dev. (GMD)* 5, 1127–1136.
- Brioude, J., Kim, S.W., Angevine, W.M., Frost, G.J., Lee, S.H., McKeen, S.A., Trainer, M., Fehsenfeld, F.C., Holloway, J.S., Ryerson, T.B., Williams, E.J., Petron, G., Fast, J.D., 2011. Top-down estimate of anthropogenic emission inventories and their interannual variability in Houston using a mesoscale inverse modeling technique. *J. Geophys. Res.* 116 (D20), D20305.
- Byers, H.R., 1974. General Meteorology, fourth ed. McGrawHill, New York, U.S.A.
- Comerón, A., Sicard, M., Rocadenbosch, F., 2013. Wavelet correlation transform method and gradient method to determine aerosol layering from lidar returns: some comments. *J. Atmos. Ocean. Technol.* 30 (6), 1189–1193.
- Damoah, R., Spichtinger, N., Forster, C., James, P., Mattis, I., Wandinger, U., Beirle, S., Wagner, T., Stohl, A., 2004. Around the world in 17 days - hemispheric-scale transport of forest fire smoke from Russia in May 2003. *Atmos. Chem. Phys.* 4 (5), 1449–1471.
- Daria, B., Carsten, A.S., Małgorzata, W., 2017. Source regions of ragweed pollen arriving in south-western Poland and the influence of meteorological data on the HYSPLIT model results. *Aerobiologia* 33 (3), 315–326.
- Decarlo, P.F., Ulbrich, I.M., Crounse, J., de Foy, B., Dunlea, E.J., Aiken, A.C., Knapp, D., Weinheimer, A.J., Campos, T., Wennberg, P.O., Jimenez, J.L., 2010. Investigation of the sources and processing of organic aerosol over the Central Mexican Plateau from aircraft measurements during MILAGRO. *Atmos. Chem. Phys.* 10 (12), 525–536.
- Deng, A., Stauffer, D., Gaudet, B., Dudhia, J., Bruyere, C., Wu, W., Vandenberghe, F., Liu, Y., Bourgeois, A., 2009. Update on WRF-ARW end-to-end multi-scale FDDA system. In: 10th WRF Users' Workshop, NCAR, 23–26 June. Boulder, CO. 1.9. 14 pp. Available: <http://www2.mmm.ucar.edu/wrf/users/workshops/WS2009/WorkshopPapers.php>.
- Derwent, R.G., 1990. Evaluation of a number of chemical mechanisms for their application in models describing the formation of photochemical ozone in Europe. *Atmos. Environ.* 24A, 2615–2624.
- De Leeuw, F.A.A.M., Leysius, V.R.H.J., Builjtjes, P.J.H., 1990. Calculation of long-term averaged ground level ozone concentrations. *Atmos. Environ.* 24A, 185–193.
- Ding, A.J., Fu, C.B., Yang, X.Q., Sun, J.N., Petäjä, T., Kermanen, V.M., Wang, T., Xie, Y., Herrmann, E., Zheng, L.F., Nie, W., Liu, Q., Wei, X.L., Kulmala, M., 2013a. Intense atmospheric pollution modifies weather: a case of mixed biomass burning with fossil fuel combustion pollution in eastern China. *Atmos. Chem. Phys.* 13 (20), 10,545–10,554.
- Ding, A.J., Huang, X., Nie, W., Sun, J.N., Kermanen, V.M., Petäjä, T., Su, H., Cheng, Y.F., Yang, X.Q., Wang, M.H., Chi, X.G., Wang, J.P., Virkkula, A., Guo, W.D., Yuan, J., Wang, S.Y., Zhang, R.J., Wu, Y.F., Song, Y., Zhu, T., Zilitinkevich, S., Kulmala, M., Fu, C.B., 2016a. Enhanced haze pollution by black carbon in megacities in China. *Geophys. Res. Lett.* 43.
- Ding, A.J., Huang, X., Nie, W., Chi, X., Xu, Z., Zheng, L., Xu, Z., Xie, Y., Qi, X., Shen, Y., Sun, P., Wang, J., Wang, L., Sun, J., Yang, X.-Q., Qin, W., Zhang, X., Cheng, W., Liu, W., Pan, L., Fu, C., 2019. Significant reduction of PM_{2.5} in eastern China due to regional-scale emission control: evidence from SORPES in 2011–2018. *Atmos. Chem. Phys.* 19 (18), 11791–11801.
