

1 **Improving the Sectional MOSAIC Aerosols of WRF-Chem with the revised**
2 **Gridpoint Statistical Interpolation System and multi-wavelength aerosol optical**
3 **measurements: DAO-K experiment 2019 at Kashi, near the Taklamakan Desert,**
4 **northwestern China**

5
6 Wenyuan Chang¹, Ying Zhang², Zhengqiang Li², Jie Chen³, Kaitao Li²
7
8

9 ¹ State Key Laboratory of Atmospheric Boundary Layer Physics and Atmospheric Chemistry
10 (LAPC), Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing 100029,
11 China.
12

13 ² State Environment Protection Key Laboratory of Satellite Remote Sensing, Aerospace
14 Information Research Institute, Chinese Academy of Sciences, Beijing 100101, China
15

16 ³ National Meteorological Information Center, China Meteorological Administration, Beijing
17 100081, China
18
19
20
21
22

23 Corresponding authors:

24 Wenyuan Chang (changwy@mail.iap.ac.cn)

25 Zhengqiang Li (lizq@radi.ac.cn)
26
27

28 **Abstract**

29 The Gridpoint Statistical Interpolation data assimilation (DA) system was developed for the
30 four-size bin sectional Model for Simulating Aerosol Interactions and Chemistry (MOSAIC)
31 aerosol mechanism in the Weather Research and Forecasting-Chemistry (WRF-Chem) model.
32 The forward and tangent linear operators for the aerosol optical depth (AOD) analysis were
33 derived from WRF-Chem aerosol optical code. We applied three-dimensional variational DA
34 to assimilate the multi-wavelength AOD, ambient aerosol scattering coefficient, and aerosol
35 absorption coefficient, measured by the sun-sky photometer, nephelometer, and aethalometer,
36 respectively. These were undertaken during a dust observation field campaign at Kashi in
37 northwestern China in April 2019. The results showed that the DA analyses decreased the low
38 biases in the model aerosols; however, it had some deficiencies. Assimilating the surface
39 particle concentration increased the coarse particles in the dust episodes, but AOD, and the
40 coefficients for aerosol scattering and absorption, were still lower than observed values.
41 Assimilating aerosol scattering coefficient separately from AOD improved the two optical
42 quantities. However, it caused an overestimation of the particle concentrations at the surface.
43 Assimilating the aerosol absorption coefficient yielded the highest positive bias in the surface
44 particle concentration, aerosol scattering coefficient, and AOD. The positive biases in the DA
45 analysis were caused by the forward operator underestimating aerosol mass scattering and
46 absorption efficiency. As a compensation, the DA system increased particle concentrations
47 excessively so as to fit the observed optical values. The best overall improvements were
48 obtained from the simultaneous assimilation of the surface particle concentration and AOD.
49 The assimilation did not substantially change the aerosol chemical fractions. After DA, the
50 clear-sky aerosol radiative forcing at Kashi was -10.4 Wm^{-2} at the top of the atmosphere,
51 which was 55% higher than the background radiative forcing value.

52 **1. Introduction**

53 Data assimilation (DA) blends the information from observations with *a priori* background
54 fields from deterministic models to obtain an optimal analysis (Wang et al., 2001; Bannister,
55 2017). With lagged emission inventories and unsatisfactory model chemistry mechanisms,
56 there are notable discrepancies between model aerosols and observed levels (He et al., 2017;
57 Chen L. et al., 2019). The DA technology incorporates aerosol measurements into the models
58 to optimize emissions (Peng et al., 2017; Ma et al., 2019), and cyclically updates the
59 background fields in forecasts. This effectively improves the air quality forecasts in China
60 (Bao et al., 2019; Cheng et al., 2019; Feng et al., 2018; Hong et al., 2020; Liu et al., 2011;
61 Pang et al., 2018; Peng et al., 2018; Xia et al., 2019a, 2019b).

62
63 Variational DA minimizes the distant scalar function measuring the misfit between model
64 states and a set of observations in each assimilation window. An effective variational DA
65 requires an appropriate tangent linear and adjoint operators, which describes the gradient or
66 sensitivity of the observed parameter to the control variable (Wang et al., 2001; Bannister
67 2017). The operator is highly dependent on the types of assimilated observations and the
68 selection of control variables; it is also sometimes dependent on the aerosol mechanism. For
69 PM_{2.5} (particulate matter with dynamic radius less than 2.5 μm) DA, the tangent linear
70 operator is the ratio of the PM_{2.5} concentration to composition of each aerosol (Pagowski et
71 al., 2010). For the aerosol optical depth (AOD) DA, the operator is generated through Mie
72 theory (Liu et al., 2011; Saide et al., 2013). With the development of aerosol mechanisms and
73 the growing body of novel aerosol observations from ground-based networks and satellites,
74 appropriate tangent linear and adjoint operators are in demand.

75
76 The community gridpoint statistical interpolation (GSI) system (Wu et al., 2002; Purser et al.,
77 2003a, 2003b) is often used to modify regional aerosol simulations with three-dimensional
78 variational (3D-Var) DA. The official GSI (version 3.7 in this study) can incorporate
79 observations of surface particulate matter concentration and AOD to constrain the aerosols
80 simulated within the aerosol mechanism of Goddard Chemistry Aerosol Radiation and
81 Transport (GOCART, Liu et al., 2011; Pagowski et al., 2014). The tangent linear operator and
82 adjoint operator for AOD were determined using the Community Radiative Transfer Model
83 (CRTM). This GSI version incorporating the Moderate Resolution Imaging
84 Spectroradiometer (MODIS) AOD in East Asia (Liu et al., 2011) revealed the simultaneous
85 DA effects of PM_{2.5} and AOD in the continental United States (Schwartz et al., 2012). This
86 GSI was used to identify DA effects that weakened during running of the succeeding model
87 as the model error grew (Jiang et al., 2013), and assessed the radiative forcing of the aerosols
88 released by wildfires (Chen et al., 2014). This version of GSI was also utilized to improve air
89 quality forecasts in China by assimilating a variety of satellite AOD data retrieved from: the
90 Geostationary Ocean Color Imager (Pang et al., 2018); Visible Infrared Imaging Radiometer
91 Suite (Pang et al., 2018); Advanced Himawari-8 Imager (Xia et al., 2019a); and the Fengyun-
92 3A/medium-resolution spectral imager (Bao et al., 2019; Xia et al., 2019b).

93
94 Despite its capabilities, the GOCART mechanism is unable to simulate nitrate and secondary
95 organic aerosols (SOA), and the GOCART aerosol size distribution uses a bulk assumption

96 for radiative transfer calculation. Strictly speaking, the lack of aerosol components violates
97 the unbiased requirements for the model states in the DA system. Lack of size-segregated
98 aerosols may introduce a bias in the calculation of optical aerosols. The official GSI can
99 assimilate the surface particle concentration from the aerosol mechanism apart from
100 GOCART, but its AOD DA is tightly bound with the GOCART aerosols. If one wished to use
101 GSI to assimilate AOD for the other aerosol mechanisms, a compromise solution was to
102 either integrate the map of the speciated aerosols of other mechanisms into that of the
103 GOCART aerosols or use a simplified formula to convert aerosol chemical mass
104 concentrations to AOD. For example, Tang et al. (2017) used the official GSI to assimilate
105 MODIS AOD with the aerosols from the Community Multi-scale Air Quality Model
106 (CMAQ). They incorporated the map of the 54 aerosol components of CMAQ into the five
107 CRTM aerosols and repartitioned the mass increments of each CMAQ aerosol according to
108 the ratio of aerosol chemical components in the background field. This repartitioning is called
109 the “ratio approach.” Cheng et al. (2019) assimilated the lidar extinction coefficient profiles
110 measured in Beijing to modify the Weather Research and Forecasting-Chemistry (WRF-
111 Chem) Model for Simulating Aerosol Interactions and Chemistry (MOSAIC) aerosols. They
112 used the ratio approach to map eight MOSAIC aerosols based on five GOCART aerosols.
113 This mapping strategy is readily implemented but introduces inconsistent size-segregated
114 aerosol information (e.g., hygroscopicity and extinction efficiency) between the aerosol
115 model and the DA system. Kumar et al. (2019) analyzed the CMAQ aerosols by assimilating
116 MODIS AOD with GSI. Their forward operator converted aerosol chemical composition into
117 AOD based on the well-known IMPROVE aerosol extinction model (Malm and Hand, 2007).
118 The IMPROVE model predicts AOD with a linear combination of aerosol chemical masses,
119 with the hydrophilic particles multiplied by a tuning factor associated with relative humidity.
120 Because building a GSI system for a new aerosol mechanism is quite technical, the official
121 GSI for the GOCART aerosols is still a primary choice for recent aerosol DA studies (Bao et
122 al., 2019; Xia et al., 2019; Hong et al., 2020).

123

124 Because of the shortcomings, the official GSI has been extended to cooperate with other
125 aerosol mechanisms in WRF-Chem. The MOSAIC mechanism in WRF-Chem simulates
126 aerosol mass and number concentrations in either four- or eight-size bins. This sectional
127 aerosol mechanism involves nitrate chemistry and can simulate SOA with the volatility basis
128 set scheme. Li et al. (2013) developed a 3D-Var scheme for assimilating the surface PM_{2.5} and
129 speciated aerosol chemical concentrations for the WRF-Chem MOSAIC aerosols. Zang et al.
130 (2016) applied this scheme to incorporate aircraft speciated aerosols in California. They
131 proved that the assimilation of aircraft profile extended the DA benefit to aerosol forecast.
132 Saide et al. (2013) proposed a revised GSI version that performed variational DA for the
133 MOSAIC aerosols. The authors generated the adjoint operator code with the automatic
134 differentiation tool (ADT), TAPENADE v3.6. The ADT used the chain rule of derivative
135 calculus on the AOD source code in WRF-Chem. They assimilated multi-source AOD data
136 with the MOSAIC aerosols over continental United States and found that incorporating multi-
137 wavelength fine-mode AOD redistributed the aerosols’ particulate mass concentration sizes.
138 The revised GSI system assimilated Korean ground-based and geostationary satellite AOD
139 datasets to improve local aerosol simulations (Saide et al., 2014, 2020). Pang et al. (2020)

140 developed the official GSI to work with the Modal Aerosol Dynamics Model for Europe with
141 the Secondary Organic Aerosol Model (MADE/SORGAM) aerosols in WRF-Chem. The
142 authors used the WRF-Chem AOD code as the forward operator to calculate the essential
143 aerosol optical properties, which were then inputted to the CRTM adjoint operator. Because
144 aerosols were externally mixed in CRTM, the setting of the internal mixture per size bin in
145 WRF-Chem was not taken into account, and the AOD of each aerosol component was
146 calculated separately.

