Sensitivity Studies of Four-Dimensional Local Ensemble Transform Kalman Filter Coupled With WRF-Chem Version 3.9.1 for Improving Particulate Matter Simulation Accuracy

Jianyu Lin[1,2], Tie Dai[2*], Lifang Sheng[1], Weihang Zhang[1], Shangfei Hai[3], Yawen Kong[4]

[1] College of Oceanic and Atmospheric Sciences, Ocean University of China, Qingdao, 266100, China

[2] State Key Laboratory of Numerical Modeling for Atmospheric Sciences and Geophysical Fluid Dynamics, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, China

[3] CMA Earth System Modeling and Prediction Centre, China Meteorological Administration (CMA), Beijing 100081

[4] State Key Laboratory of Remote Sensing Science, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100101, China

Correspondence:Tie Dai (daitie@mail.iap.ac.cn)

Abstract: Accurately simulate severe haze events through numerical models remains challenging because of uncertainty in anthropogenic emissions and physical parameterizations of particulate matter ($PM_{2.5}$ and PM_{10}). In this study, a coupled WRF-Chem/four-dimension local ensemble transform Kalman filter (4D-LETKF) data assimilation system has been successfully developed to optimize particulate matter concentration by assimilating hourly

ground-based observations in winter over the Beijing-Tianjin-Hebei region and surrounding

provinces. The effectiveness of 4D-LETKF system and its sensitivity to ensemble member size

and length of assimilation window have been investigated. The promising results show that

significant improvements have been made by analysis in the simulation of particulate matter

during severe haze event. The assimilation reduces root mean square errors of $PM_{2.5}$ from 69.93

to 31.19 μg m^{-3} and of PM_{10} from 106.88 to 76.83μg m^{-3}. Smaller root mean square errors and

larger correlation coefficients in the analysis of $PM_{2.5}$ and PM_{10} are observed across nearly all

verification stations, indicating that the 4D-LETKF assimilation optimizes the simulation of

$PM_{2.5}$ and PM_{10} concentration efficiently. Sensitivity experiments reveal that the combination

of 48 hours of assimilation window length and 40 ensemble members shows best performance

for reproducing severe haze event. In view of the performance of ensemble members, increasing

ensemble member size improves ensemble spread among each forecasting member, facilitates

the spread of state vectors about $PM_{2.5}$ and PM_{10} information in the first guess, favors the

variances between each initial condition in the next assimilation cycle and leads to better

simulation performance both in severe and moderate haze events. This study advances our

understanding about the selection of basic parameters in the 4D-LETKF assimilation system

and the performance of ensemble simulation in a particulate matter polluted environment.

Key words: 4D-LETKF, severe haze simulation, ensemble member size, length of assimilation

window

1. Introduction

Although great progress about air pollution control has been made during recent years,

45   China is facing the highest levels of particulate matter in the world (van Donkelaar et al., 2016).

46   Particulate matter consists of $PM_{2.5}$ and PM_{10}, refers to particles with aerodynamic diameters

47   of less than 2.5 and 10 μm, respectively. High concentration of particulate matter is a major

48   factor for severe haze events (air quality index larger than 200) in the Beijing-Tianjin-Hebei

49   (BTH) region of China, especially during winter (Yan et al., 2016, Zhang et al., 2018).

50   Numerical models are considered to be useful tools for simulating haze events as for taking

51   complex physical and chemical mechanisms into account, but the uncertainty in emissions and

52   physical parameterizations still remain a significant barrier in improving the simulation

53   accuracy (Gao et al., 2017, Feng et al., 2018).

54   As an effective statistical approach, data assimilation is capable of improving the accuracy

55   of pollution simulations by limiting the performance of models. Lots of data assimilation

56   approaches have been applied to the atmospheric science, including three-dimension variation

57   (3D-Var) (Lorenc 1986; Parrish and Derber 1992; Sun et al., 2020), four-dimension variation

58   (4D-Var) (Huang et al. 2009; Benedetti et al., 2009), ensemble Kalman filter algorithms and

59   their variants (Evensen 1994; Whitaker and Hamill 2002; Miyazaki et al., 2012a), etc. Among

60   them, four-dimension local ensemble transform Kalman filter (4D-LETKF) has shown unique

61   characteristics in numerical simulation (Evensen, 2003, Kong et al., 2021). Firstly, derived from

62   finite forecasting members, the background error covariance matrix of 4D-LETKF features

63   flow-dependent characteristics, and the linear combinations of ensemble members produce

64   global analysis (Hunt et al., 2007). Secondly, the computational time for 4D-LETKF remains

65   robust as the observation numbers increase, exhibiting strong computational ability in the

66   parallel architecture when assimilate various measurements (Miyoshi et al., 2007; Hunt et al.,

67   2007, Dai et al., 2021). Lastly, 4D-LETKF can assimilate time slots of asynchronous

68   observations to optimize the current state within the assimilation window, which efficiently

69   improves the quality of pollution prediction (Evensen, 2003, Ott et al., 2004, Dai et al., 2019,

70   Cheng et al., 2019).

71   The characteristics of 4D-LETKF underscore the importance of ensemble member size

72   and length of assimilation window on its effectiveness. The background error covariance matrix,

73   which represents the uncertainty in the ensemble simulations, is determined by the spread of

74   the ensemble members (Peng et al., 2017). 4D-LETKF considers approximate model

75   trajectories by linear combinations of the background ensemble trajectories. However, limited

76   numbers of ensemble members may bring about insufficient dispersion of ensemble systems

77   (Hunt et al., 2004). 4D-LETKF system can greatly improves the utilization rate of observations

78   by constraining the state variables in asynchronous hourly slot within the assimilation window.

79   A longer assimilation window efficiently reduces computational load by avoiding frequent

80   switches between state and forecast variables. But the trajectories over a long length of

81   assimilation window may diverge enough that linear combinations will not approximate the

82   model trajectories. Moreover, the model ensemble trajectory may not fit the observations well

83   over the entire interval with the presence of model errors (Dai et al., 2019). Many studies have

84   discussed the choice of these two parameters for ensemble Kalman filter algorithms and their

85   variants. When optimizing hourly aerosol fields by satellite observations, Cheng et al. (2019)

86   revealed that the forecast with a 24-hour assimilation window was comparable to those with 1-

87   hour, the root mean square error for AOD are 0.091 and 0.110, respectively, indicating the

88   weights determined at the end of the 24 hours assimilation window are valid to optimize the

89 ensemble trajectories. However, Dai et al. (2019) proposed that over 80% of the hourly

90 assimilation efficiencies for the 1-hour assimilation window are higher than those with 6- or

91 24-hours in 4D-LETKF experiments, suggesting that assimilation efficiency decreases with the

92 increase of the assimilation window interval. These different opinions reveal that there is still a

93 large uncertainty about selection of parameters in 4D-LETKF assimilation system.

94   The accuracy simulation of severe haze events with air quality index (AQI) larger than 200

95 has been a challenging problem for a long time, posing severe threats to human daily life and

96 public health (Wang et al., 2014, Kong et al., 2021, Gao et al., 2017). Although 4D-LETKF has

97 unique advantages in computational efficiency and analysis, there is little research that

98 investigate the impacts of 4D-LETKF assimilation on pollutant simulation especially in severe

99 haze events, in addition, it is also imperative to explore the basic optimal combination of

100 assimilation parameters and its explanation in this method. Our major objectives are not only

101 to evaluate the performance of 4D-LETKF in reproducing particulate matter concentration

102 during severe haze event, but also to summarize the influence rules of ensemble size and

103 assimilation window length on particulate matter simulation, and explore whether these rules

104 are applicable to moderate haze event (air quality index smaller than 200) as well. The results

105 have great significance to verify and quantify the effect of 4D-LETKF assimilation on

106 numerical simulations of $PM_{2.5}$ and PM_{10}, subsequently providing a general rule for parameter

107 selection in the 4D-LETKF during severe haze events. Herein, we utilize 4D-LETKF system

108 which was coupled with Weather Research and Forecasting with Chemistry (WRF-Chem)

109 model to improve simulative skill of particulate matter among northern China during the winter

110 of 2020. Section 2 briefly introduces detail setting of WRF-Chem model, 4D-LETKF,

111 observations and numerical experiment designs. Section 3 compares the assimilation with those

112 in the prior simulation, summarizes and explains sensitivity rules for parametric selection, and

113 followed by a conclusion in Section 4 lastly.

114 2. Methodology

115 2.1 Configuration of the forecast model

116 In our implementation, the fully coupled "online" WRF-Chem version 3.9.1 is employed

117 as numeral forward model to describe the meteorological and chemical conditions

118 simultaneously, which fully considers extensive chemical transport processes including

119 advection, convection and sedimentation processes (Grell et al., 2005). The WRF-Chem model

120 is configured with two domains (d01 and d02), both using 100 (west–east) ×100 (south–north)

121 grid points, but with horizontal resolutions of 30 and 10 km, respectively. As shown in Figure

122 1(a), the d01 domain covers most part of East Asia, and the area under the blue shadow is the

123 d02 domain. The vertical grid contains 40 full sigma levels, extending from the surface to 50

124 hPa.

125 The initial and lateral boundary conditions of meteorological fields are derived from the

126 National Centers for Environmental Prediction Final (FNL) analysis data with a spatial

127 resolution of 1°×1° and temporal interval of 6 hours. A state-of-the-art and highly non-linear

128 gas-phase chemical mechanism Regional Atmospheric Chemistry Mechanism (RACM)

129 (Stockwell et al., 1997) is selected as gas phase mechanism, and Goddard Chemistry Aerosol

130 Radiation and Transport (GOCART) (Schwartz et al., 2012) is adopted as aerosol mechanism.

131 The parameterization scheme used in research is shown in Table 1.

132 Table 1. WRF-Chem parameterization scheme in this study.

Parameterization	WRF-Chem option
Microphysics	Morrison 2–moment Scheme (Morrison et al., 2019)
Longwave radiation	RRTMG Longwave Scheme (Iacono et al., 2008)
Shortwave radiation	RRTMG Shortwave Scheme (Iacono et al., 2008)
Planetary boundary layer	YSU Scheme (Hong et al., 2006)
Cumulus parameterization	Grell 3D Ensemble Scheme (Grell et al., 1993)
Land surface model	Noah (Tewari et al., 2004)

133

134   The anthropogenic emissions are obtained from the Multi-resolution Emission Inventory

135 for China (MEIC, 2024) compiled by Tsinghua University. The inventory includes

136 anthropogenic emissions from agriculture, industry, power, residential and transportation

137 sectors (Zheng et al., 2021). Inventory with a spatial resolution of $0.25° \times 0.25°$ and has been

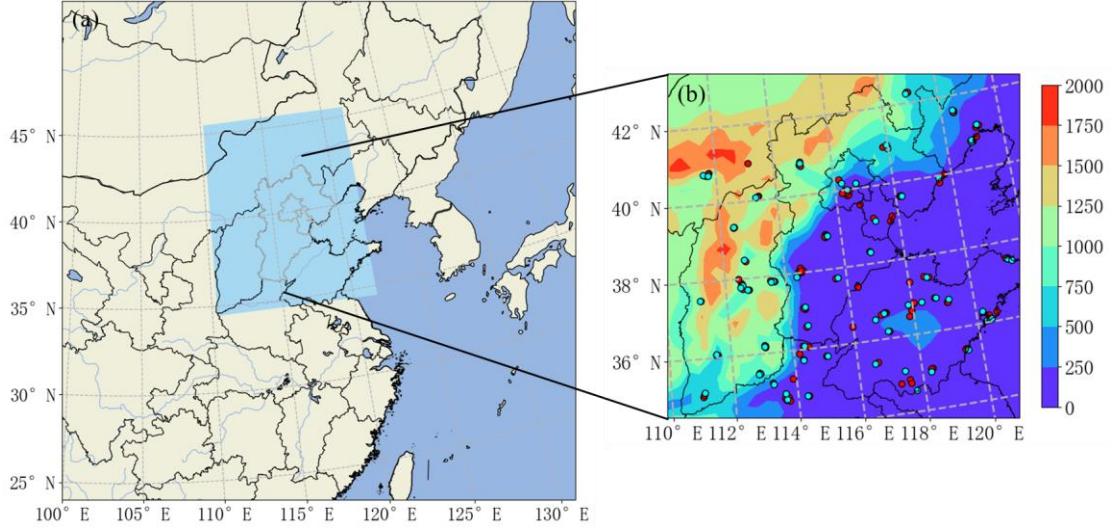
138 interpolated to match the simulation resolution. The biogenic emissions are calculated online

139 by Guenther scheme (Guenther et al. 1995). The $PM_{2.5}$, PM_{10} concentrations output from WRF-

140 Chem are linearly interpolated to site observations. The evaluation of uncertainty in the

141 emission inventory has been shown in previous research (Zhang et al., 2009).