- Ding, A.J., Nie, W., Huang, X., Chi, X., Sun, J.N., Kermanen, V., Xu, Z., Guo, W.D., Petäjä, T., Yang, X.Q., Kulmala, M., Fu, C.B., 2016b. Long-term observation of air pollution-weather/climate interactions at the SORPES station: a review and outlook. *Front. Environ. Sci. Eng.* 10 (5).
- Ding, A.J., Wang, T., Fu, C.B., 2013b. Transport characteristics and origins of carbon monoxide and ozone in Hong Kong, South China. *J. Geophys. Res. Atmos.* 118, 9475–9488.
- Ding, A.J., Wang, T., Xue, L.K., Gao, J., Stohl, A., Lei, H.C., Jin, D.Z., Ren, Y., Wang, X.Z., Wei, X.L., Qi, Y.B., Liu, J., Zhang, X.Q., 2009. Transport of north China air pollution by midlatitude cyclones: case study of aircraft measurements in summer 2007. *J. Geophys. Res.* 114, D08304.
- Draxler, Hess, 1998. An overview of the HYSPLIT_4 modelling system for trajectories, dispersion and deposition. *Aust. Meteorol. Mag.* 47, 295–308.
- Dutton, J.A., 1986. The Ceaseless Wind. An Introduction to the Theory of Atmospheric Motion. Dover, New York.
- Eckhardt, S., Stohl, A., Beirle, S., Spichtinger, N., Jennings, S.G., 2003. The North Atlantic Oscillation controls air pollution transport to the Arctic. *Atmos. Chem. Phys.* 3 (5), 3223–3240.
- Fast, J.D., Easter, R.C., 2006. A Lagrangian particle dispersion model compatible with WRF. In: 7th WRF User's Workshop, 6.2. National Center for Atmospheric Research, Boulder, CO.
- Fisher, M., Lary, D.J., 1995. Lagrangian four-dimensional variational data assimilation of chemical species. *Q. J. R. Meteorol. Soc.* 121, 1681–1704.
- Forster, C., Wandinger, U., Wotawa, G., James, P., Mattis, I., Althausen, D., Simmonds, P., O'Doherty, S.J., Jennings, S.G., Kleefeld, C., Schneider, J., Trickl, T., Kreipl, S., Jäger, H., Stohl, A., 2001. Transport of boreal forest fire emissions from Canada to Europe. *J. Geophys. Res. Atmos.* 106 (D19), 22887.
- Gerbic, C., Körner, S., Lin, J.C., 2008. Vertical mixing in atmospheric tracer transport models: error characterization and propagation. *Atmos. Chem. Phys.* 8, 591–602.
- Halse, A.K., Eckhardt, S., Schlabach, M., Stohl, A., Breivik, K., 2013. Forecasting long range atmospheric transport episodes of polychlorinated biphenyls using FLEXPART. *Atmos. Environ.* 71, 335–339.
- Harris, J.M., Draxler, R.R., Oltmans, S.J., 2005. Trajectory model sensitivity to differences in input data and vertical transport method. *J. Geophys. Res. Atmos.* 110, D14109.
- Hegarty, J., Draxler, R.R., Stein, A.F., Brioude, J., Mountain, M., Eluszkiewicz, J., Nehrkorn, T., Ngan, F., Andrews, A., 2013. Evaluation of Lagrangian particle dispersion models with measurements from controlled tracer releases. *J. Appl. Meteorol. Clim.* 52, 2623–2637.
- Hernández-Ceballos, M.A., Skjøth, C.A., Garcíamozo, H., Bolívar, J.P., Galán, C., 2014. Improvement in the accuracy of back trajectories using WRF to identify pollen sources in southern Iberian Peninsula. *Int. J. Biometeorol.* 58 (10), 2031–2043.
- Hoffmann, L., Günther, G., Li, D., Stein, O., Wu, X., Griesbach, S., Heng, Y., Konopka, P., Müller, R., Vogel, B., Wright, J.S., 2019. From ERA-Interim to ERA5: the considerable impact of ECMWF's next-generation reanalysis on Lagrangian transport simulations. *Atmos. Chem. Phys.* 19, 3097–3124.