147

148 This study provides a solution to improve the capability of the GSI 3D-Var DA system for the
149 sectional MOSAIC aerosols in WRF-Chem. We designed the tangent linear operator code for
150 AOD DA based on the WRF-Chem intrinsic aerosol optical subroutine (Fast et al., 2006), that
151 is, without using the ADT. The operator code is programmed based on the analytical
152 equations of the linear tangent model for AOD. As our revised GSI does not use the CRTM
153 module, it avoids the problem of needing to eliminate WRF-Chem aerosols characteristics
154 (e.g., aerosol mixture state and size distribution) to meet the CRTM input requirements. The
155 forward and tangent linear operators are coordinated, since they are derived from the same
156 WRF-Chem code, and are written in a single subroutine, which is coupled to the GSI at the
157 place of invoking CRTM for the AOD calculation. In addition to AOD DA, our tangent linear
158 operator has two variants to assimilate the aerosol scattering and absorption coefficients,
159 measured using a nephelometer and aethalometer, respectively.

160

161 This study verifies the effectiveness of our revised GSI system by incorporating multi-
162 wavelength aerosol optical observations that were measured during an international field
163 campaign, the Dust Aerosol Observation-Kashi, in April 2019 at Kashi city, neighboring the
164 Taklamakan Desert, northwestern China. This desert is the second largest globally, and is the
165 primary source of dust aerosols in East Asia. The dust from the desert affects the nearby
166 Tibetan Plateau (Ge et al., 2014; Jia et al., 2015; Zhao et al., 2020), air quality and climate in
167 East Asia (Huang et al., 2014), and the biogeochemical cycles in the western Pacific Ocean
168 (Calil et al., 2011). A successful DA analysis will help improve the local air quality forecast
169 and enhance our understanding of the environmental impacts of local dust storms. The
170 remainder of this paper is organized as follows. Section 2 describes the revised GSI system,
171 the experimental design, and the observed data. Section 3 presents the DA results when
172 assimilating different observations. Section 4 discusses the impact of DA on aerosol chemical
173 composition and aerosol direct radiative forcing. Finally, Section 5 provides the conclusions
174 and limitations that need further research.

175

176 **2. Methodology and Data**

177 **2.1 Forecast Model**

178 The background aerosol fields were simulated using the WRF-Chem model version 4.0 (Grell
179 et al., 2005; Fast et al., 2006). The model configurations included the Purdue Lin
180 microphysics scheme (Chen and Sun, 2002), the unified Noah land surface model (Tewari et
181 al., 2004), the Yonsei University scheme for planetary boundary layer meteorological
182 conditions (Hong et al., 2006), and the rapid radiative transfer model for general circulation
183 models (RRTMG) scheme for shortwave and longwave radiation (Iacono et al., 2008). The

184 gas-phase chemistry was simulated using the carbon bond mechanism (Zaveri and Peters,
185 1999), including aqueous-phase chemistry. The aerosol chemistry was simulated using the
186 MOSAIC mechanism (Zaveri et al., 2008), which simulated sulfate, nitrate, ammonium, black
187 carbon (BC), organic carbon (OC), sodium, calcium, chloride, carbonate, and other inorganic
188 matter (OIN, e.g., trace metals and silica). SOA was excluded from our experiments to
189 accelerate model integration. Although ignoring that SOA biased the model, the influence was
190 assumed to be small, based on low anthropogenic and biogenic emissions in the vicinity of the
191 desert. The dust emission was simulated using the GOCART dust scheme (Ginoux et al.,
192 2001), and the dust mass was included in the OIN concentration. We performed the MOSAIC
193 aerosol simulations with four-size bins (0.039–0.156 μm , 0.156–0.625 μm , 0.625–2.500 μm ,
194 and 2.5–10.0 μm dry diameters) for the anthropogenic aerosols. The sectional aerosol data in
195 the hourly model output were the aerosol dry mass mixing ratios of chemical compositions,
196 aerosol number concentration, and aerosol water content. The aerosol compositions included
197 hydrophilic particulates (i.e., SO_4^{2-} , NO_3^- , NH_4^+ , Cl^- , Na^+) and hydrophobic particulates (i.e.,
198 BC, OC, and OIN). We used the spherical particulate assumption and computed the aerosol
199 optics according to the Mie theory. The aerosol compositions were internally mixed in each
200 size bin and were externally mixed between the size bins. The internal mixing refractive index
201 was the volume-weighted mean refractive index of each composition. The WRF-Chem model
202 computed the aerosol optics at 300, 400, 600, and 999 nm and interpolated the aerosol optical
203 parameters (AOD, SSA, asymmetry factor) to eleven shortwave lengths with Ångström
204 exponents for the radiative transfer calculation.

205

206 **2.2 Assimilation System**

207 The revised GSI DA system is based on the official GSI ([https://dtcenter.org/community-](https://dtcenter.org/community-code/gridpoint-statistical-interpolation-gsi)
208 [code/gridpoint-statistical-interpolation-gsi](https://dtcenter.org/community-code/gridpoint-statistical-interpolation-gsi), Wu et al., 2002; Liu et al., 2011; Schwartz et al.,
209 2012; Pagowski et al., 2014) version 3.7. The 3D-Var DA minimizes the cost function:

210

$$211 \quad J(\mathbf{x}) = \frac{1}{2}(\mathbf{x} - \mathbf{x}_b)^T \mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}_b) + \frac{1}{2}(H(\mathbf{x}) - \mathbf{y})^T \mathbf{R}^{-1}(H(\mathbf{x}) - \mathbf{y})$$

212 (1)

213

214 where \mathbf{x} is the state vector composed of the model control variables; the subscript b denotes
215 that \mathbf{x} is the background state vector; \mathbf{y} is the vector of the observations; H is the forward
216 operator or observation operator that transfers the gridded control variables into the observed
217 quantities at the observation locations; and \mathbf{B} and \mathbf{R} are the background and observation error
218 covariance matrices, respectively.

219

220 The official GSI version only works with the GOCART aerosols for assimilating the surface-
221 layer $\text{PM}_{2.5}$ and PM_{10} (denoted as PM_x in the context) concentrations, and the 550 nm MODIS
222 AOD. Our revised GSI system assimilates PM_x concentrations, multi-wavelength aerosol
223 scattering/absorption coefficients, and AOD. Figure 1 shows the workflow of our DA system.
224 According to the AOD calculation in WRF-Chem, we can either choose the aerosol number
225 concentration (option 1), or aerosol mass concentration (option 2) as control variables. Option

226 1 is described in Li et al. (2020). In this study, we selected option 2, which is described in the
227 following subsections.

228

Figure 1

229

230 **2.2.1 Control Variables**

231 The control variables in this study were the mass mixing ratio of composition of each aerosol
232 per size bin, which corresponded to the WRF-Chem output data only. This set therefore
233 differed from previous studies that lumped aerosols per size bin as control variables. The
234 lumped aerosols avoided the burdensome task of specifying the background error statistics for
235 numerous aerosols (Li et al., 2013; Pagowski et al. 2014). Although our control variables
236 could have been further optimized, here we designed the control variable using only those that
237 substantially contributed to the total mass concentrations. We set the control variables of six
238 aerosol mass mixing ratios of SO_4^{2-} , NH_4^+ , NO_3^- , OC, BC, and OIN per size bin. Chlorine,
239 and sodium had miniscule background concentrations and remained the background values in
240 the DA analysis. There were twenty-four control variables in total for the four-size bin
241 simulations. In Kashi's case near the desert, the OIN was predominant, accounting for 62% of
242 $\text{PM}_{2.5}$ and 82% of PM_{10} .

243

244 Our design of the control variables was different from the AOD assimilation in Saide et al.
245 (2013), with theirs being the natural logarithm of the total mass mixing ratio per size bin,
246 multiplied by the thickness of the model layer. As the high model layer had a significant layer
247 thickness with low aerosol concentrations, the multiplication offset the opposite effects of
248 increasing layer thickness versus decreasing concentrations with increase in altitude. This
249 multiplication prevented the addition of many modifications for the high model layers, where
250 aerosols were low in concentration. The logarithmic transformation was used to decrease the
251 extensive value range in the control variables caused by multiplication. Since the AOD value
252 is often smaller than one, this leads to a significant negative logarithm value and a relatively
253 unconstrained DA system. Saide et al. (2013) introduced two weak constraints in their cost
254 function to cut off the user-defined "extraordinarily high" and "extraordinarily low"
255 concentrations. They repartitioned the increments of the total mass per size bin for
256 composition of each aerosol, with the background aerosol chemical mass fractions. Here,
257 neither the logarithmic transformation, nor the multiplication using layer thickness was set in
258 our DA system. Our control variable was restricted to the WRF-Chem output variable, and the
259 DA system changed the composition of each aerosol per size bin, depending on the aerosol
260 background errors.

261

262 Consistent with the set by Pang et al. (2020), aerosol water content (AWC) was not one of the
263 control variables in our GSI. Otherwise, the AWC might have increased contrary to the
264 physical constraints for the loading of hydrophilic particles, and simply as a mathematical
265 artefact. The AWC was diagnosed according to the analyzed aerosol mass concentration and
266 the background relative humidity in each DA outer loop. The hygroscopic growth was
267 calculated using the WRF-Chem code coupled with the revised GSI.

268

269 2.2.2 Tangent Linear Operator for PM_x

270 The PM₁₀ is the sum of all aerosol dry mass concentrations over the size bins, and the sum of
 271 the first three is the PM_{2.5} (Chen et al., 2019; Wang et al., 2020). Accordingly, the tangent
 272 linear operator for PM_x is the gradient of the PM_x concentration to the aerosol chemical mass
 273 concentration per size bin:

$$275 \frac{\delta[PM_x]}{\delta[C_{aer,k}]}, k = 1, \dots, n_{size}$$

276 (2)

277
 278 where n_{size} is the number of size bins and is equal to four in this study; $[.]$ denotes the mass
 279 concentration ($\mu\text{g m}^{-3}$ for PM_x); $C_{aer,k}$ is the aerosol mass mixing ratio ($\mu\text{g kg}^{-1}$) of SO_4^{2-} ,
 280 NO_3^- , NH_4^+ , OC, BC, and OIN at the k -th size bin. Because we did not multiply the chemical
 281 mass with a scaling factor to represent some unknown compositions in the summation of
 282 PM_x, Eq (2) always equals one. It means that we equally distribute the PM_x increment to
 283 each aerosol composition per size bin. The PM_{2.5} and PM₁₀ are assimilated in the same way.
 284 When the fine and coarse particles are assimilated simultaneously, we assimilate the
 285 concentration of PM_{2.5} and the coarse particulate (PM₁₀-PM_{2.5}).