142

143

144 Figure 1. (a) WRF-Chem model domains. The grey border implies the BTH region. (b)

145 Location of assimilated and independent verification observation sites with topography (units:

146 m). The red and blue dot implies the assimilated and independent verification observation

147 site, respectively.

148 2.2 The 4D-LETKF algorithm and the state variables

149 The 4D-LETKF coupled with WRF-Chem Model is implemented to investigate the

150 influence of assimilation on particulate matter simulation in this research. In this section, we

151 introduce the 4D-LETKF algorithm and corresponding state variables briefly, more detailed

152 information can be found in Hunt et al. (2007). The LETKF features a flow-dependent

153 covariance matrix from ensemble simulation and determines the analysis ensemble mean $\overline{x^a}$

154 (a posteriori) according to the following formula

155 $$\bar{x}^a = \bar{x}^b + X^b \overline{w}^a \tag{1}$$

156 where $\bar{x}^b$ and X^b denote ensemble mean of first guess and background ensemble

157 perturbations, respectively. The ensemble perturbation matrix X^b is calculated as $x^b(i) - \bar{x}^b$

158 $\{i = 1, 2, ..., k\}$, which k represents the ensemble member size. The perturbation weight matrix

159 $\overline{w}^a$ is the Kalman gain which linearly determines the increment between the analysis and the

160   first guess, and can be calculated as

161    $$\bar{w}^a = \widetilde{P^a}(Y^b)^T R^{-1}(y^0 - \bar{y}^b) \tag{2}$$

162   where $\widetilde{P^a}$ is analysis error covariance in ensemble space. y^0 and $\bar{y}^b$ denote the

163   observations vector and ensemble mean background observations, respectively. Ensemble

164   mean background observations derived from applying observation vector to ensemble member

165   state vector H ($\bar{x}^b$). The matrix R is the observation error covariance matrix. The matrix Y^b

166   represents ensemble background observation perturbations, whose i th columns is $y^b(i) - \bar{y}^b$,

167    $\{i = 1, 2, ..., k\}$. $\widetilde{P^a}$ can be obtained as

168    $$\widetilde{P^a} = [(k - 1)I/\rho + (Y^b)^T R^{-1} Y^b]^{-1} \tag{3}$$

169   where I denotes the identity matrix and k is ensemble member size. The multiplicative inflation

170   factor ρ is necessary to avoid the filter divergence (Anderson and Anderson, 1999, Sekiyama

171   et al., 2010), so which is set to 1.1 to inflate the analysis covariance (Dai et al., 2019, Anderson,

172   2007). Analysis ensemble perturbations X^a is calculated by

173    $$X^a = X^b[(k - 1)\widetilde{P^a}]^{1/2} = X^b W^a \tag{4}$$

174   Calculated by the sum of the $\bar{x}^a$ and each of the columns of X^a, the ensemble analyses are

175   served as optimal initial conditions in each ensemble member to generate the first guess in the

176   next cycle.

177   Figure 2 is the flow chart of the WRF-Chem/4D-LETKF assimilation system applied in

178   our implementation. The system conducts these processes within each assimilation cycle. The

179   4D-LETKF generates a flow-dependent background error covariance matrix by ensemble

180   member. Given that the emissions inventory is an important source of uncertainty in simulation

181   (Pagowski and Grell, 2012), the research randomly perturbs anthropogenic emissions of PM,

182 black carbon (BC) and organic carbon (OC) in January for each member to create the ensemble

183 members, and the perturbation follows a log-normal distribution in the k-dimensional space.

184 The mean values of perturbations of $PM_{2.5}$, PM_{10}, BC and OC emissions are equal to 1, and the

185 variances of these emissions are set according to corresponding uncertainty in MEIC inventory

186 (130%, 132%, 208%, 258% for $PM_{2.5}$, PM_{10}, BC, OC, respectively) (Luo et al., 2023). Such

187 ensemble anthropogenic emissions are perfect correlation in spatial and temporal dimension

188 and should not be regarded as overly restrictive (Schutgens et al., 2010, Dai et al., 2021). This

189 study only adds one times of perturbations into emissions at the first cycle of assimilation to

190 provide the information spread of particulate matter. It is enough to make most of the site

191 information affect the first-guess field. The WRF-Chem/4D-LETKF system propagates the

192 ensemble forward simulation for the entire assimilation window time and outputs the first guess

193 fields at each hourly time slot. The ensemble mean of first guess $(\bar{x}^b)$ and background ensemble

194 perturbations (X^b) can be obtained from ensemble member here. Combining observation and

195 observation operator, the innovation $(y^0 - \bar{y}^b)$ and Y^b can be obtained in each time slot. The

196 perturbation weight matrix $\bar{w}^a$ is valid within a relative short assimilation window (e.g., 24 or

197 48 hours) (Hunt et al., 2004, Cheng et al., 2019). The analysis ensemble derived from $\bar{w}^a$ at

198 the end of time slots will serve as chemical initial conditions for the next assimilation window.

199 As the cycle of assimilation proceed, a linear combination of analysis ensemble is continuously

200 obtained.

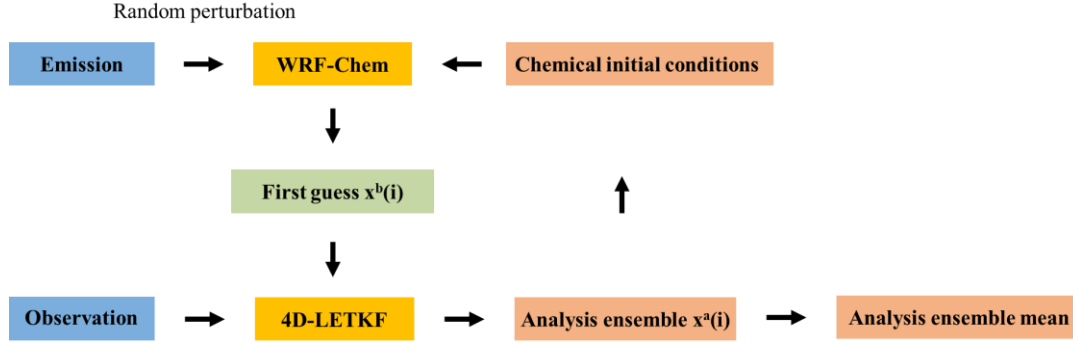



Figure 2. Flow chart of the WRF-Chem/4D-LETKF assimilation system for particulate

matter.

The ensemble Kalman filter generally encounters a spurious long-distance correlation

problem because of the limited numbers of ensemble members (Miyazaki et al., 2012a). To

avoid the problem above, it is necessary to apply observation localizations to filter observation

from a long distance. 4D-LETKF offers a flexible choice of observation localizations in

horizontal, vertical and temporal dimensions for each grid point (Cheng et al., 2019). In this

study, the horizontal localization factor is calculated as Gaussian function (Miyoshi et al., 2007),

which gradually reduces the effect of observation as the increasing departure from the analysis

grid

$$f(r) = \exp(-r^2/2\sigma^2) \tag{5}$$

Here, r represents physical distance from observation to analysis grid and σ represents

localization length. We limit the localization factor from 0 to 3.65 times the localization length

(Zhao et al., 2015), ignoring the observation beyond 3.65 times the localization length to the

analysis grid.

The selection of the state variables depends on the generative mechanism of aerosol. As a

result, 16 kinds of WRF-Chem/GOCART aerosol variables are treated as state variables. For

the $PM_{2.5}$ observations, the observation operator is described as

220    $y^b_{PM_{2.5}} = \rho_d[P_{2.5} + 1.375S + 1.8(OC_1 + OC_2) + BC_1 + BC_2 + D_1 + 0.286D_2 + S_1 +$

221    $0.942S_2]$ (6)

222   where ρ_d present the dry-air density, $P_{2.5}$ is the fine unspecified aerosol contributions,

223   S represents sulfate, OC_1 and OC_2 are hydrophobic and hydrophilic organic carbon, respectively.

224   BC_1 and BC_2 are hydrophobic and hydrophilic black carbon, D_1 and D_2 are dusts with effective

225   radii of 0.5 and 1.4 μm, and S_1 and S_2 are sea salts with effective radii of 0.3 and 1.0 μm,

226   respectively (Peng et al., 2018).

227   Similarly, the observation operator for PM_{10} is shown as below

228    $y^b_{PM_{10}} = \rho_d[P_{10} + P_{2.5} + 1.375S + 1.8(OC_1 + OC_2) + BC_1 + BC_2 + D_1 + 0.286D_2 + D_3 +$

229    $0.87D_4 + S_1 + 0.942S_2 + S_3]$ (7)

230   where P_{10} is coarse unspecified aerosol contributions, D_3 and D_4 are dusts with

231   effective radii of 2.4 and 4.5 μm. S_3 is sea salt with effective radii of 3.2 μm. Therefore, the

232   simulated $PM_{10\text{-}2.5}$ is

233    $y^b_{PM_{10-2.5}} = \rho_d[P_{10} + D_3 + 0.87D_4 + S_3]$ (8)

234   In this research, $y^o_{PM_{10-2.5}}$ calculated by $y^o_{PM_{10}} - y^o_{PM_{2.5}}$ is used to analyze state variables

235   including D_5 and S_4, which are dust with effective radii of 8 μm and sea salt with effective

236   radii of 7.5 μm, respectively.

237   2.3 Site observation data and errors

238   Ground-based observation features high temporal resolution, which can capture variation

239   of pollution concentration on an hourly scale at the bottom of the troposphere, providing

240   continuous and reliable observation. The quality-assure and quality-controlled hourly

241   observation data of $PM_{2.5}$ and PM_{10} are used to explore the influence of 4D-LETKF assimilation

242 in this research. The pollution data was obtained from China National Environmental

243 Monitoring Center (CNEMC, 2024). As the research primarily focuses on the BTH region, the

244 assimilation and verification sites are mainly located in the BTH region and neighboring

245 provinces, primarily located in urban and suburban areas. In order to obtain more reliable

246 observation data, the quality control of observation data in this study includes hourly

247 observation of default value and extreme value detection. First, during the haze period, if the

248 number of missing values for either type of pollutant at one site exceeds 24 hours, this site is

249 considered to have a certain uncertainty on observation quality, and data will not be assimilated.

250 Second, for each kind of observation in different station, the hourly observation outside the

251 range of m±3σ will not be assimilated, where the m and σ are the mean value and standard

252 deviation of daily concentration, respectively. When selecting assimilation and verification sites,

253 spatial distribution uniformity is ensured for better assimilation performance, consequently,

254 those sites are randomly selected. Finally, 127 assimilation sites and 69 verification sites in the

255 BTH region and surrounding province are selected (Figure 1b). It can be seen that the

256 assimilation and verification sites have a relatively uniform spatial distribution.

257 The observation error covariance matrix (R) is assumed to be diagonal, implying that

258 observational errors among each pollution species are uncorrelated. The observation error (r)

259 consists of measurement error (ε_0) and representation error (ε_r)

260 $$r = \sqrt{\varepsilon_0^2 + \varepsilon_r^2} \tag{9}$$

261 The measurement error ε_0 is defined as

262 $$\varepsilon_0 = ermax + 0.0075 \times \Pi_0 \tag{10}$$

263 where ermax is the base error, which is set to be 1 for $PM_{2.5}$, and PM_{10} (Chen et al., 2019a),

264 Π_0 denotes the observation of concentration. Produced by observation operator,

265 representativeness errors can be calculated by the formula (Elbern et al., 2007)

266 $\varepsilon_r = \gamma \, \varepsilon_0 \sqrt{\Delta l / L}$ (11)

267 γ is tunable scaling factor and 0.5 is set for γ, Δl is the spatial resolution of gridding (30 km

268 and 10km for d01 and d02, respectively), L depends on station location, which denotes the

269 range that an observation can reflect, here L is 2 km for calculation.