- Huang, X., Ding, A.J., Liu, L.X., Liu, Q., Ding, K., Niu, X.R., Nie, W., Xu, Z., Chi, X.G., Wang, M.H., Sun, J.N., Guo, W.D., Fu, C.B., 2016. Effects of aerosol-radiation interaction on precipitation during biomass-burning season in East China. *Atmos. Chem. Phys.* 16, 10063–10082.
- Huang, X., Wang, Z.L., Ding, A.J., 2018. Impact of aerosol-PBL interaction on haze pollution: multi-year observational evidences in north China. *Geophys. Res. Lett.* 45, 8596–8603.
- Huang, X., Ding, A., Wang, Z.L., Ding, K., Gao, J., Chai, F., Fu, C., 2020. Amplified transboundary transport of haze by aerosol-boundary layer interaction in China. *Nat. Geosci.* 13, 428–434.
- James, P., Stohl, A., Forster, C., Eckhardt, S., Seibert, P., Frank, A., 2003. A 15-year climatology of stratosphere-troposphere exchange with a Lagrangian particle dispersion model 2. Mean climate and seasonal variability. *J. Geophys. Res. Atmos.* 108 (D12), 8522.
- Kljun, N., Rotach, M.W., Schmid, H.P., 2002. A three-dimensional backward Lagrangian footprint model for a wide range of boundary-layer stratifications. *Boundary-Layer Meteorol.* 103 (2), 205–226.
- Koster, R.D., Mahanama, S.P.P., Yamada, T.J., Balsamo, G., Berg, A.A., Boisserie, M., Dirmeyer, P.A., Doblas-Reyes, F.J., Drewitt, G., Gordon, C.T., Guo, Z., Jeong, J.H., Lawrence, D.M., Lee, W.S., Li, Z., Luo, L., Malyshev, S., Merryfield, W.J., Seneviratne, S.I., Stanelle, T., van den Hurk, B.J.M., Vitart, F., Wood, E.F., 2010. Contribution of land surface initialization to subseasonal forecast skill: first results from a multi-model experiment. *Geophys. Res. Lett.* 37, L02402.
- Kumar, S.V., Peters-Lidard, C.D., Tian, Y., Houser, P.R., Geiger, J., Olden, S., Lighty, L., Eastman, J.L., Doty, B., Dirmeyer, P., Adams, J., Mitchell, K., Wood, E.F., Sheffield, J., 2006. Land information system: an interoperable framework for high resolution land surface modeling. *Environ. Model. Software* 21, 1402–1415.
- Lai, A.C.K., Chen, F.Z., 2007. Comparison of a new Eulerian model with a modified Lagrangian approach for particle distribution and deposition indoors. *Atmos. Environ.* 41, 5249–5256.
- LeMone, M.A., Tewari, M., Chen, F., Alfieri, J.G., Niyogi, D., 2008. Evaluation of the Noah land surface model using data from a fair-weather IHOP_2002 day with heterogeneous surface fluxes. *Mon. Weather Rev.* 136, 4915–4941.
- Li, H.W., Wang, D.F., Cui, L., Gao, Y., Huo, J.T., Wang, X.N., Zhang, Z.Z., Tan, Y., Huang, Y., Cao, J.J., Chow, J.C., Lee, S.C., Fu, Q.Y., 2019. Characteristics of atmospheric PM_{2.5} composition during the implementation of stringent pollution control measures in Shanghai for the 2016 G20 summit. *Sci. Total Environ.* 648, 1121–1129. <https://doi.org/10.1016/j.scitotenv.2018.08.219>.
- Li, M., Zhang, Q., Kurokawa, J.I., Woo, J.H., He, K.B., Lu, Z.F., Ohara, T., Song, Y., Streets, D.G., Carmichael, G.R., Cheng, Y.F., Hong, C.P., Huo, H., Jiang, X.J., Kang, S.C., Liu, F., Su, H., Zheng, B., 2017. MIX: a mosaic Asian anthropogenic emission inventory under the international collaboration framework of the MICS-Asia and HTAP. *Atmos. Chem. Phys.* 17, 935–963.
- Li, X.S., Choi, Y., Czader, B., Roy, A., Kim, H., Lefer, B., Pan, S., 2016. The impact of observation nudging on simulated meteorology and ozone concentrations during DISCOVER-AQ 2013 Texas campaign. *Atmos. Chem. Phys.* 16, 3127–3144.