287 2.2.3 Forward Operator for Aerosol Optics in WRF-Chem

288 We used the original forward operator in WRF-Chem for the aerosol optical parameters (Fast
 289 et al., 2006). AOD is calculated as a function of wavelength according to Mie theory. The
 290 columnar AOD τ is the sum of layer AOD across the n_z model layers:

$$292 \tau = \sum_{z=1}^{n_z} \tau_z = \sum_{z=1}^{n_z} \sum_{k=1}^{n_{size}} e_{ext,z,k} \cdot n_{z,k} \cdot H_z$$

293 (3)

294
 295 where $e_{ext,z,k}$ is the extinction cross section of a single mixing particle in the k -th size bin at the
 296 z -th model layer, $n_{z,k}$ is the aerosol number concentration, and H_z is the layer thickness. At the
 297 surface, the ambient aerosol scattering (E_{sca}) and absorbing (E_{abs}) coefficients that are
 298 measured by the nephelometer and aethalometer, respectively, are represented in the model as
 299

$$300 E_{sca} = \sum_{k=1}^{n_{size}} e_{sca,1,k} \cdot n_{1,k}$$

$$301 E_{abs} = \sum_{k=1}^{n_{size}} e_{abs,1,k} \cdot n_{1,k}$$

302 (4)

303
 304 where $e_{sca,1,k}$ and $e_{abs,1,k}$ are the scattering and absorption cross section of a particle at the
 305 surface. There is a relationship:

306

$$e_{ext,z,k} = e_{sca,z,k} + e_{abs,z,k} \quad (5)$$

309

The extinction cross section $e_{ext,z,k}$ of a wet particle with radius $r_{wet,z,k}$ is:

311

$$e_{ext,z,k} = p_{ext,z,k} \cdot \pi \cdot r_{wet,z,k}^2 \quad (6)$$

314

where $p_{ext,z,k}$ is the extinction efficiency, given the desired mixing refractive indexes and the wet particle radius. The $p_{ext,z,k}$ is attained through the Chebyshev polynomial interpolation:

317

$$p_{ext,z,k} = \exp \left\{ \sum_{j=1}^{n_{coef}} c_{ch}(j) \cdot c_{ext,z,k}(j) \right\} \quad (7)$$

319

where c_{ch} is the coefficient of n_{coef} order Chebyshev polynomials, $c_{ext,z,k}$ is the polynomial value for the extinction efficiency of the particle, which is an internal mixture of all aerosol compositions (i.e., the control variables plus chlorine, sodium, and AWC). The radius in the AOD subroutine code is in a logarithmic transform to handle the broad particle size range from 0.039 μm to 10 μm . The exponential function in Eq. (7) transforms the logarithm radius back to the normal radius. The aerosol number concentration $n_{z,k}$, and the aerosol dry (wet) mass concentration $m_{i,z,k}$ have a linkage through the dry (wet) particle radius $r_{dry,z,k}$ ($r_{wet,z,k}$) and the density ρ_i of each aerosol chemical composition:

328

$$n_{z,k} = \sum_i^{n_{wet_aer}} \frac{m_{i,z,k}}{\rho_i} \cdot \frac{3}{4\pi \cdot r_{wet,z,k}^3} = \sum_i^{n_{dry_aer}} \frac{m_{i,z,k}}{\rho_i} \cdot \frac{3}{4\pi \cdot r_{dry,z,k}^3} \quad (8)$$

330

331

Both the dry and wet particle radius will appear in the tangent linear operator. The difference between the second and the third terms in Eq (8) is whether aerosol water content is counted.

n_{wet_aer} is the number of aerosol chemical composition plus aerosol water content ($n_{wet_aer} = n_{dry_aer} + 1$).

336

2.2.4 Tangent Linear Operator Developed for AOD

338

339

340

341

As per the forward operator in Eq. (3) in WRF-Chem, we developed the tangent linear operator for AOD, which requires the derivative of τ in Eq. (3) to the aerosol dry mass concentration (aerosol water content is not a control variable), $m_{i,z,k}$:

342

$$\frac{\delta\tau}{\delta m_{i,z,k}} = \frac{\delta\tau_z}{\delta m_{i,z,k}} = \frac{\delta e_{ext,z,k} \cdot n_{z,k} \cdot H_z}{\delta m_{i,z,k}} + \frac{e_{ext,z,k} \cdot \delta n_{z,k} \cdot H_z}{\delta m_{i,z,k}} + \frac{e_{ext,z,k} \cdot n_{z,k} \cdot \delta H_z}{\delta m_{i,z,k}} \quad (9)$$

343

344

345 The first term on the righthand side of Eq. (9) indicates the change in AOD as the perturbation
 346 of extinction cross section. According to Eq. (6), considering that the particle radius is
 347 constant, $\delta e_{ext,z,k}$ is represented as:

$$348 \delta e_{ext,z,k} = \delta p_{ext,z,k} \cdot \pi \cdot r_{wet,z,k}^2 \quad (10)$$

349 where $\delta ch(j) = 0$ assuming that the particle radius is constant. This assumption simplifies
 350 the tangent linear operator and is also employed in Saide et al. (2013).

351 Equation (10) is expanded with the derivative of Eq. (7):

$$352 \delta p_{ext,z,k} = p_{ext,z,k} \cdot \left\{ \sum_{j=1}^{n_{coef}} c_{ch}(j) \cdot \delta c_{ext,z,k}(j) \right\} \quad (11)$$

353 By expanding $\delta c_{ext,z,k}$ in Eq. (11), we have:

$$354 \delta c_{ext,z,k}(j) = \delta w_{00} \cdot E_{ext,00}(j) + \delta w_{01} \cdot E_{ext,01}(j) + \delta w_{10} \cdot E_{ext,10}(j) + \delta w_{11} \cdot E_{ext,11}(j) \quad (12)$$

355 where the four parameters of E_{ext} indicate the extinction efficiencies in the Mie lookup table
 356 surrounding the point with the desired mixing refractive indexes, and the wet particle radius.
 357 The interpolation weights δw are determined as:

$$358 \delta w_{00} = (v - 1)\delta u + (u - 1)\delta v \quad \delta w_{01} = (1 - v)\delta u - u\delta v$$

$$359 \delta w_{10} = (1 - u)\delta v - v\delta u \quad \delta w_{11} = u\delta v + v\delta u \quad (13)$$

360 where

$$361 u = \frac{R_{mix} - R_{low}}{R_{up} - R_{low}} \quad \delta u = \frac{\delta R_{mix}}{R_{up} - R_{low}}$$

$$362 v = \frac{I_{mix} - I_{low}}{I_{up} - I_{low}} \quad \delta v = \frac{\delta I_{mix}}{I_{up} - I_{low}} \quad (14)$$

363 In Eq. (14), R_{mix} and I_{mix} are the aerosol volume-weighted mean real and imaginary parts of
 364 complex refractive indices, respectively. R_{up} (I_{up}) and R_{low} (I_{low}) are the nearest upper and
 365 lower limits for R_{mix} (I_{mix}) in the Mie table. Considering $V_{wet,z,k}$ is the volume of all aerosol dry
 366 masses plus aerosol water content, the real and imaginary parts and their derivatives are:

$$367 R_{mix,z,k} = \sum_i^{n_{wet,aer}} R_i \cdot \frac{m_{i,z,k}}{\rho_i \cdot V_{wet,z,k}} \quad \delta R_{mix,z,k} = \frac{R_i}{\rho_i \cdot V_{wet,z,k}} \cdot \delta m_{i,z,k}$$

384
$$I_{mix,z,k} = \sum_i^{n_{wet,aer}} I_i \cdot \frac{m_{i,z,k}}{\rho_i \cdot V_{wet,z,k}} \quad \delta I_{mix,z,k} = \frac{I_i}{\rho_i \cdot V_{wet,z,k}} \cdot \delta m_{i,z,k}$$

385

386

387

(15)

where

388
$$V_{wet,z,k} = \sum_i^{n_{wet,aer}} \frac{m_{i,z,k}}{\rho_i}$$

389

390

(16)

Put Eq. (12), Eq. (13) into Eq. (11) leads to:

391

392

393
$$\delta p_{ext,z,k} = [(v-1)\alpha_{sca,00} + (1-v)\alpha_{sca,01} - v\alpha_{sca,10} + v\alpha_{sca,11}]\delta u +$$

394
$$[(u-1)\alpha_{abs,00} - u\alpha_{abs,01} + (1-u)\alpha_{abs,10} + u\alpha_{abs,11}]\delta v$$

395

(17)

where

397
$$\alpha_{sca,00} = p_{sca,1,k} \cdot \sum_{j=1}^{n_{coef}} c_{ch}(j) \cdot E_{sca,00}(j) \quad \alpha_{sca,01} = p_{sca,1,k} \cdot \sum_{j=1}^{n_{coef}} c_{ch}(j) \cdot E_{sca,01}(j)$$

398
$$\alpha_{sca,10} = p_{sca,1,k} \cdot \sum_{j=1}^{n_{coef}} c_{ch}(j) \cdot E_{sca,10}(j) \quad \alpha_{sca,11} = p_{sca,1,k} \cdot \sum_{j=1}^{n_{coef}} c_{ch}(j) \cdot E_{sca,11}(j)$$

399
$$\alpha_{abs,00} = p_{abs,1,k} \cdot \sum_{j=1}^{n_{coef}} c_{ch}(j) \cdot E_{abs,00}(j) \quad \alpha_{abs,01} = p_{abs,1,k} \cdot \sum_{j=1}^{n_{coef}} c_{ch}(j) \cdot E_{abs,01}(j)$$

400
$$\alpha_{abs,10} = p_{abs,1,k} \cdot \sum_{j=1}^{n_{coef}} c_{ch}(j) \cdot E_{abs,10}(j) \quad \alpha_{abs,11} = p_{abs,1,k} \cdot \sum_{j=1}^{n_{coef}} c_{ch}(j) \cdot E_{abs,11}(j)$$

401

402

403

(18)

404 The subscripts of *sca* and *abs* in Eq. (17) and (18) denote “scattering” and “absorption”,

405 respectively. The first term on the righthand side of Eq. (9) is determined using Eq. (10) and

406 Eq. (17). The second term on the righthand side of Eq. (9) indicates the linkage of the aerosol

407 number and mass concentrations. It is the derivative of dry particle in Eq. (8) by assuming a

408 constant radius:

409

410
$$\delta n_{z,k} = \frac{3 \cdot \delta m_{i,z,k}}{4\pi \cdot r_{dry,z,k}^3 \cdot \rho_i}$$

411

(19)

412

413 The third term on the righthand side of Eq. (9) contains the derivative of the layer thickness to

414 the concentrations in this layer. This indicates that the light attenuation length based on per

415 unit concentration, which can be intuitively represented by the ratio of layer thickness to the
 416 aerosol mass concentration in this layer. Putting Eq. (10) and Eq. (19) into Eq. (9), we have
 417 the original formula of the tangent linear operator for AOD for the aerosol dry mass
 418 concentration:

419
 420

$$\begin{aligned}
 421 \quad \frac{\delta\tau}{\delta m_{i,z,k}} &= \frac{\delta\tau_z}{\delta m_{i,z,k}} = \frac{\delta e_{ext,z,k} \cdot n_{z,k} \cdot H_z}{\delta m_{i,z,k}} + \frac{e_{ext,z,k} \cdot \delta n_{z,k} \cdot H_z}{\delta m_{i,z,k}} + \frac{e_{ext,z,k} \cdot n_{z,k} \cdot \delta H_z}{\delta m_{i,z,k}} = \\
 422 \quad & \{[(v-1)\alpha_{sca,00} + (1-v)\alpha_{sca,01} - v\alpha_{sca,10} + v\alpha_{sca,11}] \cdot \frac{\pi \cdot r_{wet,z,k}^2 \cdot R_i \cdot n_{z,k} \cdot H_z}{\rho_i \cdot V_{wet,z,k} \cdot (R_{up,z,k} - R_{low,z,k})} \\
 423 \quad & + \\
 424 \quad & [(u-1)\alpha_{abs,00} - u\alpha_{abs,01} + (1-u)\alpha_{abs,10} + u\alpha_{abs,11}] \cdot \frac{\pi \cdot r_{wet,z,k}^2 \cdot I_i \cdot n_{z,k} \cdot H_z}{\rho_i \cdot V_{wet,z,k} \cdot (I_{up,z,k} - I_{low,z,k})} + \\
 425 \quad & \frac{3e_{ext,z,k} \cdot H_z}{4\pi \cdot r_{dry,z,k}^3 \cdot \rho_i} + \frac{e_{ext,z,k} \cdot n_{z,k} \cdot H_z}{m_{i,z,k}} \} \cdot \beta \\
 426 \quad & \\
 427 \quad & \hspace{15em} (20) \\
 428 \quad &
 \end{aligned}$$

429 where β is the factor that changes the unit of mass from $\mu\text{g kg}^{-1}$ to $\mu\text{g m}^{-3}$. The last righthand
 430 term in Eq. (20) may not have a quick convergence in the DA outer loops because the aerosol
 431 mass concentration $m_{i,z,k}$ in the denominator often has a low bias, which introduces an error
 432 into the operator. The error is amplified by the layer thickness H_z in the numerator. Thus, the
 433 operator of Eq. (20) cannot lead to a stable analysis. For this reason, we changed the operator
 434 to account for the columnar mean aerosol extinction coefficient which is described as follows:
 435

$$\begin{aligned}
 436 \quad \frac{\delta(\overline{e_{ext} \cdot n})}{\delta m_{i,z,k}} &= \frac{H_z}{\sum H_z} \cdot \frac{\delta(e_{ext,z,k} \cdot n_{z,k})}{\delta m_{i,z,k}} = \frac{H_z}{\sum H_z} \cdot \left[\frac{\delta e_{ext,z,k} \cdot n_{z,k}}{\delta m_{i,z,k}} + \frac{e_{ext,z,k} \cdot \delta n_{z,k}}{\delta m_{i,z,k}} \right] = \\
 437 \quad & \{[(v-1)\alpha_{sca,00} + (1-v)\alpha_{sca,01} - v\alpha_{sca,10} + v\alpha_{sca,11}] \cdot \frac{\pi \cdot r_{wet,z,k}^2 \cdot R_i \cdot n_{z,k}}{\rho_i \cdot V_{wet,z,k} \cdot (R_{up,z,k} - R_{low,z,k})} \\
 438 \quad & + \\
 439 \quad & [(u-1)\alpha_{abs,00} - u\alpha_{abs,01} + (1-u)\alpha_{abs,10} + u\alpha_{abs,11}] \cdot \frac{\pi \cdot r_{wet,z,k}^2 \cdot I_i \cdot n_{z,k}}{\rho_i \cdot V_{wet,z,k} \cdot (I_{up,z,k} - I_{low,z,k})} + \\
 440 \quad & \frac{3e_{ext,z,k}}{4\pi \cdot r_{dry,z,k}^3 \cdot \rho_i} \} \cdot \beta \cdot \frac{H_z}{\sum H_z} \\
 441 \quad & \\
 442 \quad & \hspace{15em} (21) \\
 443 \quad &
 \end{aligned}$$

444 In Eq. (21), the operator is based on the extinction coefficient at each layer, weighted by the
 445 layer thickness normalized to the total model layer thickness. Correspondingly, the AOD
 446 observations and AOD observation error are divided by the total layer thickness at the

447 observation location. Note that the dry ($r_{dry,z,k}$) and wet ($r_{wet,z,k}$) particle radiuses are both
 448 present in Eq (21). Because aerosol water content is not a control variable, $r_{dry,z,k}$ is used in Eq
 449 (19) and appears in Eq (21). Aerosol water content participates the computation of internal
 450 mixing refractive indexes, and thus $r_{wet,z,k}$ is also present in Eq (21). Equation (21) is the final
 451 tangent linear operator for AOD DA in this study.

452

453 **2.2.5 Tangent Linear Operator Developed for Surface Aerosol Attenuation Coefficients**

454 The aerosol scattering and absorption coefficients measured by the nephelometer and
 455 aethalometer, respectively, are similar to the aerosol extinction coefficient at the surface in
 456 Eq. (21). Neither of the two coefficients address the layer thickness. The operator for the
 457 aerosol scattering coefficient measured by nephelometer is described as follows:

458

$$\begin{aligned}
 460 \quad \frac{\delta(e_{sca,1,k} \cdot n_{1,k})}{\delta m_{i,1,k}} &= \{[(v-1)\alpha_{sca,00} + (1-v)\alpha_{sca,01} - v\alpha_{sca,10} + v\alpha_{sca,11}] \\
 461 \quad &\cdot \frac{\pi \cdot r_{wet,1,k}^2 \cdot R_i \cdot n_{1,k}}{\rho_i \cdot V_{wet,1,k} \cdot (R_{up,1,k} - R_{low,1,k})} + \frac{3e_{sca,1,k}}{4\pi \cdot r_{dry,1,k}^3 \cdot \rho_i}\} \cdot \beta \\
 459 \quad & \hspace{15em} (22)
 \end{aligned}$$

462

463 where the symbols have the same meaning as before, and the subscript one in Eq. (22)
 464 denotes the surface layer. The operator for the aerosol absorption coefficient measured by
 465 aethalometer is

466

$$\begin{aligned}
 468 \quad \frac{\delta(e_{abs,1,k} \cdot n_{1,k})}{\delta m_{i,1,k}} &= \{[(u-1)\alpha_{abs,00} - u\alpha_{abs,01} + (1-u)\alpha_{abs,10} + u\alpha_{abs,11}] \\
 469 \quad &\cdot \frac{\pi \cdot r_{wet,1,k}^2 \cdot I_i \cdot n_{1,k}}{\rho_i \cdot V_{wet,1,k} \cdot (I_{up,1,k} - I_{low,1,k})} + \frac{3e_{abs,1,k}}{4\pi \cdot r_{dry,1,k}^3 \cdot \rho_i}\} \cdot \beta \\
 467 \quad & \hspace{15em} (23)
 \end{aligned}$$

470

471 As shown in the operators, the gradients of the aerosol mass concentrations rely on the aerosol
 472 number concentration; meanwhile, the number concentration is estimated according to the
 473 mass concentration and the particle radius. The two concentrations are intertwined in the DA
 474 system, indicating the nonlinearity of the operator. This nonlinearity is handled with a
 475 succeeding minimization of the cost function within the GSI. That is, the cost function is first
 476 minimized with the number concentration in the background field, and the number
 477 concentration is updated with the first analyzed aerosol mass concentrations. In the second
 478 minimization, the number concentration assessed in the first analysis constructs a new
 479 operator value, resulting in a new analysis of mass concentrations. This iterative process is
 480 denoted as the ‘‘outer loop,’’ which is repeated several times to attain the final analysis
 481 (Massart et al., 2010). We set ten maximum iterations to handle the nonlinearity in the
 482 operator. The cost function in most analyses reaches the minimum in two or three outer loops.
 483 The WRF-Chem AOD code is coupled into the GSI subroutine at the place of invoking
 484 CRTM. The tangent linear operators of Eq. (21), Eq. (22), and Eq. (23) are simultaneously
 485 determined in the subroutines, which are cyclically invoked in the outer loops within the GSI.
 486

487 **2.2.6 Aerosol Complex Refractive Indexes in GSI**

488 Table S1 in the supplementary document shows the complex refractive indexes for each
489 aerosol chemical composition in the revised GSI. The refractive indexes are for eleven
490 wavelengths, including four for CE318, three for nephelometer, three for aethalometer, and
491 one for 550 nm MODIS AOD (not assimilated in this study). The real parts of refractive
492 indexes of sulfate, nitrate, and ammonium are similar and refer to Toon et al.'s (1976) data.
493 The real part is 1.53 at 440 nm and decreases to 1.52 at 1020 nm. The refractive indexes of
494 OC and BC are constant across the wavelengths, being $1.55-0.001i$ for OC (Chen and Bond,
495 2010) and $1.95-0.79i$ for BC (Bond and Berstrom, 2006). The dust refractive index's real part
496 is a constant value of 1.54 (Zhao et al., 2010). The dust refractive index's imaginary part
497 depends on the dust mineralogy, size distribution, and shape, which are associated with the
498 dust sources. The imaginary part varies a lot at the same dust source. Cheng et al. (2006)
499 reported the desert dust refractive index in winter and spring at Dunhuang, a city adjacent to
500 the northeast side of the Taklamakan desert. Their imaginary part value was approximately in
501 the ranges of 0.0008 to 0.0028 at 440 nm, 0.0006 to 0.0030 at 670 nm, 0.0005 to 0.0036 at
502 870 nm, and 0.0005 to 0.0040 at 1020 nm (See Figure 9 in their paper). Recently, Di Biagio et
503 al. (2019) retrieved the dust's imaginary part in the Taklimakan desert's north edge (41.83°N,
504 85.88°E). Their dust imaginary part decreased from 0.0018 ± 0.0008 at 370 nm to
505 0.0005 ± 0.0002 at 950 nm, much lower than the generic values in climate models. The
506 imaginary part's retrieval uncertainty is related to the iron oxide in dust samples, the cutoff
507 coarse particle size ($<10\ \mu\text{m}$ in Di Biagio et al., 2019), and the assumption of spherical
508 particles applied in the retrieval algorithm. Here, we admit the high uncertainty and use the
509 imaginary part following the generic model values (Table S1), which are higher than the
510 upper limits of the data of Di Biagio et al. (2019) and are close to the values of Cheng et al.
511 (2006). The desert dust has a stronger absorption at shortwave wavelengths. The refractive
512 index of a wavelength without exact literature data uses the nearby wavelength's data in
513 literature. The supplement also shows the aerosol density (Table S2) that follows the density
514 data in Barnard et al. (2010). The aerosol density is necessitated to compute aerosol optical
515 parameters in the AOD forward operator and construct our tangent linear operator.