270 Meteorological data were collected from National Climatic Data Center (NCDC, 2024),

271 which provides hourly air temperature, dew point, and windspeed data. The observational

272 meteorological data are used to validate the performance of simulations in this study.

273 2.4 Experiment design

274 A series of control and data assimilation experiments during severe and moderate haze

275 events, as listed in Table 2, have been carried out to achieve our major objective. The control

276 experiments refer to numerical experiments without data assimilation. The Severe-FR

277 experiment with 48 hours spin up time is performed firstly to quantify the necessity of adjusting

278 particulate matter concentration during severe haze event. Severe-FR-24h, Severe-FR-48h, and

279 Severe-FR-72h accompany with restart every 24, 48, and 72 hours respectively and update

280 meteorological boundary conditions. Except Severe-FR, the rest of the experiments all have 24

281 hours of spin up time at the beginning of each restart or assimilation cycle. The settings of

282 assimilation experiments cycle time are same with corresponding control experiments. The

283 detail descriptions of experiment cycle time are shown in Figure S1. Since the effectiveness of

284 4D-LETKF is highly related with ensemble member size and length of assimilation window

285 (Rubin et al., 2016), the sensitivity analysis is employed to investigate the influence from two

286 parameters on assimilation effect (Kong et al., 2023). The selection of assimilation parameters

287 for the sensitivity experiments includes 20, 40 and 60 for ensemble members, and 24, 48 and

288 72 hours for the length of assimilation window empirically (Kong et al., 2021, Dai et al., 2021).

289 All sensitivity experiments use identical WRF-Chem physical parameterizations,

290 anthropogenic emission and random perturbations. Through the comparison between all

291 assimilation experiments, the influence rules of 4D-LETKF assimilation on the simulation of

292 particulate matter in severe haze can be retrieved. Lastly, aiming to determinate the applicable

293 range of obtained influence rules above, two assimilation experiments in a moderate haze event

294 are performed to validate whether the rules are also suitable to a less-polluted environment. The

295 detail reasons for selection of parameters will be fully described in the next section.

296 Table 2. design of numerical experiments in this research.

	Experiment	Design of simulation
Control experiments	Severe-FR	Free run experiment in severe haze event and without restart in integration process.
	Severe-FR-24h	Free run experiment in severe haze event and with restart every 24, 48 and 72 hours, provide deterministic simulation corresponding to data assimilation experiment.
	Severe-FR-48h	
	Severe-FR-72h	
	Moderate-FR-48h	Free run experiment in moderate haze events and with restart every 48 hours, provide deterministic simulation corresponding to data assimilation experiment.

Data assimilation experiments in severe haze event	Severe-20m-24h Severe-20m-48h Severe-20m-72h	Assimilation experiment in severe haze event with 20 ensemble members and 24, 48, 72 hours of assimilation window length respectively.
	Severe-40m-24h Severe-40m-48h Severe-40m-72h	Assimilation experiment in severe haze event with 40 ensemble members and 24, 48, 72 hours of assimilation window length respectively.
	Severe-60m-24h Severe-60m-48h Severe-60m-72h	Assimilation experiment in severe haze event with 60 ensemble members and 24, 48, 72 hours of assimilation window length respectively.
Data assimilation experiments in moderate haze event	Moderate-20m-48h Moderate-40m-48h	Assimilation experiment in moderate haze event with 20 and 40 ensemble members combine with 48 hours of assimilation window length.

297   Root mean square error (RMSE), mean errors (BIAS), mean absolute error (MAE) and

298 correlation coefficient are calculated in this study to evaluate the performance of each numerical

299 experiment. The assimilation efficiency (AE) for estimating the data assimilation performance

300 is also calculated from the formulation below (Yumimoto and Takemura, 2011)

301    $AE = \frac{RMSE^f - RMSE^a}{RMSE^f} \times 100\%$ (12)

302 where $RMSE^a$ and $RMSE^f$ is RMSE with and without assimilation, respectively. According

303 to the definition, if AE is positive, it means that RMSE has decreased due to assimilation effect.

304 When AE is equal to 1, RMSE in analysis completely disappears, and analysis is equal to

305 observation. The formulations of correlation coefficient and root mean square error are shown in

306 supplement.

307 3. Results

308 3.1 Comparison of the analysis with control experiment

309 3.1.1 The reproduction of the severe haze event in BTH

310   It is essential to discuss the basic evolution of pollutant and the necessity of pollutant data

311 assimilation in severe haze event before conducting the assimilation experiments. The severe

312 haze event selected in this study occurred from 00:00 UTC 15 January 2020 to 00:00 UTC 21.

313 The AQI is a comprehensive indictor of overall air pollution and criterion for severe haze events

314 (Zhan et al., 2018, Bao et al., 2015). As a result, Figure 3(a) shows the temporal variation of

315 AQI at the six sites among BTH region during the investigated period. The peak value of AQI

316 mainly appeared on 18 January, and then rapidly decreased on 19 and 20 January. The temporal

317 averaged of AQI have exceeded 200, with particulate matter identified as the primary pollutant.

318 Fig. 3(b) provides the correlation coefficients and standardized standard deviations of five

319 parameters from Severe-FR against observations. Meteorological variables including air

320 temperature, dew point temperature and wind speed are well simulated when compared with

321 $PM_{2.5}$ and PM_{10}. The correlation coefficients of meteorological factors are all larger than 0.6,

322 while that of pollutant concentrations are all below 0.4. Therefore, when the meteorological

323 conditions can be retrieved relatively accurately, particulate matter assimilation is the key to

324 improving the simulative skill of pollutants.

325

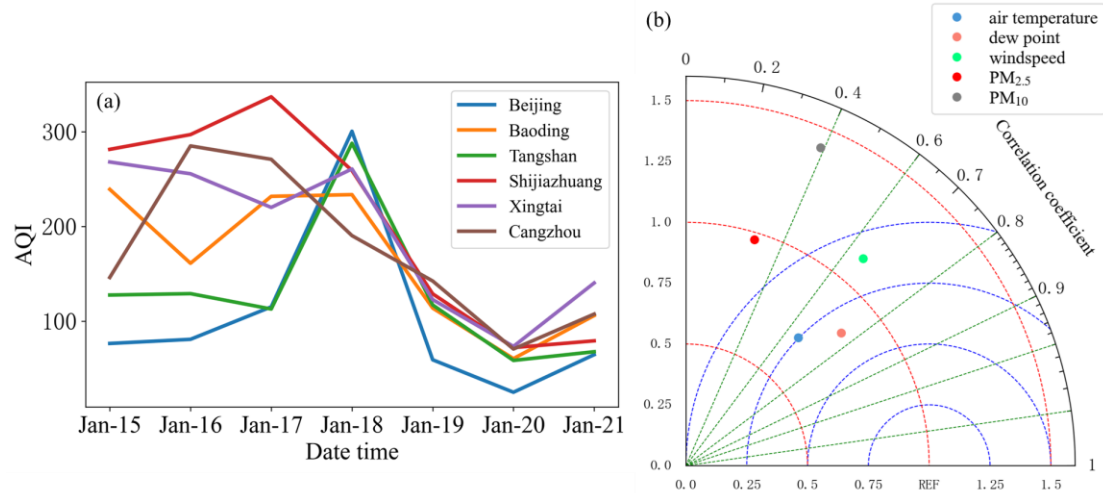



Figure3. (a) Temporal variation about AQI at six sites in severe haze event. (b) A Taylor graph

describing simulation from Severe-FR about five kinds of parameters compared with the

observed ones in BTH region.

3.1.2 The improvement in the reproduction of severe haze achieved by 4D-LETKF

The divergence between assimilation and control experiment reflects the contribution from

4D-LETKF initial condition adjustment. Consequently, the study takes an ensemble member

size of 40 and assimilation window length of 48 hours to conduct sensitivity experiment and

compare with Severe-FR-48h which has the same integration time in each cycle to validate the

effectiveness of 4D-LETKF assimilation system (the analysis from the selection of 40 ensemble

members and 48 hours of assimilation window length is presented here because it shows the

best performance among sensitivity experiments in the next section). Figure 4 reveals the

performance of control and assimilation experiments in severe haze event. The RMSE values

of $PM_{2.5}$ and PM_{10} in Severe-FR-48h are 69.93 and 106.88μg m^{-3} and both with scattered

distribution, indicating substantial uncertainty exist in reproducing this severe haze event. In

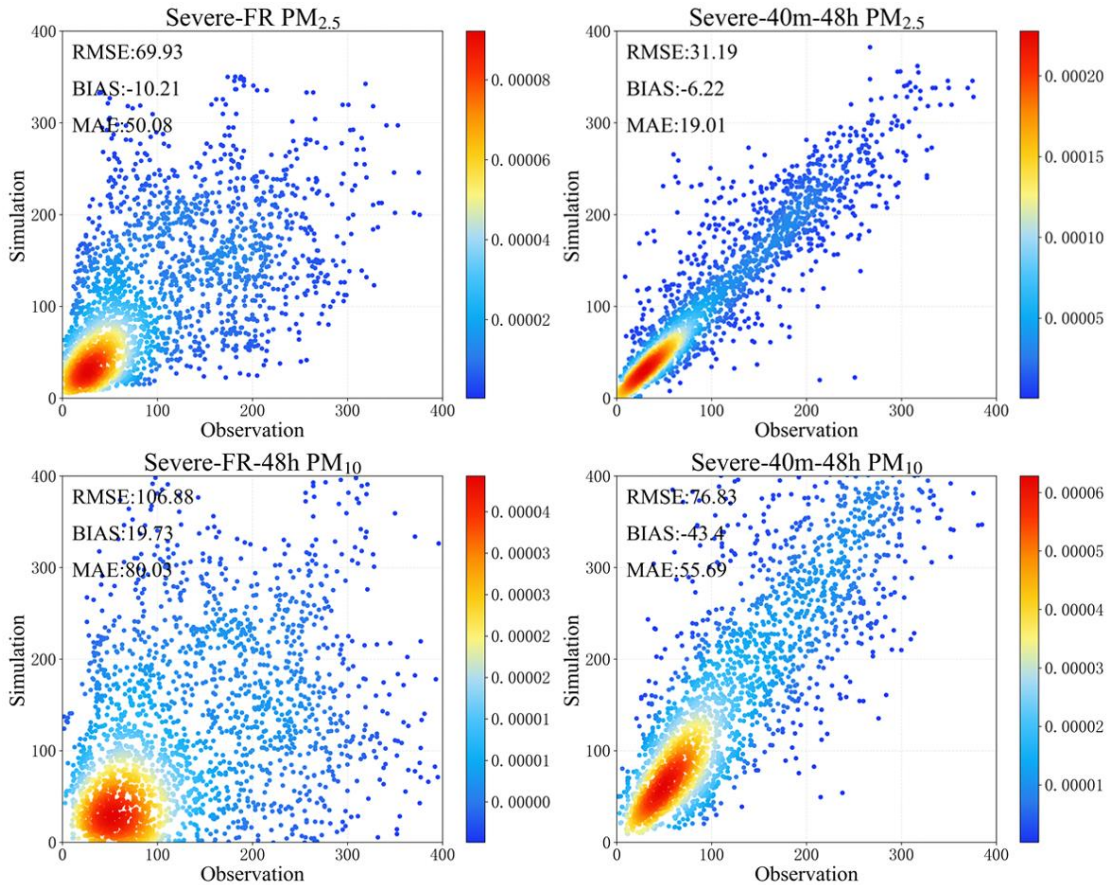
Severe-40m-48h, the RMSE values of $PM_{2.5}$ and PM_{10} are 31.19 and 76.83μg m^{-3}, decreasing

by 55.40% and 28.12% respectively in a high particulate matter concentration environment.

343 The decreased RMSE values also imply that the assimilation system has reached a well-

344 calibrated stage. Not only more points are getting together, but smaller simulation errors for

345 PM$_{2.5}$ and PM$_{10}$ also imply that the Severe-40m-48h outperforms the Severe-FR-48h in this

346 severe haze event.