- Lin, J.C., Gerbig, C., 2005. Accounting for the effect of transport errors on tracer inversions. *Geophys. Res. Lett.* 32, L01802.
- Lin, L.X., Huang, X., Ding, A.J., Fu, C.B., 2016. Dust-induced radiative feedbacks in north China: a dust storm episode modeling study using WRF-Chem. *Atmos. Environ.* 129, 43–54.
- Liu, Y., Shao, M., Fu, L.L., Lu, S.H., Zeng, L.M., Tang, D.G., 2008. Source profiles of volatile organic compounds (VOCs) measured in China: part I. *Atmos. Environ.* 42 (25), 6247–6260.
- Lou, S., Yang, Y., Wang, H., Smith, S.J., Qian, Y., Rasch, P.J., 2019. Black carbon amplifies haze over the North China Plain by weakening the East Asian winter monsoon. *Geophys. Res. Lett.* 46, 452–460.

- Madala, S., Satyanarayana, A.N.V., Srinivas, C.V., Boadh, R., Pinaka Pani, V.V.S.N., Kumar, M., 2015. Simulation of atmospheric pollution dispersion over complex terrain region of Jharkhand with FLEXPART-WRF with incorporation of improved turbulence intensity relationships. *Atmos. Environ.* 123, 139–155.
- Mahmud, M., 2012. Assessment of Atmospheric impacts of biomass open burning in Kalimantan, Borneo during 2004. *Atmos. Environ.* 78, 242–249.
- Ngan, F., Byun, D., Kim, H., Lee, D., Rappenglück, B., Pour-Bazar, A., 2012. Performance assessment of retrospective meteorological inputs for use in air quality modeling during TexAQS 2006. *Atmos. Environ.* 54, 86–96.
- Nguyen, K.C., Noonan, J.A., Galbally, I.E., Physick, W.L., 1997. Predictions of plume dispersion in complex terrain: Eulerian versus Lagrangian models. *Atmos. Environ.* 31, 947–958.
- Peters, L.K., Berkowitz, C.M., Carmichael, G.R., Easter, R.C., Fairweather, G., Ghan, S.J., Hales, J.M., Leung, L.R., Pennell, W.R., Potra, F.A., Saylor, R.D., Tsang, T.T., 1995. The current state and future direction of Eulerian models in simulating the tropospheric chemistry and transport of trace species: a review. *Atmos. Environ.* 29 (2), 189–222.
- Satellite Services Division, Office of Satellite Data Processing and Distribution, NESDIS, NOAA, U.S. Department of Commerce, and National Centers for Environmental Prediction, National Weather Service, NOAA, U.S. Department of Commerce, 2004. NCEP ADP global upper air observational weather data, october 1999 - continuing. Research data archive at the national center for atmospheric research, computational and information systems laboratory. updated daily. <https://doi.org/10.5065/39C5-Z211>. (Accessed 8 January 2018).
- Spitinger, N., Wenig, M., James, P., Wagner, T., Platt, U., Stohl, A., 2001. Satellite detection of a continental-scale plume of nitrogen oxides from boreal forest fires. *Geophys. Res. Lett.* 28 (24), 4579–4582.
- Srinivas, C.V., Hari Prasad, K.B.R.R., Naidu, C.V., Baskaran, R., Venkatraman, B., 2016. Sensitivity analysis of atmospheric dispersion simulations by FLEXPART to the WRF-simulated meteorological predictions in a coastal environment. *Pure Appl. Geophys.* 173 (2), 675–700.
- Stauffer, D.R., Seaman, N.L., 1994. Multiscale four-dimensional data assimilation. *J. Appl. Meteorol.* 33, 416–434.
- Stein, A.F., Draxler, R.R., Rolph, G.D., Stunder, B.J.B., Cohen, M.D., Ngan, F., 2015. NOAA's HYSPLIT atmospheric transport and dispersion modeling system. *Bull. Am. Meteorol. Soc.* 96, 2059–2077.
- Stohl, A., 1998. Computation, accuracy and applications of trajectories-a review and bibliography. *Atmos. Environ.* 32 (6), 947–966.
- Stohl, A., Eckhardt, S., Forster, C., James, P., Spichtinger, N., 2002. On the pathways and timescales of intercontinental air pollution transport. *J. Geophys. Res. Atmos.* 107 (D23), 4684.