516

517 **2.3 Background Error Covariance (BEC)**

518 Many aerosol DA studies used the National Meteorological Center (NMC) method (Parrish
519 and Derber, 1992) to model the BEC matrix. The NMC method uses long-term archived
520 weather data that are created in the forecast cycles. It computes the statistical differences
521 between two forecasts with different leading lengths (e.g., 24 h and 48 h), but which are valid
522 at the same time. The NMC method is workable because solving global weather forecasts is
523 an initial value problem of mathematical physics. That is, a slight difference in the initial
524 atmospheric state would lead to a substantially different prediction, because of the chaos in
525 the atmosphere. However, a regional model is a boundary value problem. Meteorological
526 reanalysis data drive the regional chemistry simulation, and the driving data quality affects the
527 simulation (Giorgi and Mearns, 1999). The WRF-Chem simulations in the NMC method only
528 reflected the influences of using different initial conditions. As the model runs, the influence
529 of the initial conditions becomes weak, while the influence of lateral boundary conditions
530 always takes effect. Because the same reanalysis data drive the paring regional model

531 simulations, the following lateral boundary conditions for the simulations of the two leading-
532 lengths are similar. This leads to a limited regional model difference when using the NMC
533 method. That is, the NMC method's BEC on the meteorology would underestimate the
534 aerosol error in WRF-Chem. Kumar et al. (2019) assimilated AOD in the contiguous United
535 States based on the NMC method's BEC. They perturbed the background emissions by adding
536 the gridded mean differences of four emission inventories. Their analysis with the BEC
537 accounting for meteorology and emissions uncertainties reduced the AOD bias by 38%,
538 superior to 10% bias reduction counting the meteorology uncertainty alone.

539

540 Some aerosol DA studies have created background error variance using the ensemble
541 simulations by randomly disturbing model lateral boundary conditions and surface emissions
542 (Peng et al., 2017; Ma et al., 2020). The ensemble experiments better represent the model
543 error, but significantly increase the computational burden. Here, we used the standard
544 deviation of hourly aerosol concentrations in April in the background field (first guess field)
545 to represent the background error variance. The rationale of this approach is that the Tarim
546 Basin acts as a “dust reservoir” and traps dust particles for a period, before being carried long-
547 distance by wind (Fan et al., 2020). The model bias in dust dominates the model aerosol error,
548 and is correlated with the aerosol variation as the weather fluctuates. The model bias is small
549 on clear days when the aerosol concentration is low. Conversely, the bias is large when the
550 mean concentration is high: that is, on heavily-polluted days. Because the mean aerosol
551 concentration correlated positively with the aerosol variation, we used the standard deviation
552 of aerosol concentration to represent the background aerosol error. This approach was similar
553 to Sič et al. (2016), who set a percentage of the first guess field for the background error
554 variance. Our approach prioritizes DA modification of aerosols which have high background
555 mean concentrations.

556

557 We calculated the statistics of the background error, including the aerosol standard deviation
558 and the horizontal and vertical correlation length scales, using the GENerate the Background
559 Errors (GEN-BE) software (Descombes et al., 2015), based on the one-month hourly aerosol
560 concentrations in WRF-Chem. We obtained the statistics of four static BECs for the four DA
561 analysis hours (i.e., 00:00, 06:00, 12:00, and 18:00 UTC), respectively. The DA procedures
562 for the April 2019 data repeatedly use the statistics of the background error at the
563 corresponding analysis time. A usual strategy to enrich the samples of model results for
564 calculating the statistics is to gather model grid points with similar characteristics of the
565 atmosphere, referred to as “binning.” The statistics are spatially averaged over the binned grid
566 points. The default strategy in the GEN_BE for GSI is latitude-binning, which creates a
567 latitude-dependent error correlation function (Figure 2a). The latitude binning is generally
568 used for latitude flow dependency and works for large and global domains (Wu et al., 2002).
569 However, we found that using the latitude-binning strategy overestimated the surface PM_x
570 concentration when assimilating aerosol optical observations. One reason for this was related
571 to the model bias in particle extinction efficiency, as discussed in Section 3.3. Another
572 plausible reason is related to the vertical profile of the background model error. The
573 maximum dust error occurred at the surface of the desert (Figure 2e) because of the local dust
574 emission sources, but the maximum error at Kashi was at the dust transporting layer above the

575 surface (Figure 2d). Owing to the vast extent of the Taklamakan Desert, the latitude-binning
576 suppressed the local error characteristics at Kashi, and led to a vertical error profile (Figure
577 2c) similar to that over the desert (Figure 2e).

578

579 For this reason, we used the standard deviation of the control variable at each model grid to
580 replace the latitude-binning standard deviation. The horizontal and vertical correlation length
581 scales were calculated based on the latitude-binning data. Figure 3 shows the background
582 error statistics generated by the GEN_BE software, which provided the input to the GSI. The
583 OIN component showed high background errors in the third and fourth particle sizes at the
584 transporting layer above the surface (Figure 3f). The aerosol compositions related to
585 anthropogenic emissions (i.e., sulfate, nitrate, ammonium, OC, and BC, referred to here as
586 ‘anthropogenic aerosols’) had maximum errors in the second particle size and showed the
587 greatest vertical error at the surface. The background error for OIN composition was higher
588 than that for anthropogenic aerosols by a factor of two or three, because of the high
589 background dust concentration in the city.

590

591 The horizontal and vertical correlation length scales determine the range of observation
592 innovations spreading from the observation locations. The horizontal influences had small
593 changes in altitude within the lowest 15 model layers (below a height of ~5 km), indicating
594 that the dust transport layer was well-mixed in the lower atmosphere. This deep dust layer
595 was consistent with the dust simulation by Meng et al. (2019). They showed that the dust in
596 spring was vertically mixed in a thick boundary layer to a height of 3–5 km in the Tarim
597 Basin. The vertical correlation length scales first increased from low values at the surface, to
598 high values at ~2.5 km in height (for the 8–9 layers), indicating that strong winds yielded
599 intense aerosol upward flux. The vertical correlation length scale quickly decreased from the
600 maximum value, with further increase in altitude corresponding to the large particle gradient
601 at the upper edge of the transporting layer. The latter was associated with laminar air motion
602 during the dust storm.

603

604 The background model errors were independent of particle size, which would have tended to
605 accumulate the DA modification in a single size bin that had the maximum background error
606 (e.g., the OIN in the fourth particle size). To avoid excessive accumulation of increments in a
607 single size bin, we added a one-dimensional recursive filter for the background covariances of
608 control variables across the size bins within the GSI. The inter-size bin correlation length
609 scale was four bin units.

610

Figure 2, Figure 3

611

612 **2.4 Observational Data and Errors**

613 The Dust Aerosol Observation–Kashi field campaign was performed at Kashi from
614 00:00UTC 25 March to 00:00 UTC 1 May 2019. The site was placed in the Kashi campus of
615 the Aerospace Information Research Institute, Chinese Academy of Sciences (39.50°N,
616 75.93°E; Li et al., 2018), about 4 km in the northwest to the Kashi city. The site aerosol
617 observations used for our DA analysis included: (1) the multi-wavelength AOD measured by

618 the sun-sky photometer (Cimel CE318); (2) the multi-wavelength aerosol scattering and
619 absorption coefficients at the surface, measured with a nephelometer (Aurora 3000) and
620 aethalometer (Magee AE-33), respectively, during the campaign; and (3) the hourly PM_{2.5} and
621 PM₁₀ observations, measured with a METONE BAM-1020 continuous particulate monitor.
622 All the instruments were deployed at the roof of a three stories height building on the campus.
623 Please refer to Li et al. (2020) for more details about the field campaign.
624

625 Table 1 summarizes the observation periods, the wavelengths of the aerosol optical data, and
626 the observation errors. The multi-wavelength data of each type of observation were
627 assimilated simultaneously. The observation errors of PM_x are handled in the conventional
628 way (Schwartz et al., 2012; Chen et al., 2019), which contains the measurement error (e_1) and
629 the representative error (e_2). The measurement error is the sum of a baseline error of 1.5 μg
630 m^{-3} and 0.75% of the observed PM_x concentration. The representative error is the
631 measurement error multiplied by the half-squared ratio of the grid spacing to the scale
632 distance. The scale distance denotes the site representation in GSI and has four default values
633 of 2, 3, 4, and 10 km, corresponding to the urban, unknown, suburban, and rural sites. We
634 used 3 km for the scale distance in this study. As we had a single site in Kashi, it is difficult to
635 estimate the site representation error. Since the DA analysis was based on the child model
636 domain with a horizontal resolution of 5 km, close to the site distance to the Kashi urban area,
637 we assumed the aerosol optical measurement had good representativeness of the model grid
638 covering the site. The observation error of CE318 AOD took the AERONET AOD
639 uncertainty of 0.01 in cloud-free conditions (Holben et al., 1998). The AOD observational
640 error was further divided by the total model layer thickness in GSI. It is difficult to determine
641 instrumental errors in nephelometers and aethalometers, and we empirically set their
642 instrumental errors to 10 Mm^{-1} , equivalent to the magnitude of the Rayleigh extinction
643 coefficient. The observational errors were uncorrelated, with \mathbf{R} being a diagonal matrix.
644

Table 1

645

646 2.5 Experimental Design

647 The WRF-Chem simulations were configured in a two-nested domain centered at 82.9 °E,
648 41.5 °N. The coarse domain was a 120×100 (west-east × north-south) grid with a horizontal
649 resolution of 20 km that covered the Taklamakan Desert, and the fine domain was an 81×61
650 grid with a resolution of 5 km, focusing on Kashi and environs (Figure 4a). Both domains had
651 41 vertical levels extending from the surface to 50 hPa. The lowest model layer at the site was
652 approximately 25-meter height from the ground. The two domains were two-way coupled.
653 The parent domain covered the entire dust emission source, providing dust transport fluxes at
654 the lateral boundaries of the fine domain. The aerosol radiative effect was set to provide
655 feedback on the meteorology. The indirect effect of aerosols was not set in the experiments.
656 Initial and lateral boundary meteorological conditions for WRF-Chem were the one-degree
657 resolution of the National Centers for Environmental Prediction Final Analysis data created
658 by the Global Forecast System model. The meteorological lateral boundary conditions for the
659 coarse domain were updated every six hours, and were linearly interpolated between the
660 updates in WRF-Chem. We did not set the chemical boundary conditions for the coarse

661 domain. The Multiresolution Emission Inventory of China (MEIC) for the year 2010
662 (www.meicmodel.org) provided anthropogenic emission levels. The yearly emission
663 differences in 2010-2019 may bias the aerosol chemical simulation, but this bias is hard to be
664 quantified as lack of aerosol chemical observations in this city. As the significant pollutant at
665 Kashi is dust, we just ignore the model uncertainties due to the yearly differences in
666 anthropogenic emission inventories. The biogenic emission levels were estimated online
667 using the Model of Emissions of Gases and Aerosols from Nature (Guenther et al., 2006).
668 Wildfire emissions were not set in the experiments.