348 Figure 4. Scatter and density plot of PM$_{2.5}$ and PM$_{10}$ in Severe-FR-48h and Severe-40m-

349 48h versus observations from verification stations (units: μg m^{-3}). The color bar represents the

350 Gaussian Kernel Density estimation.

351 In order to acquire basis distribution of simulation errors for particulate matter, Figure 5

352 presents the frequency distribution of deviations between observed and simulated particulate

353 matter concentrations in Sevre-FR-48 and Severe-40m-48h experiments. It is obviously that

354 Severe-40m-48h increases the frequency of low deviations and decrease those of high

355 deviations in the simulation of $PM_{2.5}$. The deviation pattern of $PM_{2.5}$ in Severe-40m-48h is

356 generally squeezed with higher peaking and symmetrical to the value of 0 than Severe-FR-48h.

357 For the deviation distribution pattern of PM_{10}, it shows high frequency of negative deviations

358 and great underestimation in the Severe-FR-48h, and this underestimation has been effectively

359 corrected by the adjustment of initial conditions and step analysis in Severe-40m-48h. Specially,

360 the proportion of deviation within 20 µg m^{-3}in the Severe-40m-48h is 69.98% for $PM_{2.5}$ and

361 31.90% for PM_{10}.

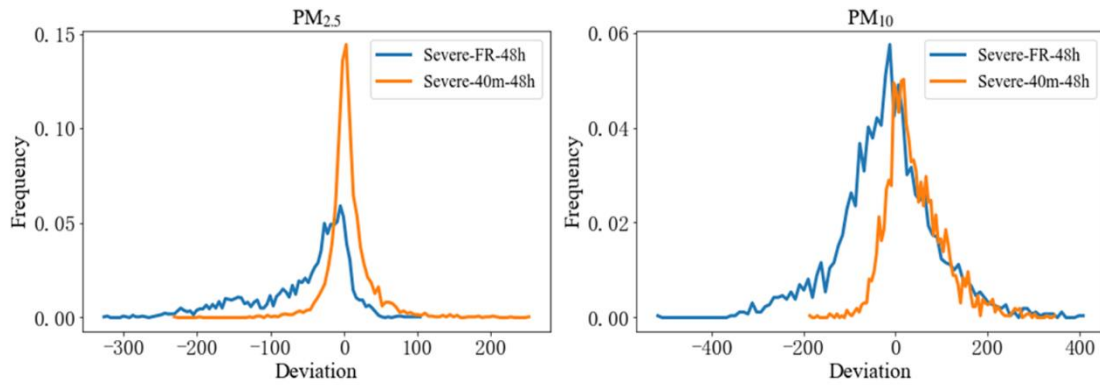



Figure 5. Frequency distribution of the deviations about the simulated $PM_{2.5}$ and PM_{10}

concentrations in Severe-40m-48h and Severe-FR-48h minus the observed ones (units for

deviation: µg m^{-3}).

366 Figure 6 exhibits the spatial distribution of four statistical parameters about RMSE for

367 particulate matter among the BTH region. By comparison from the Severe-FR-48h and Severe-

368 40m-48h, there are significant RMSE reduction for $PM_{2.5}$ after assimilation, implying that the

369 actual evolution of $PM_{2.5}$ can be better represented by Severe-40m-48h. For instance, the RMSE

370 values of $PM_{2.5}$ in Baoding, Hengshui and Cangzhou, have significantly decreased to 29.85,

371 18.98, and 19.06µg m^{-3}, respectively, compared to 80.55, 55.22 and 76.32 µg m^{-3}in the Severe-

372 FR-48h. AE in most verification stations has exceeded 50% also suggests the high efficiency

of 4D-LETKF assimilation for the simulation of $PM_{2.5}$. Although the performance of

assimilation experiment in Shijiazhuang city does not have a good agreement with observation

and shows a positive difference, high values of AE in most of verification stations also proves

the validation of assimilating effect for PM_{10}. Compared to the Severe-FR-48h, the Severe-

40m-48h productively reduces the RMSE of PM_{10}, accompany with high values of 61.18%,

59.17% and 52.18% about AE on Zhangjiakou, Tangshan and Hengshui, respectively.

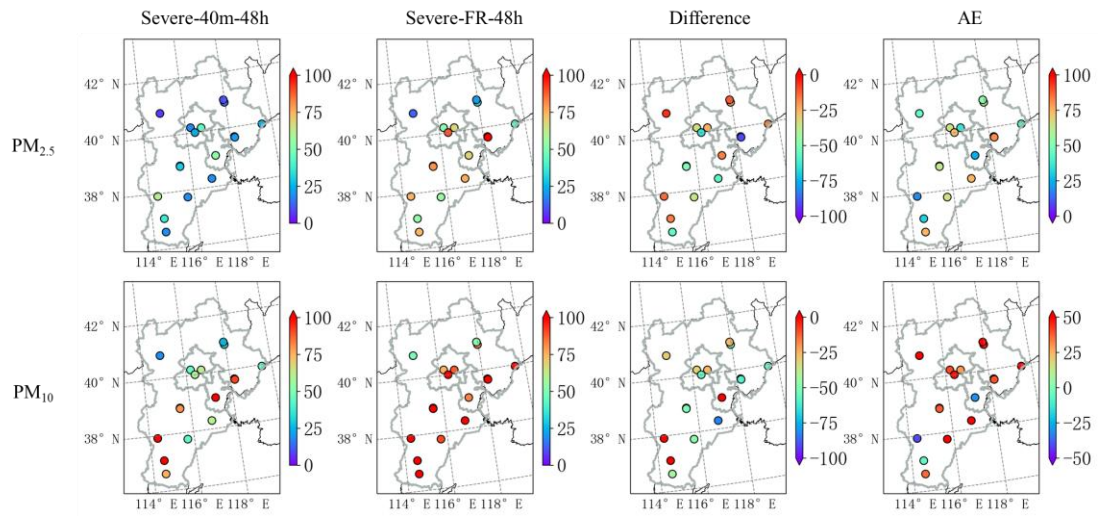



Figure 6. Spatial distribution of RMSE values from Severe-40m-48h (first column), Severe-

FR-48h (second column), their difference (third column) and AE (fourth column) for $PM_{2.5}$

(first row) and PM_{10} (second row) from 15 January to 21 January among verification station

in BTH region. The difference implies the RMSE in Severe-40m-48h minus those in Severe-

FR-48h (units: µg m^{-3}). AE is assimilation efficiency and has been described in methodology

before. The grey border implies the BTH region.

The spatial distribution of correlation coefficients from Severe-40m-48h, Severe-FR-48h,

their difference for $PM_{2.5}$ and PM_{10} are also illustrated in Figure 7. The assimilation experiment

increases the correlation coefficients to more than 0.6 at all sites in the simulations of $PM_{2.5}$ and

exceed 0.7 among the southern BTH region in the simulations of PM_{10}. The Severe-40m-48h

390   also reverses the opposite trend of $PM_{2.5}$ and PM_{10} series in Severe-FR-48h versus observations,

391   for example, the correlation coefficients in Severe-FR-48h at Chengde and Zhangjiakou are -

392   0.42 and -0.53, but increase to 0.52 and 0.69 after assimilation in the simulations of PM_{10}.

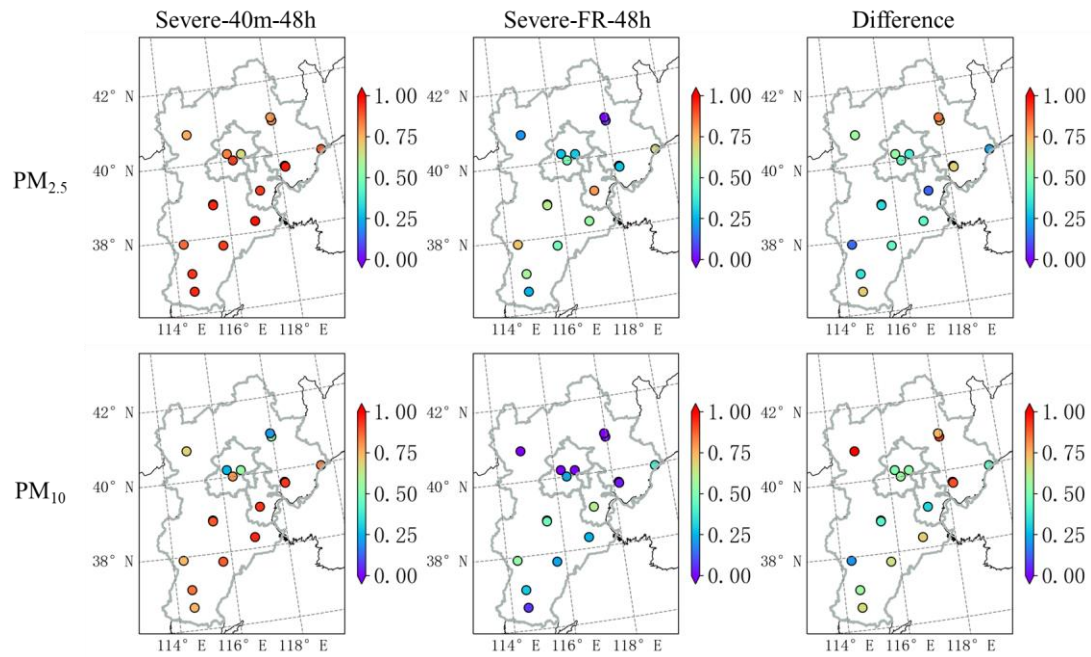
393   Incorporating more assimilable observations may further increase the correlation coefficient in

394   the simulation of particulate matter (Kong et al., 2021). Data assimilation by multiple

395   observations from diverse platform is necessary because it can integrate and coordinate

396   observational information into aerosol forecasts well and then improve air pollutant forecast

397   accuracy (Barbu et al., 2009, Ma et al., 2020).

398

400   Figure 7. Spatial distribution of correlation coefficients from Severe-40m-48h (first column),

401   Severe-FR-48h (second column), their difference (third column) for $PM_{2.5}$ (first row) and

402   PM_{10} (second row) from 15 January to 21 January among verification station in BTH region.

403   The difference implies the correlation coefficient in Severe-40m-48h minus those in Severe-

404   FR-48h. The grey border implies the BTH region.

405   The temporal variations of particulate matter from Severe-40m-48h, Severe-FR-48h and

406   observation at six independent verification stations are shown in Figure S2 and Figure S3. The

407   six independent verification stations have experienced different levels of air pollution and

408   distributed uniformly over BTH region. It is apparently that the analysis at six stations have

409   good agreement with observations both for $PM_{2.5}$ and PM_{10}, which can better characterize the

410   peaks and valleys of particulate matter concentration over investigated period.

411   Table 3 lists the Δ RMSE, Δ CORR and AE in the simulations of particulate matter at

412   independent stations outside the BTH region. The RMSEs and correlation coefficients have

413   decreased and increased respectively after assimilate ground-based observations, suggesting

414   that the uncertainty in Severe-FR-48h has been well optimized not only in the BTH region, but

415   also includes the whole simulation domain. The main source of these gains is generated from

416   local initial field assimilation. Compared to the Severe-FR-48h, the analysis in Yuncheng shows

417   that the RMSE values of $PM_{2.5}$ and PM_{10} have decreased by 98.26 and 144.56 $\mu g\ m^{-3}$ remarkably,

418   such a great improvement may relate to the enhanced estimation capability about state variables

419   of particulate matter. The high values of AE also suggest that verification observation sites

420   outside the BTH region have achieved a good Kalman gain. In previous researches, predicting

421   heavy haze events in northern China, especially over the BTH Region, remained a challenge

422   when compared to other regions like Pearl River Delta and Yangtze River Delta in China (Feng

423   et al., 2018, Gao et al., 2017). In this research, the analysis is propagated by meteorological

424   elements including temperature, air pressure and wind fields come from NCEP Final analysis

425   data, which may provide an optimal meteorological boundary conditions for the assimilation of

426   pollutant concentration.

427 Table 3. Statistics about PM$_{2.5}$ and PM$_{10}$ from analysis in the cities among neighboring

428 provinces of BTH region. ΔRMSE (ΔCORR) represent the RMSE (correlation coefficient)

429 from analysis minus those from Severe-FR-48h (units for ΔRMSE: μg m^{-3}).