- Stohl, A., Forster, C., Eckhardt, S., Spichtinger, N., Huntrieser, H., Heland, J., Schlager, H., Wilhelm, S., Arnold, F., Cooper, O., 2003. A backward modeling study of intercontinental pollution transport using aircraft measurements. *J. Geophys. Res. Atmos.* 108 (D12), 4370.
- Stohl, A., Forster, C., Frank, A., Seibert, P., Wotawa, G., 2005. Technical note: the Lagrangian particle dispersion model FLEXPART version 6.2. *Atmos. Chem. Phys.* 5 (9), 2461–2474.
- Stohl, A., Forster, C., Huntrieser, H., Mannstein, H., McMillan, W.W., Petzold, A., Schlager, H., Weinzierl, B., 2007. Aircraft measurements over Europe of an air pollution plume from Southeast Asia - aerosol and chemical characterization. *Atmos. Chem. Phys.* 7, 913–937.
- Stohl, A., James, P., 2004. A Lagrangian analysis of the atmospheric branch of the global water cycle. Part I: method description, validation, and demonstration for the august 2002 flooding in central europe. *J.Hydrometeo* 5 (4), 656.
- Stohl, A., Trickl, T., 1999. A textbook example of long-range transport: simultaneous observation of ozone maxima of stratospheric and North American origin in the free troposphere over Europe. *J. Geophys. Res. Atmos.* 104 (D23), 30445–30462.
- Su, L., Yuan, Z., Fung, J.C.H., Lau, A.K.H., 2015. A comparison of HYSPLIT backward trajectories generated from two GDAS datasets. *Sci. Total Environ.* 506–507, 527–537.
- Wang, H., Rasch, P.J., Easter, R.C., Singh, B., Zhang, R., Ma, P.L., Qian, Y., Ghan, S.J., Beagley, N., 2014. Using an explicit emission tagging method in global modeling of source-receptor relationships for black carbon in the Arctic: variations, sources, and transport pathways. *J. Geophys. Res. Atmos.* 119 (12), 888–12,909.
- Wang, Z.L., Huang, X., Ding, A.J., 2018. Dome effect of black carbon and its key influencing factors: a one-dimensional modelling study. *Atmos. Chem. Phys.* 18 (4), 2821–2834.
- Wang, Z.L., Huang, X., Ding, A.J., 2019. Optimization of vertical grid setting for air quality modelling in China considering the effect of aerosol-boundary layer interaction. *Atmos. Environ.* 210, 1–13.
- Yang, Y., Russell, L.M., Lou, S., Liao, H., Guo, J., Liu, Y., Singh, B., Ghan, S.J., 2017a. Dust-wind interactions can intensify aerosol pollution over eastern China. *Nat. Commun.* 8, 15333.
- Yang, Y., Wang, H., Smith, S.J., Ma, P.L., Rasch, P.J., 2017b. Source attribution of black carbon and its direct radiative forcing in China. *Atmos. Chem. Phys.* 17, 4319–4336.
- Zhong, Q.R., Huang, Y., Shen, H.Z., Chen, Y.L., Chen, H., Huang, T.B., Zeng, E.Y., Tao, S., 2017. Global estimates of carbon monoxide emissions from 1960 to 2013. *Environ. Sci. Pollut. Res.* 24, 864–873.
- Zhang, J.M., Wang, T., Ding, A.J., Zhou, X.H., Xue, L.K., Poon, C.N., Wu, W.S., Gao, J., Zuo, H.C., Chen, J.M., Zhang, X.C., Fan, S.J., 2009. Continuous measurement of peroxyacetyl nitrate (PAN) in suburban and remote areas of western China. *Atmos. Environ.* 43 (2), 228–237.
- Zhang, Y., Ding, A.J., Mao, H.T., Nie, W., Zhou, D.R., Liu, L.X., Huang, X., Fu, C.B., 2016. Impact of synoptic weather patterns and inter-decadal climate variability on air quality in the North China Plain during 1980–2013. *Atmos. Environ.* 124, 119–128.
- Zhou, D.R., Ding, K., Huang, X., Liu, L.X., Liu, Q., Xu, Z.N., Jiang, F., Fu, C.B., Ding, A.J., 2018. Transport, mixing and feedback of dust, biomass burning and anthropogenic pollutants in eastern Asia: a case study. *Atmos. Chem. Phys.* 18 (22), 16345–16361.