669
670 We conducted a one-month WRF-Chem simulation for April 2019, starting at 00:00 UTC 27
671 March and discarding the first five days for spin-up. The revised GSI system modified the
672 aerosols in the fine domain at 00:00, 06:00, 12:00, and 18:00 UTC each day starting from
673 00:00 UTC 1 April until the end of the month. We assimilated the observations four times a
674 day because the reanalyzed meteorological data were available for the four time slices, which
675 facilitated the model restarting from the DA analyses. The hourly PM_x observations were
676 assimilated at the exact time of analysis. The observed AOD and aerosol scattering/absorption
677 coefficients were assimilated when they fell within 3 hours before the time of analysis. Table
678 2 shows the DA experiments. The literal meanings of the experimental names denote the
679 observations that were individually or simultaneously assimilated. To study the impact of DA
680 on aerosol direct radiative forcing (ADRF), we modified the WRF-Chem code to calculate the
681 shortwave irradiance with and without aerosols at each model integration step. The modified
682 WRF-Chem model restarted from each DA analysis and ran to the next analysis time. Each
683 running performed the radiation transfer calculation twice, and each calculation saw the
684 aerosols and clean air, respectively. The irradiance difference between the two pairing calls
685 was aerosol radiative forcing. Section 4.2 shows the DA effects on the clear-sky ADRF
686 values.

687

Table 2, Figure 4

688

689 **3. Results**

690 **3.1 Evaluation of Control Experiment**

691 Table 2 shows the monthly mean values and correlations between the observed data and the
692 model results. The statistical values were based on the pairing data between the model results
693 and the observations. Figures 6 show the surface PM_x concentrations, aerosol scattering
694 coefficients, and AOD when assimilating the observations at 00:00, 06:00, 12:00, and 18:00
695 UTC each day in April.

696

697 Kashi is in the junction between the Tian Shan Mountains to the west and the Taklamakan
698 Desert to the east (Figure 5a). In the Tarim Basin, the prevailing surface wind is easterly or
699 northeasterly, which raises dust levels and carries the particles westward (Figure 5b). An
700 intense dust storm hit the city at noon on 24 April 2019, with a peak PM₁₀ concentration
701 exceeding 3,000 $\mu\text{g m}^{-3}$. The dust storm travelled across the northern part of the desert and
702 carried the dust particles to Kashi and the mountainous area (Figure 5c, d). A few mild dust
703 storms occurred at Kashi on April 3–5, April 8–11, and April 14–17 (Figure 6b), and the

704 maximum PM_{10} concentrations were in the range of 400–600 μgm^{-3} . The time series of $PM_{2.5}$,
705 aerosol scattering/absorption coefficient, and AOD showed patterns, similar to those for PM_{10}
706 (Figure 6).

707

708 WRF-Chem captured the main dust episodes, but significantly underestimated the aerosols at
709 Kashi (Table 2). The background monthly mean concentrations of $PM_{2.5}$ and PM_{10} were 17%
710 and 41% lower than the observed values, respectively, with a low correlation ($R < 0.3$). The
711 simulated dust storm on 24 April was a mild dust event and had a maximum PM_{10} of $\sim 300 \mu\text{g}$
712 m^{-3} , one-tenth of the observed value. The model underestimates the aerosol
713 scattering/absorption coefficients and AOD by 40–70%.

714

715 The OIN component accounted for the model bias in PM_{10} on dusty days. Zhao et al. (2020)
716 proposed that the GOCART scheme reproduced dust emission fluxes under conditions of
717 weak wind erosion but underestimated the emissions in conditions of strong wind erosion. We
718 did not assimilate meteorology. The model bias in the surface wind introduces errors in dust
719 emission, and places bias on the number of dust particles entering the city. In the non-dust
720 days with the PM_{10} lower than the 25th percentile PM_{10} in April, the model hourly $PM_{2.5}$ on
721 average only accounted for 60% of the observed data levels. The $PM_{2.5}$ low bias could be due
722 to the lack of SOA chemistry in our experiments and the emission low bias in the residential
723 sector which is a major source of anthropogenic emissions for $PM_{2.5}$, BC, and OC in the
724 developing western area. The residential sector accounts for 36–82% of these emissions,
725 according to the MEIC emission inventory (Li et al., 2017) and is the primary source of
726 uncertainty in anthropogenic emissions inventories in China.

727

Figure 5

728

729 **3.2 Assimilating $PM_{2.5}$ and PM_{10} Concentrations**

730 Simultaneous assimilation of the observed PM_x (DA_ PM_x) improved both the fine and
731 coarse particle concentrations, with a substantial improvement in the third and fourth particle
732 sizes of the OIN composition (Figure 8f). The analyzed monthly mean PM_{10} increased to
733 $329.3 \mu\text{g m}^{-3}$, with a high correlation of 0.99. The analyzed monthly mean $PM_{2.5}$ was
734 improved to $89.3 \mu\text{g m}^{-3}$, although it was still lower than the observed levels, with a high
735 correlation of 0.89. The low bias in $PM_{2.5}$ and the high bias in PM_{10} in the analyses were both
736 mainly in the dust storm on 24–25 April (Figure 6a, d). Applying the inter-size bin correlation
737 length caused the interlinked analyses of $PM_{2.5}$ and PM_{10} . In the desert area, the coarse and
738 fine dust are readily affected by the magnitude of BEC of the fourth size-bin OIN (oin_a04).
739 We intentionally decreased the BEC of oin_a04 by 10% each time to 30% of its original
740 value. The magnitude of 30% of oin_a04 was comparable to the magnitude of the third size-
741 bin (oin_a03) OIN's background error. As shown in Table S3, because the oin_a04's BEC
742 reduction relaxes the constraint on the coarse particle, the PM_{10} bias becomes more negative
743 along with the decrease in oin_a04's BEC. Meanwhile, the $PM_{2.5}$ bias becomes more positive.
744 Correspondingly, the ratio of $PM_{2.5}$ to PM_{10} was exaggerated to 0.33 with 30% of oin_a04's
745 BEC, higher than the observed value of 0.28. Overall, the original BEC of oin_a04 is a
746 reasonable tradeoff in our DA experiments. The inter-size bin correlation length tunes the

747 cross size-bin modifications, and it indeed does matter to the DA performance compared with
748 those without inter-size bin correlation. Although the correlation length of four in our DA
749 experiment is a little bit arbitrary, we found that the impact on the analysis due to using
750 different correlation length is almost ignorable.

751

752 The DA system preferentially modified the coarse particle concentrations because of the
753 coarse particles's high background model error according to our BEC modeling strategy.
754 Intuitively, our modification that mainly focused on the highest concentration of coarse
755 particles was reasonable. It decreased the model biases by raising the heaviest loading
756 aerosols. As a result, the ratio of PM_{2.5} to PM₁₀ decreased from 0.39 in the background to 0.27
757 in DA_PMx, approaching the observed ratio of 0.28. Such improvement was consistent with
758 the correction required to the model desert dust in literature. Kok et al. (2011) found that
759 regional and global circulation models underestimate the fraction of emitted coast dust (>~5
760 μm), overestimates the fraction of fine dust (<2μm diameter). Adebisi and Kok (2020)
761 claimed that too rapid deposition of coarse dust out of the atmosphere accounts for the
762 missing coarse dust in models. Similarly, WRF-Chem assimilated too much smaller dust
763 particles than the observed. According to Kashi's AOD between 440 nm and 1020 nm, the
764 observed Ångström exponent (AE) was 0.18 in this case, but the background value was 0.54
765 (Table 3). DA_PMx reduced the AE value to 0.30, a little improvement but not sufficient.

766

767 As the particle concentration increased, the 635 nm aerosol scattering coefficient in DA_PMx
768 moderately increased to 170.4 Mm⁻¹, with a high correlation of 0.89, still lower than the
769 observed level of 231.5 Mm⁻¹. The analyzed 660 nm absorption coefficient was 15.8 Mm⁻¹,
770 67% lower than observed levels, with a correlation of 0.42. The analyzed AOD showed a
771 monthly mean value of 0.38 in DA_PMx, 42% lower than observed levels, with a low
772 correlation of 0.35.

773

774 Figure 9a shows the diurnal concentrations of PM₁₀ in the analyses in April. The observed
775 PM₁₀ showed a substantial variation at 18:00 UTC, the (local midnight). This substantial
776 nocturnal variation was partly owing to the dust storm that started on 24 April and ended the
777 next day. This midnight variation was also related to a nocturnal low-level jet. Ge et al.
778 (2016) pointed out that there was a nocturnal low-level jet at a height of 100–400 m, with a
779 wind speed of 4–10 m s⁻¹ throughout the year in the Tarim Basin. They stressed that the low-
780 level jet broke down in the morning, transporting its momentum toward the surface, and
781 increased dust emissions. The nocturnal low-level jet increased the possibility of dust
782 particles moving towards the city at night, causing a high PM₁₀ variation at 18:00 UTC. The
783 diurnal changes in the DA analyses followed the observed levels, but had higher mean values.

784

785 **3.3 Assimilating AOD**

786 Assimilating AOD (DA_AOD) improved the monthly mean 870 nm AOD to 0.59,
787 approaching to the observed value of 0.66, with a high correlation of 0.98 (Figure 6u). The
788 monthly mean PM_{2.5} was improved to 92.6 μg m⁻³, quite close to the observed level of 91 μg
789 m⁻³, but the analyzed PM₁₀ was 541.7 μg m⁻³, 68% higher than the observed value. The DA
790 system improved the AOD at the price of deteriorating the data quality of surface coarse

791 particle concentrations. Surface particle overestimations have been reported in previous
792 studies (Liu et al., 2011; Ma et al., 2020; Saide et al., 2020). In the arid area of Kashi, the ratio
793 of PM_{2.5} to PM₁₀ therefore reduced to 0.17 in DA_AOD, which was too far comparing with
794 the observed ratio of 0.28.