City/	PM$_{2.5}$			PM$_{10}$		
Statistical variable	ΔRMSE	ΔCORR	AE	ΔRMSE	ΔCORR	AE
Taiyuan	-21.30	+0.41	23.93%	-54.8	+0.71	39.05%
Changzhi	-38.39	+0.38	65.93%	-63.06	+0.68	63.22%
Jincheng	-37.94	+0.37	66.85%	-94.32	+0.89	72.45%
Shuozhou	-27.07	+0.31	58.96%	-100.08	+0.84	69.96%
Yuncheng	-98.26	+0.67	77.85%	-144.56	+1.25	80.64%
Hohhot	-92.30	+0.67	74.53%	-121.79	+1.41	68.92%
Chifeng	-16.90	+0.55	64.95%	-38.18	+0.95	60.85%
Huludao	-38.56	+0.20	59.11%	-95.95	+1.04	65.76%
Jinzhou	-42.97	+0.21	61.17%	-46.26	+0.83	45.83%
Chaoyang	-39.37	+0.37	51.14%	-83.04	+1.23	61.86%
Jinan	-44.90	+0.63	69.71%	-71.22	+0.59	62.93%
Qingdao	-23.99	+0.27	37.98%	-72.05	+0.44	72.21%
Shouguang	-28.70	+0.21	58.03%	-48.92	+0.39	47.57%
Anyang	-35.87	+0.41	62.53%	-32.08	+0.60	33.15%
Zhengzhou	-26.26	+0.37	37.51%	-3.64	+0.42	3.73%

430 3.2 The sensitivity of 4D-LETKF to ensemble member size and length of assimilation

431   window

432   In previous section, the research has compared the performance from assimilation

433   experiment with 40 ensemble members and 48 hours of assimilation window length against that

434   do not integrate hourly pollutant observations. The results fully demonstrate the ability of 4D-

435   LETKF assimilation method to reproduce severe haze events in spatial and temporal

436   dimensions. However, the 4D-LETKF assimilation effect is highly rely on selection of

437   ensemble member size and length of assimilation window, so how does the assimilation

438   approach vary to the parameterized selection in severe haze event? It is of great meaning to

439   conduct sensitivity experiments based on ensemble member size and length of assimilation

440   window, compare each performance of them by statistical metrics, and summarize the general

441   influence rule of 4D-LETKF parameter selection. Consequently, nine panels of sensitivity

442   experiments are conducted with the selection of ensemble member size (20, 40, 60 members)

443   and the length of assimilation window (24, 48, 72 hours) to maximize the positive innovation

444   in this section.

445   Figure 8 reveals the heatmap about RMSE in each sensitivity experiment of particulate

446   matter over verification sites among the BTH region. The results of free run experiment with

447   different integration times (24, 48, 72 hours) are offered here for comparison with analysis

448   which with same assimilation cycle time. The RMSEs of $PM_{2.5}$ and PM_{10} in each free run

449   experiment exceed 60μg m^{-3} and 100μg m^{-3}, respectively. It is apparently that the 4D-LETKF

450   performs better than the FR experiment in the simulation about $PM_{2.5}$ and PM_{10} over wide range

451   of ensemble member sizes and assimilation window lengths, illustrating the broad applicability

452   of 4D-LETKF data assimilation to these parameters. However, it can be found that the analysis

453   of $PM_{2.5}$ and PM_{10} are dependent on length of assimilation window and dramatically related to

454   ensemble member size in all sensitivity experiments. Unlike the short-lived and chemical

455   reactive species (such as SO_2 and NO_2) which easily undergo complex and nonlinear

456   photochemical reactions, a relative longer assimilation window length seems more suitable for

457   assimilating ground-based particulate matter observations (Peng et al., 2017, Kong et al., 2021).

458   A longer assimilation window length could also avoid underestimation of model spread, which

459   implies overconfidence in the first guess state estimate by enough integration time of each

460   member (Schutgens et al., 2010, Miyazaki et al., 2012a, Hunt et al., 2007). Hence, 48 or 72

461   hours of assimilation window length are advised to optimize the ensemble concentration

462   trajectories. On the other hands, increasing ensemble member size efficiently reduces

463   uncertainty in $PM_{2.5}$ and PM_{10}, as evidenced by the decrease of RMSEs from free run to

464   assimilation experiments with 20 and 40 members. However, when compared with the results

465   from 40 ensemble members, the accuracy of numerical simulations has not significantly

466   improved for both $PM_{2.5}$ and PM_{10} with 60 ensemble members, indicating that 40 members are

467   sufficient and feasible to provide a reliable estimation of the background error and analysis

468   rather than more numerical source consumption. Considering numerical source consumption

469   and RMSE values in the simulations of $PM_{2.5}$ and PM_{10}, the Severe-40m-48h shows more

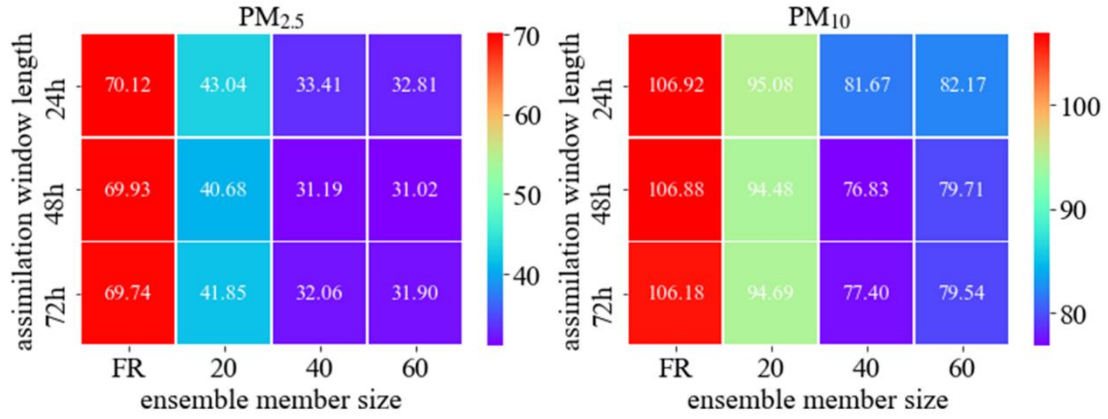
470   comparable to the observations when compared with the other eight panels of sensitivity

471   experiments.

Figure 8. Heatmap about RMSE in each sensitivity experiment of particulate matter over

verification sites (units: μg m^{-3}). The number in each small square represents the RMSE

between observation and simulation for each combination of ensemble member size and the

length of assimilation window methods.

3.3 The influence from ensemble member size to the ensemble spread

In order to explore why increasing ensemble member size can efficiently reduce the

uncertainty in the analysis of PM$_{2.5}$ and PM$_{10}$ as revealed in Figure 8, the study investigates the

spatial distribution of standard deviations of PM$_{2.5}$ and PM$_{10-2.5}$ among first guess ($x^b(i)$) and

analysis field ($x^a(i)$) in terms of ensemble members. The standard deviations of ensemble

members describe how the emission perturbation propagates among the forward model, and

this perturbation is driven by the underlying surface pollution emission inputs and the

meteorological conditions. Therefore, the standard deviation in the first guess fields quantifies

the dispersion degree of the ensemble background, substantially impacts the calculation of

assimilation parameters such as ensemble state vector perturbations, and further affects the

performance of particulate matter predictions.

Since the RMSE decreases with the increasing ensemble member size when 20 and 40

members are setting, and 48 hours of assimilation window length corresponds to a smaller

490   RMSE, the study compares the spatial distribution of ensemble standard deviations from

491   Severe-20m-48h and Severe-40m-48h to explain the relationship between ensemble member

492   size and simulation errors in analysis result. Figure 9 depicts contour maps of the spatial

493   distribution of temporal averaged standard deviations in the first guess and analysis of Severe-

494   40m-48h, Severe-20m-48h and their difference for $PM_{2.5}$ and $PM_{10-2.5}$ during severe haze event.

495   The first guess in Severe-40m-48h and Severe-20m-48h shows that the relatively high standard

496   deviations are generally observed in southern of BTH region, while those in the northern areas

497   are close to zero for both $PM_{2.5}$ and $PM_{10-2.5}$. High value centers are distributed in densely

498   populated areas and urban centers including Shijiazhuang, Xingtai, Tianjin and Tangshan city,

499   where the standard deviations have generally exceeded $30\mu g\ m^{-3}$. Combined with Figure S4, it

500   can be seen that the areas with large concentration standard deviations correspond well with the

501   spatial distribution of anthropogenic emission and the areas with large standard deviations of

502   emission sources. The standard deviations of concentrations of $PM_{2.5}$ and $PM_{10-2.5}$ have closely

503   relationship with the allocation and configuration of anthropogenic emission sources, because

504   disturbances are only added to emission sources for each ensemble member, without disturbing

505   the meteorological field in this haze event. The variation of difference in the third column

506   entirely comes from increasing ensemble member size. The positive difference between Severe-

507   40m-48h and Severe-20m-48h in first guess suggests that increasing ensemble member size

508   leads to greater differences among each ensemble for both $PM_{2.5}$ and $PM_{10-2.5}$ over BTH areas.

509   The high efficiency of 4D-LETKF is strongly influenced by sufficient information spread

510   among ensemble members, which integrate spreading observational information to produce

511   analysis from the first guess (Rubin et al., 2016). As a result, the increasing ensemble member

512 size improves divergence for each member and facilitates the state vectors about $PM_{2.5}$ and

513 $PM_{10-2.5}$ information spread in the first guess, which makes a better performance for Severe-

514 40m-48h rather than Severe-20m-48h in this severe haze event. The standard deviations of

515 $PM_{2.5}$ in analysis are generally lower than those in first guess. Due to the localization of 4D-

516 LETKF, that is the ground-based observation data only optimized for simulation grid within a

517 certain range, square-like areas of low standard deviations appear in the analysis of $PM_{2.5}$ both

518 for 40 and 20 ensemble members. Nearly all assimilated stations are located at the center of

519 low value square areas suggesting that 4D-LETKF tunes all $PM_{2.5}$ trajectories into a small range

520 with low standard deviation at each slot of analysis by the assimilation of ground-based

521 observations. For $PM_{10-2.5}$, there are no square-like areas of low standard deviations in the

522 analysis both for 40 and 20 ensemble members, indicating that the 4D-LETKF does not has an

523 obvious limitation for $PM_{10-2.5}$ trajectories, however, the decreased standard deviations effect

524 from the 4D-LETKF is still distinct for the particulate matter because PM_{10} consist of $PM_{2.5}$

525 and $PM_{10-2.5}$ in simulation. Enlarging ensemble member size is benefit to the improving of

526 standard deviations of $PM_{2.5}$ and $PM_{10-2.5}$ in analysis, while the improving magnitude of $PM_{2.5}$

527 is obviously smaller than $PM_{10-2.5}$. The assimilation results are not directly influenced by the

528 increased standard deviations in analysis. Such low increasement of standard deviations

529 (generally below 3 $\mu g\ m^{-3}$) is unlikely to induce uncertainty in the fitting and averaging process,

530 but facilitates divergence in initial conditions between forecasting members in the next

531 assimilation cycle. In addition, Figure S5 depict the spatial distribution of standard deviation

532 from Severe-60m-48h, Severe-20m-48h and their difference in the first guess and analysis field.

533 It can be seen that increasing the number of ensemble member generally also improves the

534   standard deviation in first guess and analysis over the BTH region both for $PM_{2.5}$ and $PM_{10-2.5}$.

535   Overall, the increasement of standard deviations generated by increasing ensemble member

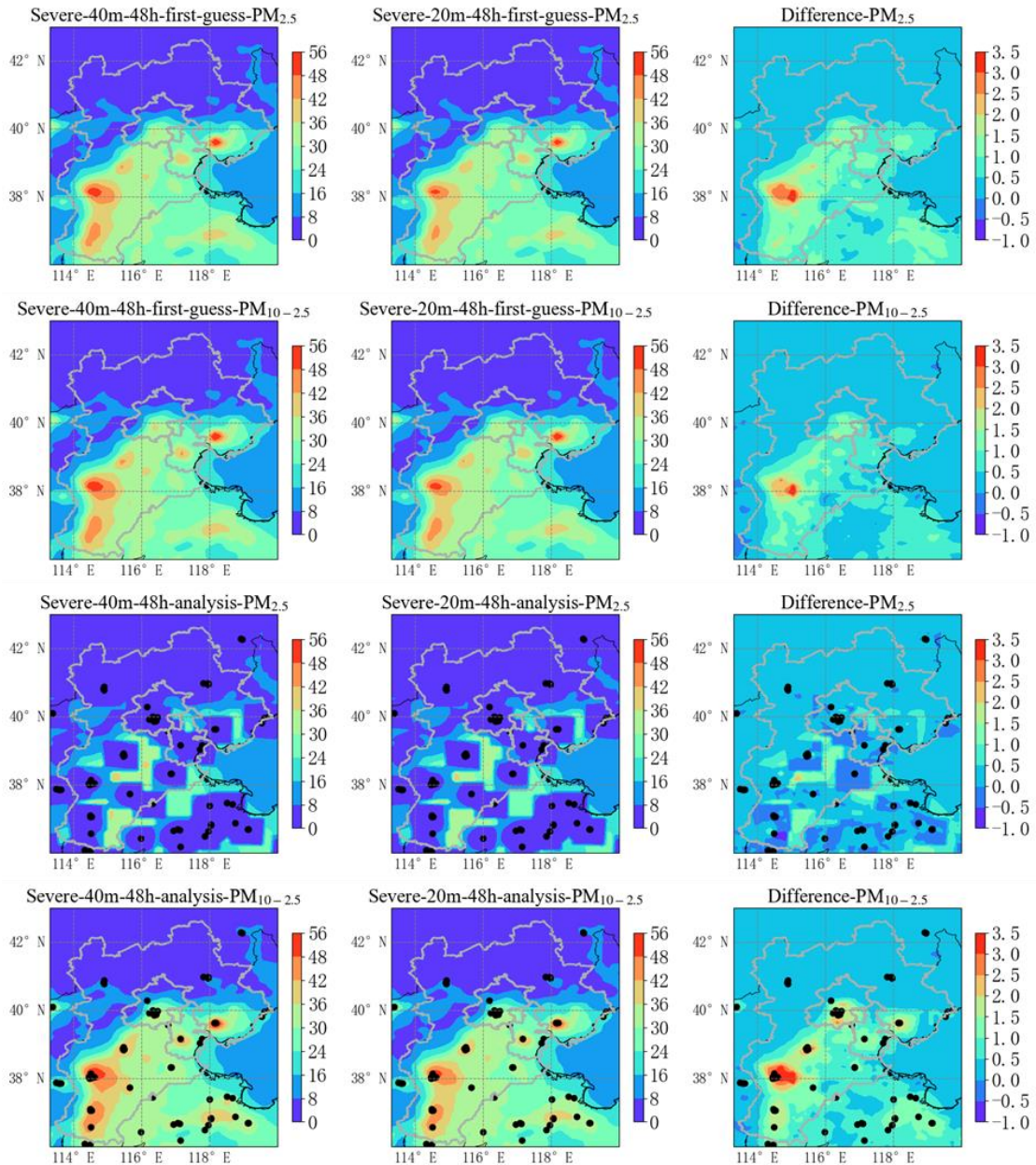
536   size directly improves the information spread of ensemble members in the first guess field and

537   the assimilation effect of 4D-LETKF, while the positive difference of standard deviation in

538   analysis favors the variances between each initial condition in the next assimilation window

539   during severe haze event.

540

541

542   Figure 9. Contour maps of spatial distributions of temporal averaged $PM_{2.5}$ and $PM_{10-2.5}$

543   standard deviations in the first guess (first and second row) and analysis (third and fourth

544   row) of Severe-40m-48h, Severe-20m-48h and their difference (Severe-40m-48h minus

545   Severe-20m-48h) within simulation period (units: $\mu g\ m^{-3}$). The black dots in analysis of $PM_{2.5}$

546   and $PM_{10-2.5}$ implies the location of assimilated stations. The grey border implies the BTH

547   region.

548   In the other sides, no matter for the first guess from Severe-40m-48h or Severe-20m-48h,

549   the high standard deviations are found near Shijiazhuang region in Figure 9, but Shijiazhuang

550   station (37.91N°, 114.35°E) still with larger RMSE and smaller AE in Figure 6. This seems

551   contrary to the opinion that increasing standard deviations in the first-guess field is beneficial

552   to raising the accuracy of pollutant simulations. Therefore, Shijiazhuang station and the stations

553   which with high values of AE (exceed 50) and difference of standard deviation in first guess

554   (exceed 1 $\mu g\ m^{-3}$) including Beijing, Tangshan, Handan, Baoding, Cangzhou, and Hengshui

555   regions are selected to explore the temporal distribution of standard deviations difference

556   between 40 and 20 ensemble members, so as to further advance our understanding about the

557   relationship between ensemble member size and simulation uncertainty in 4D-LETKF system.

558   The geographical locations of observations are shown in Figure S6 of the supplement. Figure

559   10 examines the temporal distribution of the standard deviation difference for $PM_{2.5}$ and PM_{10-}

560     $_{2.5}$ during the investigated period at Shijiazhuang station and results averaged from the selected

561   stations. From January 17 to January 18, the standard deviation difference in first guess at

562   Shijiazhuang station has increased drastically and exceeded up to 10 $\mu g\ m^{-3}$ for both $PM_{2.5}$ and

563     $PM_{10-2.5}$. This uneven temporal distribution results in a large standard deviation difference of

first guess in Figure 9. This huge divergency between ensemble member may attributed to the peak pollutant levels with AQI exceeds 300 at Shijiazhuang station occurs on the January 17 as shown in Figure 3. In highly polluted environments, 40 forecasting members with different perturbations in emission sources are more likely to differ the concentration of particulate matter in first guess fields. Excessive large spread of $PM_{2.5}$ and $PM_{10-2.5}$ for ensemble members may arise an over-high estimation about background error variance and obtain a poor Kalman gain. Moreover, it can be found that the standard deviation difference of $PM_{2.5}$ and $PM_{10-2.5}$ at Shijiazhuang station are obviously lower than the averaged from selected stations except the high dispersion time, suggesting the increasing number of ensemble members has limited impact on the spread between each ensemble member at Shijiazhuang during these dates. Too low standard deviations imply filter divergence near Shijiazhuang station, which may induce the underestimation of model spread, reduce the effect of observation information, and make system more certain of state estimate about particulate matter concentrations in first guess (Hunt et al., 2007). In addition, reducing uncertainty in the mixed anthropogenic emission inventory may be an important approach to avoid filter convergence near the Shijiazhuang region. Generally edited by empirical and statistical data such as anthropogenic emission factors and activity dataset, the anthropogenic emissions based on bottom-up method can hardly capture the real spatiotemporal distribution of anthropogenic emissions over China as frequently variations in energy consumption, even though the latest version. Among the southern of BTH region, the great positive innovations of particulate matter emissions in posterior estimation have been discovered in previous researches, implying that the update of underestimated emissions in this region may enlarge ensemble spread since a large quantity of

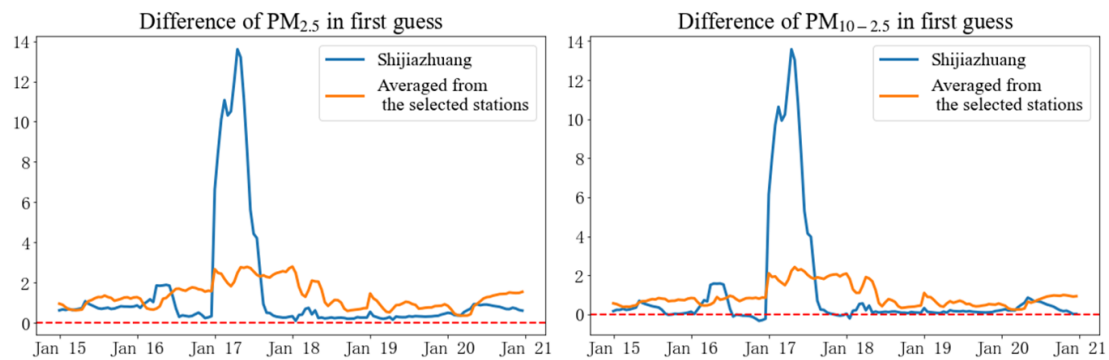
586 emissions corresponds to a higher degree of perturbation (Peng et al., 2017, Feng et al., 2023).

587 In a word, the perturbations added to emissions and meteorological fields needs to be executed

588 carefully in 4D-LETKF system to avoid too high or too low ensemble spread degree because

589 which determinate how analysis results weight toward observations information and first guess

590 fields (Dai et al., 2021).

591

592 Figure 10. Temporal distribution of standard deviation difference (Severe-40m-48h minus

593 Severe-20m-48h) in first guess for PM$_{2.5}$ and PM$_{10-2.5}$ at Shijiazhuang station and averaged

594 from the selected stations (units: μg m^{-3}). The red dash line is zero.

595 The results above suggest that the increasing ensemble member size enlarges ensemble

596 spread, benefits the information spread in the first guess and finally improves the simulative

597 skill in severe haze event. However, it has not been testified whether these influence rules are

598 also practical for a more common, and less polluted condition. Therefore, two assimilation

599 experiments in moderate haze event, Moderate-20m-48h and Moderate-40m-48h, are

600 performed to examine the applicable range. As shown in Figure S7, the moderate haze event

601 spans from 00:00 UTC 15 January 2019 to 00:00 UTC 21. This moderate event began on 15

602 January, with AQI increasing until 18 January, reaching a moderate level but not lasting for a

603 long time, and then decreased on 19 and 20 January. Most areas experienced mild or moderate

604 air pollution, with AQI generally below 200, the primary pollutant was particulate matter after

605 calculation. The simulations of moderate haze event utilize the same anthropogenic emission

606 inventory as used in severe haze event since two events both happen in January, thereby avoids

607 the additional influence introduce from emission source variation and the perturbations to

608 information spread and assimilation effect.

609   Figure S8 shows the simulated concentrations of $PM_{2.5}$ and PM_{10} against ground-based

610 observations during moderate air pollution event. The RMSEs of $PM_{2.5}$ in Moderate-FR-48h,

611 Moderate-20m-48h and Moderate-40m-48h are 40.40, 24.12 and 18.52$\mu g\ m^{-3}$, respectively, and

612 the RMSEs of PM_{10} are 73.47, 67.81 and 57.04 $\mu g\ m^{-3}$ respectively. The concentrations of $PM_{2.5}$

613 and PM_{10} in assimilation experiments are more in agreement with observations, suggesting the

614 validation of 4D-LETKF initial condition adjustment in moderate haze event. The phenomena

615 that the simulation error of $PM_{2.5}$ and PM_{10} decrease with increasing ensemble member size are

616 same with those characteristics have shown in severe haze event before.

617   Similar to Figure 9, Figure S9 presents the spatial distributions of standard deviations about

618 $PM_{2.5}$ and PM_{10} in the first guess of Moderate-40m-48h, Moderate-20m-48h and their

619 difference. The relatively smaller magnitude of standard deviation difference in first guess may

620 relate to relatively low $PM_{2.5}$ and PM_{10} concentrations in moderate haze event. Positive

621 difference in first guess and analysis for particulate matter implies the Moderate-40m-48h

622 obtains a higher diversity of ensemble members than Moderate-20m-48h, and which are also

623 similar with those happen in the severe haze event.

624 4. Summary

625   The numerical reproduction of severe haze events with AQI larger than 200 has been a

626 challenging problem in the field of atmospheric pollution for a long time. In this research, a

WRF-Chem/4D-LETKF coupled data assimilation system has been successfully developed by ensemble member with perturbed anthropogenic emissions to improve the simulative skill of particulate matter in severe haze event during the winter of 2020. The research validated the effectiveness of 4D-LETKF data assimilation, discussed the optimal parameter combination of ensemble member size and length of assimilation window for 4D-LETKF assimilation system, summarized and explained the influence rules from parametric selection to the 4D-LETKF assimilation effect during severe and moderate haze event.