795

796 The revised GSI updates aerosol number concentration according to the analyzed aerosol
797 mass concentration and the background ratio between mass and number concentrations. Thus,
798 an overestimation of aerosol mass concentration inclines to raise aerosol number
799 concentration, resulting in high scattering/absorption coefficients. In Kashi, the analyzed 635
800 nm scattering coefficient in DA_AOD was 222.6 Mm⁻¹, slightly lower than the observed
801 value. The analyzed 660 nm absorption coefficient was 17.0 Mm⁻¹, 64% lower than the
802 observed value. It indicates that WRF-Chem strongly underestimated the
803 scattering/absorption cross section. This underestimation resulted in too many coarse particles
804 as compensation to fit the observed AOD, and hence decreased the PM_{2.5}/PM₁₀ ratio further.

805

806 Table 4 shows the ratios of the AOD and aerosol scattering/absorption coefficients to the
807 surface PM₁₀ concentrations. The ratio of AOD to PM₁₀ in the background model result was
808 one-third of the observed levels. The observed mass scattering coefficient (Esca/PM₁₀) was
809 1.05 Mm⁻¹ μg⁻¹ m³, while the background value was only 0.65 Mm⁻¹ μg⁻¹ m³. DA_AOD
810 did not eliminate the low bias but enlarged the low bias to 0.51 Mm⁻¹ μg⁻¹ m³. The same
811 thing occurred for Eabs/PM₁₀, which was 0.09 in the background and 0.05 in DA_AOD,
812 much lower than the observed value of 0.25. Figure 10 shows these mean ratios at the other
813 wavelengths. The low bias in AOD/PM₁₀ was comparable at each wavelength. All DA
814 experiments yielded close bias in extinction/scattering/absorption efficiency. Such low bias in
815 AOD/PM₁₀ imposed the DA system to overestimate the PM₁₀ to fit the observed AOD data.

816

817 Additionally, we computed the surface single scattering albedo (SSAsrf) with the 525 nm
818 scattering coefficient and 520 nm absorption coefficient. We did not use the Ångström
819 exponent to interpolate the scattering/absorption coefficients to a similar wavelength because
820 the AE itself had a large model bias even after DA (Table 3). The observed SSAsrf value was
821 0.78, indicating an emphatic absorption particle, probably due to the mixture of anthropogenic
822 black carbon and natural desert dust in the local air. The model background SSAsrf was 0.86,
823 while the DA analyses gave even higher SSAsrf (0.88 to 0.9).

824

825 The low bias in mass scattering/absorption efficiency is related to the aerosol optical module,
826 which is based on Mie theory in WRF-Chem. First, the simulations used four-size bin particle
827 segregation. This coarse size representation aggregated many aerosols in the accumulation
828 mode (Figure 8f). Because small particles have a strong of light attenuation capability,
829 according to the Mie theory, too many coarse particles would not effectively increase the
830 AOD. Saide et al. (2020) linked the aerosol optics to the size bin representation (from 4 to 16
831 bins) for hazes in South Korea. They showed that WRF-Chem underestimated the dry aerosol
832 extinction, and the underestimation could be relieved when using a finer size bin than four.
833 Okada and Kai (2004) found that the dust particle radius in the Taklamakan Desert was in the

834 range of 0.1–4 μm , indicating the dominant fine-mode particles in the desert. Using the four-
835 size bin would simultaneously obtain better analyses of both AOD and PM_x.

836

837 Second, the dust particles are irregular in shape (Okada and Kai, 2004), while the spherical
838 particle is a common assumption for the aerosol optics in the Mie theory in current models,
839 which is an essential source of uncertainty in the forward operator of WRF-Chem when the
840 assumption of spherical particles for dust fails. The irregular morphology had a significant
841 influence on the dust simulation. Okada et al. (2001) found that the aspect ratio (the ratio of
842 the longest dimension to its orthogonal width) of the mineral dust particles (0.1–6 μm) in
843 China arid regions exhibited a median of 1.4. Dubovik et al. (2006) suggested the aspect ratio
844 of ~ 1.5 and higher in desert dust plumes. Kok et al. (2017) found that the dust' sphericity
845 assumption underestimated dust extinction efficiency by ~ 20 – 60% for the dust particle larger
846 than $1\mu\text{m}$. Tian et al. (2020) found that using a dust ellipsoid model could increase the
847 concentration of coarse dust particle (5–10 μm) by $\sim 5\%$ in eastern china and $\sim 10\%$ in the
848 Taklimakan area because of the decrease in gravitational settling, comparing with the
849 simulations with dust sphericity model. Nevertheless, the aspect ratio of the spheroid dust is
850 uncertain. Even after applying the spheroidal approximation, Soorbas et al. (2015) found that
851 the model underestimated 550 nm aerosol scattering and backscattering values by 49% and
852 11%, respectively, because of the uncertainties in particle axial ratio, complex refractive
853 index, and the particle size distribution. To date, the assumption of spherical particles has
854 been widespread in models (including WRF-Chem) for computational efficiency. Impact of
855 dust morphology to DA deserves a further investigation.

856

857 To reduce the overestimate in PM_x concentrations, we set the gridded standard deviation in
858 place of the latitude-binning standard deviation, as discussed in Section 2.3. Figure 11 shows
859 the analyzed vertical profiles of PM_x concentrations. Higher PM₁₀ concentrations were
860 observed in the low atmosphere than at the surface. These vertical error profiles decreased the
861 surface PM₁₀ particles and tended to increase the PM_{2.5}/PM₁₀ ratio, contrary to the effects of
862 low model bias in particle extinction efficiency. For the net effect of the compensation, the
863 mass extinction efficiency in the analysis was still almost equivalent to the background value
864 (Table 4). That is, our tuned BEC vertical profile at Kashi, to some extent canceled out the
865 effects of other model error sources (e.g., the positive bias in the coarse particle of BEC, and
866 the low bias in extinction efficiency) but was not sufficient to increase the mass extinction
867 efficiency to the observed value. Finer aerosol size representation and a better advanced
868 aerosol optical calculation for dust are essential solutions.

869

870 Assimilating the AOD seems to increase the diurnal variation in the DA analyses, but this
871 variation was not conclusive since there were different amounts of AOD data for DA at 00:00,
872 06:00, and 12:00. The AOD data were not always available as the data quality control (i.e.,
873 cloud screening). There was a higher increase in the concentration at noon (06:00 UTC)
874 (Figure 9b), corresponding to a few high AOD during mild dust episodes at that hour. The DA
875 system had to raise the PM₁₀ to fit the observed high AOD values. Because the CE318 AOD
876 was only available in the daytime, none DA analysis was performed at 18:00. Also, due to the
877 limited AOD data, assimilating AOD did not substantially increase the correlation of PM_x.

878 The analyzed PM_{2.5} and PM₁₀ still had low correlations with the observed levels
879 ($R=0.31\sim 0.35$).

880

881 **3.4 Assimilating Aerosol Scattering Coefficient**

882 Assimilating the aerosol scattering coefficient (DA_Esca) yielded overall analyses similar to
883 the phenomenon in DA_AOD. The analyzed 635 nm scattering coefficient (192.1 Mm^{-1}) was
884 lower than the observation (231.5 Mm^{-1}), with a high correlation of 0.97. The analyzed
885 monthly mean AOD was 0.53, better than the AOD of 0.38 when assimilating PM_x.
886 However, the surface particle concentrations were overestimated (i.e., positive biases by 14%
887 for PM_{2.5}, and 37% for PM₁₀), with a substantial increase in the coarse particle of OIN.
888 Overestimations appeared during a few mild dust episodes (Figure 7d). This again indicated
889 that WRF-Chem underestimated the dust scattering efficiency, in accordance with the low
890 bias in the ratio of the scattering coefficient to PM₁₀ ($0.52 \text{ Mm}^{-1} \mu\text{g}^{-1} \text{ m}^3$; Table 4). The DA
891 system thus overfitted the PM_x concentration to approach the observed scattering coefficient.
892 The diurnal PM₁₀ in the analysis was similar to the assimilation of PM_x, showing a maximum
893 improvement and a robust nocturnal variation at 18:00 UTC (Figure 9c). Assimilating the
894 scattering coefficient failed to improve the absorption coefficient. The monthly mean
895 absorption coefficient was 16.5 Mm^{-1} , 65% lower than the observed value.

896

897 **3.5 Assimilating Aerosol Absorption Coefficient**

898 In contrast to the above results, assimilating the absorption coefficient (DA_Eabs) degraded
899 all the analyses other than the absorption coefficient itself. The analyses showed substantial
900 daily variations, and strong positive biases appeared in the dust episodes (Figure 7). The
901 PM_{2.5} was overestimated by a factor of three, and the PM₁₀ was overestimated by a factor of
902 four. The increases occurred each hour (Figure 9d). Because of the constant ratio between
903 mass and number concentration, the particle number concentration increased. As a result, the
904 aerosol scattering coefficient was overfitted to 612.2 Mm^{-1} , higher than the observed levels by
905 a factor of three. The monthly mean AOD improbably rose up to 1.73. Nevertheless, the
906 absorption coefficient (40 Mm^{-1}) was improved to the observed level (47.4 Mm^{-1}).

907

908 Improving the absorption coefficient at the cost of PM₁₀ overestimation indicates the model
909 biases in the representation of the particle mixture and the other absorbing particles (e.g.,
910 black carbon, brown carbon and aged dust). With respect to the current model, this failure is
911 related to the aerosol absorption represented in WRF-Chem. The leading absorption aerosol in
912 WRF-Chem is BC. The BC particle in the second size ($0.156\text{--}0.625 \mu\text{m}$) had the maximum
913 absorption, according to Mie theory, and had the maximum DA modifications in the second-
914 size bin (Figure 8e). However, because the BC had a small background concentration, the BC
915 showed a small DA improvement ($<1.5 \mu\text{g m}^{-3}$) and had small effects on increasing the
916 particle absorption. Meanwhile, the coarse dust particle concentration was primarily
917 increased, but the dust particles did not have a strong absorption as BC. As a result, the model
918 lowered the ratio of the absorption coefficient of PM₁₀ by an order of magnitude (Table 4).
919 The lower mass absorption efficiency was comparable at each wavelength and was close to
920 the other DA experiment (Figure 10c). Because of the constraint of the observed absorption
921 coefficient, the DA system dramatically overestimated the particle concentrations and induced

922 too much higher aerosol scattering coefficient and AOD. Because the overestimation of the
923 scattering coefficient was higher than that of the absorption coefficient, DA_abs even gave
924 the strongest SSArf (0.9; Table 3) in all DA experiments, opposite to our expectation that the
925 assimilation of absorption coefficient should improve SSA.