It is concluded that the Severe-40m-48h experiment shows the best performance in the simulations of $PM_{2.5}$ and PM_{10} after comparing the statistical errors and computing resource consumption across multiple sensitivity analyses, with the RMSEs of 31.19 and 76.83μg m^{-3} for $PM_{2.5}$ and PM_{10} in severe haze event. Severe-40m-48h optimizes the underestimation of particulate matter concentrations in Severe-FR-48h, and remarkably improves the simulation accuracy in the entire BTH region and neighboring provinces. For example, the RMSEs of $PM_{2.5}$ in Baoding, Hengshui and Cangzhou decrease to 29.85, 18.98 and 19.06 μg m^{-3} respectively, from 80.55, 55.22 and 76.32 μg m^{-3} in Severe-FR-48h. Severe-40m-48h is also capable of retrieving the peaks and valleys of particulate matter concentration over the investigated period. To examine the dependence of the assimilation effect of 4D-LETKF, nine panels of sensitivity tests were conducted according to ensemble member size and length of assimilation window. The findings suggest that the simulation accuracy of $PM_{2.5}$ and PM_{10} can be strongly improved by the increasing ensemble member size from 20 to 40. A relative longer assimilation window length such as 48 or 72 hours combine with 40 ensemble member size is advised in 4D-LETKF assimilation system. In view of performance of ensemble member,

649   increasing ensemble member size improves ensemble spread among each forecasting member,

650   facilitates the spread of state vectors about $PM_{2.5}$ and PM_{10} information in the first guess, favors

651   the variances between each initial condition in the next assimilation window and leads to better

652   performance in simulation of severe haze event. A similar conclusion can also be draw from the

653   moderate haze event, suggesting that this influence rule is applicable in both severe and

654   moderate haze conditions.

655   There are still some deficiencies in this research. Although we have performed data quality

656   control in this study, we did not use approaches such as super-observations to improving the

657   correspondence between grid points and observations (Jin et al., 2022, Miyazaki et al., 2012a),

658   which may increase the representational error and result in the possibility of two stations with

659   different concentrations interpolating in the same grid. Improving the spatial resolution of

660   forward model or introducing super observations may mitigate this problem (Miyazaki et al.,

661   2012b, Feng et al., 2020b). Furthermore, the concentration of state variables about particulate

662   matters in initial conditions are optimized in this study, but there still remain large uncertainties

663   in anthropogenic emission data, which is an important chemical boundary input for pollutant

664   simulations. These uncertainties sources may play a significant role in the over- or

665   underestimation of pollutant ensemble modeling. The anthropogenic emissions inversion based

666   on Ensemble Kalman filter and their variants is recognized as an effective approach for

667   reducing uncertainty in anthropogenic emission sources (Peng et al., 2018, Feng et al., 2020a,

668   Chen et al., 2019b). The jointly adjust initial conditions and emissions source with 4D-LETKF

669   is the focus of future work to further improving the forecast skills about air pollutants during

670   heavy pollution events.

671

CRediT authorship contribution statement

Jianyu Lin: Conceptualization, Formal analysis, Visualization. **Tie Dai**: Investigation, Methodology, Resources, Supervision. **LiFang Sheng**: Funding acquisition, Project administration. **Weihang Zhang**: Writing - original draft, Data curation. **Shangfei Hai**: Writing - review and editing. **Yawen Kong**: Validation.

Declaration of competing interest

The authors declare that they have no conflict of interest.

Acknowledgements

We are grateful to the relevant researchers who make contributions to LETKF method (Miyoshi, 2024).

Financial support. This work is supported by the National Natural Science Foundation of China (No. 42275191), National Natural Science Funds of China (42375190).

Code and data availability

The WRF-Chem Version 3.9.1, LETKF source code and data in this research are available in https://zenodo.org/records/14010521 (Lin, 2024). The observation data are available in https://zenodo.org/records/14835237 (Lin, 2025).

692

693   Reference

695   Anderson, J. L.: An adaptive covariance inflation error correction algorithm for ensemble filters,

696   Tellus, 59A, 210–224, https://doi.org/10.1111/j.1600-0870.2006.00216.x, 2007.

698   Anderson, J. L. and Anderson, S. L.: A Monte Carlo implementation of the nonlinear filtering

699   problem to produce ensemble assimilations and forecasts, Mon. Weather Rev., 127, 2741–2758,

700   1999.

702   Bao, J., Yang, X., Zhao, Z., Wang, Z., Yu, C., Li, X.: The spatial-temporal characteristics of air

703   pollution in China from 2001-2014, Int. J. Environ. Res. Publ. Health, 12 (12), 15875e15887,

704   https://doi.org/10.3390/ijerph121215029, 2015.

706   Barbu, A. L., Segers, A. J., Schaap, M., Heemink, A. W., and Builtjes, P. J. H.: A multi-

707   component data assimilation experiment directed to sulphur dioxide and sulphate over Europe,

708   Atmospheric Environment, 43(9), 1622–1631, https://doi.org/10.1016/j.atmosenv.2008.12.005,

709   2009.

711   Benedetti, A., Morcrette, J., Boucher, O., Dethof, A., Engelen, R. J., Fisher, M., Flentje, H.,

712   Huneeus, N., Jones, L., Kaiser, J. W., Kinne, S., Mangold, A., Razinger, M., Simmons, A. J.,

713   and Suttie, M.: Aerosol analysis and forecast in the European Centre for Medium-Range

714   Weather Forecasts Integrated Forecast System: 2. Data assimilation, J. Geophys. Res., 114,

D13205, https://doi.org/10.1029/2008JD011115, 2009.

Cheng, Y., Dai, T., Goto, D., Schutgens, N. A. J., Shi, G., and Nakajima, T.: Investigating the assimilation of CALIPSO global aerosol vertical observations using a four-dimensional ensemble Kalman filter, Atmos. Chem. Phys, 19, 13445–13467, https://doi.org/10.5194/acp-19-13445-2019, 2019.

Chen, D., Liu, Z., Ban, J., and Chen, M.: The 2015 and 2016 wintertime air pollution in China: SO_2 emission changes derived from a WRF-Chem/EnKF coupled data assimilation system, Atmos. Chem. Phys, 19, 8619–8650, https://doi.org/10.5194/acp-19-8619-2019, 2019a.

Chen, D., Liu, Z., Ban, J., Zhao, P., and Chen, M.: Retrospective analysis of 2015–2017 wintertime $PM_{2.5}$ in China: response to emission regulations and the role of meteorology, Atmos. Chem. Phys., 19, 7409–7427, https://doi.org/10.5194/acp-19-7409-2019, 2019b.

CNEMC: China National Environmental Monitoring Center, http://106.37.208.233:20035/ (last access: 31 December 2024), 2024.

Dai, T., Cheng, Y., Goto, D., Li, Y., Tang, X., Shi, G., and Nakajima, T.: Revealing the sulfur dioxide emission reductions in China by assimilating surface observations in WRF-Chem, Atmos. Chem. Phys., 21, 4357–4379, https://doi.org/10.5194/acp-21-4357-2021, 2021.

737   Dai, T., Cheng, Y., Suzuki, K., Goto, D., Kikuchi, M., Schutgens, N. A. J., Yoshida, M., Zhang,

738   P., Husi, L., Shi, G., and Nakajima, T.: Hourly Aerosol Assimilation of Himawari-8 AOT Using

739   the Four-Dimensional Local Ensemble Transform Kalman Filter, J. Adv. Model. Earth Syst.,

740   11, 680–711, https://doi.org/10.1029/2018MS001475, 2019.

741

742   Elbern, H., Strunk, A., Schmidt, H., and Talagrand, O.: Emission rate and chemical state

743   estimation by 4-dimensional variational inversion, Atmospheric Chemistry and Physics, 7(14),

744   3749–3769, https://doi.org/10.5194/acp-7-3749-2007, 2007.

745

746   Evensen, G.: Sequential data assimilation with a nonlinear quasi-geostrophic model using

747   Monte Carlo methods to forecast error statistics, Journal of Geophysical Research: Oceans, 99,

748   10143–10162, https://doi.org/10.1029/94JC00572, 1994.

749

750   Evensen, G.: The Ensemble Kalman Filter: theoretical formulation and practical

751   implementation, Ocean Dynam, 53, 343–367, https://doi.org/10.1007/s10236-003-0036-9,

752   2003.

753

754   Feng, S., Jiang, F., Jiang, Z., Wang, H., Cai, Z., and Zhang, L.: Impact of 3DVAR assimilation

755   of surface $PM_{2.5}$ observations on $PM_{2.5}$ forecasts over China during wintertime, Atmospheric

756   Environment, 187, 34-49, https://doi.org/10.1016/j.atmosenv.2018.05.049, 2018.

757

758   Feng, S., Jiang, F., Wang, H., Wang, H., Ju, W., Shen, Y., Zheng, Y., Wu, Z., and Ding,

759   A.: NOx emission changes over China during the COVID-19 epidemic inferred from surface

760   NO_2 observations, Geophysical   Research   Letters, 47,   e2020GL090080,

761   https://doi.org/10.1029/2020GL090080, 2020a.

762

763   Feng, S., Jiang, F., Wu, Z., Wang, H., Ju, W., and Wang, H.: CO emissions inferred from surface

764   CO observations over China in December 2013 and 2017, Journal of Geophysical Research:

765   Atmospheres, 125, https://doi.org/10.1029/2019JD031808, 2020b.

766

767   Feng, S., Jiang, F., Wu, Z., Wang, H., He, W., Shen, Y., Zhang, L., Zheng, Y., Lou, C., Jiang, Z.,

768   and Ju, W.: A Regional multi-Air Pollutant Assimilation System (RAPAS v1.0) for emission

769   estimates: system development and application, Geosci. Model Dev., 16, 5949–5977,

770   https://doi.org/10.5194/gmd-16-5949-2023, 2023.

771

772   Gao, M., Saide, P. E., Xin, J., Wang, Y., Liu, Z., Wang, Y., Wang, Z., Pagowski, M., Guttikunda,

773   S. K., and Carmichael, G. R.: Estimates of Health Impacts and Radiative Forcing in Winter

774   Haze in Eastern China through Constraints of Surface $PM_{2.5}$ Predictions, Environmental

775   Science & Technology, 51(4), 2178-2185, https://doi.org/10.1021/acs.est.6b03745, 2017.

776

777   Grell, G. A.: Prognostic Evaluation of Assumptions Used by Cumulus Parameterizations, Mon.

778   Wea.   Rev., 121,   764–787,   https://doi.org/10.1175/1520-

779   0493(1993)121<0764:PEOAUB>2.0.CO;2, 1993.

780

781   Grell, G., Peckham, S. E., Schmitz, R., McKeen, S. A., Frost, G., Skamarock, W. C., and Eder,

782   B.: Fully coupled "online" chemistry within the WRF model, Atmos. Environ, 39, 6957–6975,

783   https://doi.org/10.1016/j.atmosenv.2005.04.027, 2005.

784   Guenther, A., Hewitt, C. N., Erickson, D., Fall, R., Geron, C., Graedel, T., Harley, P., Klinger,

785   L., Lerdau, M., Mckay, W. A., Pierce, T., Scholes, B., Steinbrecher, R., Tallamraju, R., Taylor,

786   J., and Zimmerman, P.: A global model of natural volatile organic compound emissions, Journal

787   of Geophysical Research: Atmospheres, 100, 8873–8892, https://doi.org/10.1029/94JD02950,

788   1995.

789

790   Hong, S., Noh, Y., and Dudhia, J.: A new vertical diffusion package with an explicit treatment

791   of entrainment processes, Mon. Wea. Rev, 134, 2318–2341,

792   https://doi.org/10.1175/MWR3199.1, 2006.

793

794   Huang, X.Y., Xiao, Q., Barker, D. M., Zhang, X., Michalakes, J., Huang, W., Henderson, T.,

795   Bray, J., Chen, Y., Ma, Z., Dudhia, J., Guo, Y., Zhang, X., Won, D., Lin, H., and Kuo, Y.: Four-

796   dimensional variational data assimilation for WRF: Formulation and preliminary results, Mon.

797   Weather Rev, 137, 299–314, https://doi.org/10.1175/2008MWR2577.1, 2009.

798

799   Hunt, B. R., Kalnay, E., Kostelich, E. J., Ott, E., Patil, D. J., Sauer, T., Szunyogh, I., Yorke, J.

800   A., and Zimin, A. V.: Four-dimensional ensemble Kalman filtering, Tellus A: Dynamic

801   Meteorology and Oceanography, 56(4), 273–277, https://doi.org/10.3402/tellusa.v56i4.14424,

802   2004.

803

Hunt, B. R., Kostelich, E. J., and Szunyogh, I.: Efficient data assimilation for spatiotemporal chaos: a local ensemble transform Kalman filter, Physica D: Nonlinear Phenomena, 230, 112–126, https://doi.org/10.1016/j.physd.2006.11.008, 2007.

Iacono, M. J., Delamere, J. S., Mlawer, E. J., Shephard, M. W., Clough, S. A., and Collins, W. D.: Radiative forcing by long–lived greenhouse gases: Calculations with the AER radiative transfer models, J. Geophys. Res., 113, D13103, https://doi.org/10.1029/2008JD009944, 2008.

Jin, J., Pang, M., Segers, A., Han, W., Fang, L., Li, B., Feng, H., Lin, H. X., and Liao, H.: Inverse modeling of the 2021 spring super dust storms in East Asia, Atmos, Chem. Phys., 22, 6393–6410, https://doi.org/10.5194/acp-22-6393-2022, 2022.

Kong, L., Tang, X., Zhu, J., Wang, Z., Sun, Y., Fu, P., Gao, M., Wu, H., Lu, M., Wu, Q., Huang, S., Sui, W., Li, J., Pan, X., Wu, L., Akimoto, H., and Carmichael, G. R.: Unbalanced emission reductions of different species and sectors in China during COVID-19 lockdown derived by multi-species surface observation assimilation, Atmos. Chem. Phys., 23, 6217–6240, https://doi.org/10.5194/acp-23-6217-2023, 2023.

Kong, Y., Sheng, L., Li, Y., Zhang, W., Zhou, Y., Wang, W., and Zhao, Y.: Improving $PM_{2.5}$ forecast during haze episodes over China based on a coupled 4D-LETKF and WRF-Chem system, Atmospheric Research, 249, 105366, https://doi.org/10.1016/j.atmosres.2020.105366, 2021.

827   Lin J: LETKF-WRF-Chem, Zenodo [code and data set], https://zenodo.org/records/14010521,

828   2024.

829

830   Lin J: Observation, Zenodo [Data set], https://doi.org/10.5281/zenodo.14835237, 2025.

831

832   Lorenc, A.C.: Analysis methods for numerical weather prediction, Q. J. R. Meteorol. Soc., 112,

833   1177–1194, https://doi.org/10.1002/qj.49711247414, 1986.

834

835   Luo, X., Tang, X., Wang, H., Kong, L., Wu, H., Wang, W., SONG, Y., Luo, H., Wang, Y., Zhu,

836   J., and Wang, Z.: Investigating the Changes in Air Pollutant Emissions over the Beijing-Tianjin-

837   Hebei Region in February from 2014 to 2019 through an Inverse Emission Method, Adv. Atmos.

838   Sci. 40, 601–618, https://doi.org/10.1007/s00376-022-2039-9, 2023.

839

840   Ma, C., Wang, T., Jiang, Z., Wu, H., Zhao, M., Zhuang, B., Li, S., Xie, M., Li, M., Liu, J., and

841   Wu, R.: Importance of bias correction in data assimilation of multiple observations over eastern

842   China   using   WRF-Chem/DART, Journal   of   Geophysical   Research: Atmospheres, 125,

843   e2019JD031465, https://doi.org/10.1029/2019JD031465, 2020.

844

845   MEIC: Multi-resolution Emission Inventory for China, http://www.meicmodel.org/ (last access:

846   31 December 2024), 2024.

847

848   Miyoshi, T., Yamane, S., and Enomoto, T.: Localizing the Error Covariance by Physical

Distances within a Local Ensemble Transform Kalman Filter (LETKF), Scient. Online Lett. Atmos., 3, 89–92, https://doi.org/10.2151/sola.2007-023, 2007.

Morrison, H., Thompson, G., and Tatarskii, V.: Impact of Cloud Microphysics on the Development of Trailing Stratiform Precipitation in a Simulated Squall Line: Comparison of One– and Two–Moment Schemes, Mon. Wea. Rev., **137**, 991–1007, https://doi.org/10.1175/2008MWR2556.1, 2009.

Miyazaki, K., Eskes, H. J., and Sudo, K.: Global NOx emission estimates derived from an assimilation of OMI tropospheric NO_2 columns, Atmos. Chem. Phys., 12, 2263–2288, https://doi.org/10.5194/acp-12-2263-2012, 2012a.

Miyazaki, K., Eskes, H. J., Sudo, K., Takigawa, M., Weele, M. van., and Boersma, K. E.: Simultaneous assimilation of satellite NO_2, O_3, CO, and HNO_3 data for the analysis of tropospheric chemical composition and emissions, Atmos. Chem. Phys., 12, 9545–9579, https://doi.org/10.5194/acp-12-9545-2012, 2012b.

Miyoshi, T.: LETKF source codes, GitHub, available at: https://github.com/takemasa-miyoshi/letkf, last access: 1 January 2024.

NCDC: National Climatic Data Center, https://wwww.ncei.noaa.gov/ (last access: 31 December 2024), 2024.

871

872   Ott, E., Hunt, B. R., Szunyogh, I., Zimin, A. V., Kostelich, E. J., Corazza, M., Kalnay, E., Patil,

873   D. J., and Yorke, J. A.: A local ensemble Kalman filter for atmospheric data assimilation. Tellus

874   A, 56, 415–428, https://doi.org/10.1111/j.1600-0870.2004.00076.x, 2004.

875

876   Pagowski, M., and Grell, G. A.: Experiments with the assimilation of fine aerosols using an

877   ensemble Kalman filter, Journal of Geophysical Research, 117, D21302,

878   https://doi.org/10.1029/2012JD018333, 2012.

879

880   Parrish, D. F. and Derber, J. C.: The National Meteorological Center's Spectral Statistical-

881   Interpolation Analysis System, Mon. Weather Rev., 120, 1747–1763,

882   https://doi.org/10.1175/1520-0493(1992)120<1747:TNMCSS>2.0.CO;2, 1992.

883

884   Peng, Z., Liu, Z., Chen, D., and Ban, J.: Improving $PM_{2.5}$ forecast over China by the joint

885   adjustment of initial conditions and source emissions with an ensemble Kalman filter, Atmos.

886   Chem. Phys., 17, 4837–4855, https://doi.org/10.5194/acp-17-4837-2017, 2017.

887

888   Peng, Z., Lei, L., Liu, Z., Sun, J., Ding, A., Ban, J., Chen, D., Kou, X., and Chu, K.: The impact

889   of multi-species surface chemical observation assimilation on air quality forecasts in China,

890   Atmos. Chem. Phys., 18, 17387–17404, https://doi.org/10.5194/acp-18-17387-2018, 2018.

891

892   Rubin, J. I., Reid, J. S., Hansen, J. A., Anderson, J. L., Collins, N., Hoar, T. J., Hogan, T., Lynch,

893   P., McLay, J., Reynolds, C. A., Sessions, W. R., Westphal, D. L., and Zhang, J.: Development

of the Ensemble Navy Aerosol Analysis Prediction System (ENAAPS) and its application of the Data Assimilation Research Testbed (DART) in support of aerosol forecasting, Atmos. Chem. Phys., 16, 3927–3951, https://doi.org/10.5194/acp-16-3927-2016, 2016.

Schutgens, N. A. J., Miyoshi, T., Takemura, T., and Nakajima, T.: Sensitivity tests for an ensemble Kalman filter for aerosol assimilation, Atmos. Chem. Phys., 10, 6583–6600, https://doi.org/10.5194/acp-10-6583-2010, 2010.

Sekiyama, T. T., Tanaka, T. Y., Shimizu, A., and Miyoshi, T.: Data assimilation of CALIPSO aerosol observations, Atmos. Chem. Phys., 10, 39–49, doi:10.5194/acp-10-39-2010, 2010.

Sun, W., Liu, Z., Chen, D., Zhao, P., and Chen, M.: Development and application of the WRFDA-Chem three-dimensional variational (3DVAR) system: aiming to improve air quality forecasting and diagnose model deficiencies, Atmos. Chem. Phys., 20, 9311–9329, https://doi.org/10.5194/acp-20-9311-2020, 2020.

Schwartz, C. S., Liu, Z., Lin, H. C., and McKeen, S. A.: Simultaneous three-dimensional variational assimilation of surface fine particulate matter and MODIS aerosol optical depth, J. Geophys. Res., 117, D13202, https://doi.org/10.1029/2011JD017383, 2012.

Stockwell, W. R., Kirchner, F., Kuhn, M., and Seefeld, S.: A new mechanism for regional atmospheric chemistry modeling, J. Geophys. Res., 102, 25847–25879,

916   https://doi.org/10.1029/97JD00849, 1997.

918   Tewari, M., Chen, F., Wang, W., Dudhia, J., LeMone, M. A., Mitchell, K., Ek, M., Gayno, G.,

919   Wegiel, J., and Cuenca, R. H.: Implementation and verification of the unified NOAH land

920   surface model in the WRF model, 20th conference on weather analysis and forecasting/16th

921   conference on numerical weather prediction, pp. 11–15, 2004.

923   van Donkelaar, A., Martin, R.V., Brauer, M., Hsu, N. C., Kahn, R. A., Levy, R. C., Lyapustin,

924   A., Sayer, A. M., and Winker, D. M.: Global estimates of fine particulate matter using a

925   combined geophysical-statistical method with information from satellites, models, and

926   monitors, Environ. Sci. Technol, 50, 3762–3772, https://doi.org/10.1021/acs.est.5b05833, 2016.

928   Wang, Z., Li, J., Wang, Z., Yang, W., Tang, X., Ge, B., Yan, P., Zhu, L., Chen, X., and Chen, H.:

929   Modeling study of regional severe hazes over mid-eastern China in January 2013 and its

930   implications on pollution prevention and control, Sci. China-Earth Sci., 57, 3–13,

931   https://doi.org/10.1007/s11430-013-4793-0, 2014.

933   Whitaker, J. S. and Hamill, T. M.: Ensemble data assimilation without perturbed observations,

934   Mon. Weather Rev., 130, 1913–1924, https://doi.org/10.1175/1520-

935   0493(2002)130<1913:EDAWPO>2.0.CO;2, 2002.

937   Yan, S., Cao, H., Chen, Y., Wu, C., and Fan, H.: Spatial and temporal characteristics of air

quality and air pollutants in 2013 in Beijing, Environ Sci Pollut Res., 23, 13996–14007, https://doi.org/10.1007/s11356-016-6518-3, 2016.

Yumimoto, K., and Takemura, T.: Direct radiative effect of aerosols estimated using ensemble-based data assimilation in a global aerosol climate model, Geophysical Research Letters, 38, L21802, https://doi.org/10.1029/2011GL049258, 2011.

Zheng, B., Cheng, J., Geng, G., Wang, X., Li, M., Shi, Q., Qi, J., Lei, Y., Zhang, Q., He, K.: Mapping anthropogenic emissions in China at 1 km spatial resolution and its application in air quality modeling, Science Bulletin, 66(6), 612-620, https://doi.org/10.1016/j.scib.2020.12.008, 2021.

Zhang, L., Li, Q., Wang, T., Ahmadov, R., Zhang, Q., Li, M., and Lv, M.: Combined impacts of nitrous acid and nitryl chloride on lower-tropospheric ozone: new module development in WRF-Chem and application to China, Atmos. Chem. Phys., 17, 9733–9750, https://doi.org/10.5194/acp-17-9733-2017, 2017.

Zhang, Q., Streets, D. G., Carmichael, G. R., He, K. B., Huo, H., Kannari, A., Klimont, Z., Park, I. S., Reddy, S., Fu, J. S., Chen, D., Duan, L., Lei, Y., Wang, L. T., and Yao, Z. L.: Asian emissions in 2006 for the NASA INTEX-B mission, Atmos. Chem. Phys., 9, 5131–5153, https://doi.org/10.5194/acp-9-5131-2009, 2009.

960 Zhao, Y., Greybush, S. J., Wilson, R. J., Hoffman, R. N., and Kalnay, E.: Impact of assimilation

961 window length on diurnal features in a Mars atmospheric analysis, Tellus A, 67, 26042,

962 https://doi.org/10.3402/tellusa.v67.26042, 2015.

963

964 Zhang, Q., Ma, Q., Zhao, B., Liu, X., Wang, Y., Jia, B., and Zhang, X.: Winter haze over North

965 China Plain from 2009 to 2016: Influence of emission and meteorology, Environmental

966 Pollution, 242(B), 1308-1318, https://doi.org/10.1016/j.envpol.2018.08.019, 2018.

967

968 Zhan, D., Kwan, M., Zhang, W., Yu, X., Meng, B., and Liu, Q : The driving factors of air quality

969 index in China, Journal of Cleaner Production, 197, 1, 1342-1351,

970 https://doi.org/10.1016/j.jclepro.2018.06.108, 2018.