926

927 To understand the DA_Eabs's failure, we performed a few trials by changing the imaginary
928 part of the dust refractive index on 1200UTC on April 9. The results are present in the
929 supplementary Table S4a and S4b. The trials show that a high imaginary part of the dust
930 refractive index decreases the aerosol absorption coefficient. This paradox is due to the BC's
931 reduction. Specifically, a high imaginary part increases the absorption efficiency of coarse
932 dust and decreases the coarse dust number concentration (num_a04; Table S4a). This
933 reduction also led to less fine aerosol number concentrations (e.g., num_a02) because of the
934 inter-size bin correlation. BC is abundant in the second and third size bins, and its imaginary
935 part of refractive index is two orders of magnitude higher than dust. Less BC caused a weak
936 absorption coefficient (Table S4b). On the contrary, the low dust imaginary part would not
937 largely increase dust numbers in the coarse size bin because the DA system attempts to
938 increase BC to enhance the absorption coefficient. In an extreme case with zero value of
939 imaginary part of dust, the improvement of absorption coefficient exclusively relies on BC;
940 the num_a02 is increased by order of magnitude (Table S4a), and 660 nm Eabs rose up to
941 92.5 Mm^{-1} (Table S4b), much higher than the observed level.

942

943 At Kashi, BC has a low background concentration and low background error. The innovation
944 of BC was limited. Thus, tuning the imaginary part of dust would not change the SSAsrf
945 value a lot (0.89 to 0.92). Excluding the contribution from OIN in PM_{10} , the scattering
946 coefficient was associated with sulfate. The sulfate's background error was higher than the
947 BC's by order of magnitude. The DA system prioritized sulfate modification even when
948 assimilating absorption coefficient, resulting in a smaller BC mass fraction in PM_{10} (Figure
949 12f) and a high SSAsrf of 0.90.

950

951 We did another set of trials by increasing the original BC's BEC per size bin. As shown in the
952 supplementary Table S5, increasing the BC's BECs would not much degrade the absorption
953 coefficient but significantly decrease the positive biases in PM_x , AOD, and scattering
954 coefficient; the SSAsrf approached the observation. Increasing BC's BECs by a factor of
955 seven (DA_Eabs_BC*7) shows the best analyses. This trial suppressed the positive biases
956 without decreasing the accuracy of absorption coefficient (Figure 7), and the BC mass
957 fraction was increased (Figure 12g). Nevertheless, the disadvantage of the enlargement of
958 BC's BEC is that the simultaneous assimilation of scattering and absorption coefficient is not
959 convergent as well as before. After four outer loops and each with 50 inner iterations, the
960 analyzed absorption coefficient in DA_Eabs_BC*7 was still higher than the observed value
961 by 47% (Figure S1j). It indicates there is a low bias in BC's background concentration that
962 violates the unbiased condition of DA.

963

964 **3.6 Assimilating Multi-source Observations**

965 Assimilating an individual observation improves the corresponding model parameter (i.e.,
966 PM_{2.5}, PM₁₀, Esca, Eabs, and AOD) but may worsen other parameters. The reasons for the
967 inconsistent improvements are relevant to the aerosol model itself. These are: (1) the model
968 parameters have opposite signs in biases (e.g., one model parameter has a positive bias while
969 another has a negative bias); (2) the model biases have vast differences in magnitude (e.g., a
970 good fit of a parameter may lead to another's overfit) and the different biases in magnitude
971 cannot be reconciled because the forward operator is inaccurate to represent the linkage
972 between aerosol mass and aerosol optics (e.g., lower particle mass extinction efficiency).

973
974 In our case, simultaneous assimilation of the scattering and absorption coefficients
975 (DA_Esca_Eabs) resulted in the analyses when assimilating the scattering coefficient alone
976 (DA_Esca), and the inferior analysis in DA_Eabs vanished. This was because incorporating
977 the scattering coefficient constrained the aerosol number concentrations, which also benefited
978 from incorporating the observed absorption coefficient. Simultaneous assimilation of PM_x
979 and AOD (DA_PM_x_AOD) gave the best overall DA results, in which all the analyses except
980 the absorption coefficient were not significantly different in the month mean values from the
981 observations. Furthermore, DA_PM_x_AOD substantially improved the Ångström exponent,
982 with an analyzed value of 0.17, consistent with the observed value of 0.18 (Table 3).
983 Simultaneous assimilation of all observations (DA_PM_x_Esca_Eabs_AOD) did not
984 substantially improve the analyses when compared with DA_PM_x_AOD because the surface
985 coefficients, and AOD had overlapped information of the light attenuation. A redundant
986 information source did not introduce extra constraints on the DA system.

987

Table 3, 4; Figure 6, 7, 8, 9, 10

988

989 3.7 Vertical Profiles of Aerosol Concentrations

990 Figure 11 shows the vertical concentration profiles of PM_{2.5} and PM₁₀. The DA system
991 increased the aerosol concentrations up to a height of 4 km, which is consistent with previous
992 studies on the Taklamakan Desert. Meng et al. (2019) simulated a deep dust layer thickness in
993 spring, with a depth of 3–5 km. Ge et al. (2014) analyzed the Cloud-Aerosol Lidar Orthogonal
994 Polarization data from 2006–2012 in the desert. They showed that dust could be lifted up to
995 5 km above the Tarim Basin, and even higher along the northern slope of the Tibetan Plateau.
996 Among our DA experiments, the analyzed PM_x in the lower atmosphere followed PM_x at the
997 surface. The vertical PM₁₀ concentration increased quickly in the lowest three model layers
998 and maintained high values at heights of less than 3 km. This vertical profile corresponded to
999 the background vertical error profile, reflecting the deep dust transporting layer. The PM_{2.5}
1000 vertical profiles showed a rapid reduction with an increase in altitude. The figure clearly
1001 shows that DA_PM_x improved the PM_{2.5} and PM₁₀ better, whereas DA_AOD preferentially
1002 adjusted the coarse particles and overestimated the PM₁₀. Also shown in the figure are the
1003 vertical profiles normalized to their own respective surface particulate concentrations. The
1004 assimilations not only added a larger fraction of the mass in these layers but also adjusted the
1005 shapes of the PM₁₀ profiles within 3 km above the ground (Figure 11d), following the BEC's
1006 vertical correlation length scales (Figure 3r).

1007

Figure 11

1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027

4. Discussions

4.1 DA Impact on Aerosol Chemical Composition

For control variable design, our DA system modifies the chemical composition of each aerosol according to the BEC values. The PM₁₀ chemical fractions remain close to their background values (Figure 12). As discussed in section 3.5, the assimilation of the aerosol absorption coefficient alone (DA_Eabs) increased the sulfate fraction. The DA modification increased aerosol number concentration, and the rising number concentration increased the tangent linear operator value for the scattering component. Sulfate was the predominant anthropogenic aerosol at Kashi and had a high background error value. The DA system prioritized the modification of sulfate and prevented a rise in the BC fraction in DA_Eabs. As the enlarged BC BEC in DA_Eabs_BC*7, the BC mass fraction showed the largest increase. The model bias in aerosol background concentration and the background error determine the analyzed aerosol chemical fraction. Overall, it seems that differences in aerosol chemical composition from assimilating the aerosol optical data are smaller than the difference in model setting (e.g., using other aerosol chemistry mechanisms, or using finer aerosol size bins). The assimilation of the total aerosol quantities cannot eliminate the intrinsic bias in aerosol composition. Thus, accurate aerosol chemistry and optical modules are crucial to attain a better background aerosol chemical data for DA analysis (Saide et al., 2020).

Figure 12

1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048

4.2 DA Impact on Aerosol Direct Radiative Forcing

Table 5 shows the instantaneous clear-sky ADRF in the background data and the analyses of DA_PMx and DA_PMx_AOD. After the analyses, the DA effect (various DA frequencies for assimilating AOD and the surface particle concentrations) gradually faded away after restarting the model run. We therefore focused on the instantaneous radiative forcing values one hour after assimilating AOD data. This ensured that the comparison was based on similar analysis times and showed effective DA effects.

Aerosol redistributes the energy between the land and the atmosphere. The atmosphere gains more shortwave energy as the dust and black carbon particle absorption; the warming atmosphere also emits more longwave energy as it absorbs shortwave energy. The change in energy budget at the surface is correspondingly the opposite of that in the atmosphere. As shown in Table 5, the enhancements in surface cooling forces were slightly stronger than the atmospheric warming forcings. The differences between the surface forcing and atmospheric forcing indicate the ADRF at the top of the atmosphere (TOA). The TOA ADRF when assimilating the surface particle concentrations was enhanced by 21% in the shortwave, 100% in the longwave, and 18% in the net forcing values, and enhanced by 34%, 67%, and 32%, respectively, when assimilating the AOD. Apparently, assimilating PMx alone is not sufficient to accurately estimate the ADRF value. At Kashi, the total net (shortwave plus longwave) clear-sky ADRF with assimilating surface particles and AOD were -10.4 Wm^{-2} at

1049 the TOA, $+20.8 \text{ Wm}^{-2}$ within the atmosphere, and -31.2 Wm^{-2} at the surface, respectively,
1050 enhanced by 55%, 48%, and 50% respectively, compared to the background ADRF values.

1051

1052 It is noteworthy to say that the ADRF estimation remains uncertain even after DA. The AOD
1053 observation is only sporadically available because of cloud screening in retrieval data. The
1054 DA experiments cannot eliminate the low bias in AOD in WRF-Chem. The ADRF values in
1055 the DA experiments are likely to be weaker than the plausible aerosol radiative forcing at
1056 Kashi. Neither DA experiment lowers SSAsrf to approach the observation. The observed
1057 SSAsrf (0.78) indicates likely warming forcing of aerosol at Kashi, while WRF-Chem and the
1058 DA analyses impose cooling forcing. The ADRF uncertainty is associated with the
1059 background aerosols. WRF-Chem simulates aerosol size up to $10 \mu\text{m}$, whereas larger particles
1060 ($>10 \mu\text{m}$) exhibit substantial absorption relative to scattering in the visible wavelength (Kok
1061 et al., 2017). Anthropogenic emission inventories need an update for the year 2019, which
1062 may reduce the potential low bias in BC concentration. Additionally, the revised GSI does not
1063 concern the change in particle effective radius per size bin when calculating the aerosol
1064 number concentration in each outer loop. Low absorption cross section rises aerosol number
1065 concentration as compensation, increasing aerosol scattering coefficient too much. If our
1066 tangent operator concerns the change in particle effective radius per size bin, we can use
1067 aerosol mass and number concentration as control variables simultaneously. The DA would
1068 have a higher degree of freedom to balance the particle radius and number concentration and
1069 improve the absorption coefficient. All these need further research in the future.

1070

1071 **5. Conclusions**

1072 This study described our revised GSI DA system for assimilating aerosol observed data for
1073 the four-size bin sectional MOSAIC aerosol mechanism in WRF-Chem. The DA system has
1074 new design tangent linear operators for the multi-wavelength AOD, aerosol scattering, and
1075 absorption coefficients measured by the sun-sky radiometer, nephelometer, and aethalometer,
1076 respectively. We examined the DA system for Kashi city in northwestern China by
1077 assimilating the multi-wavelength aerosol optical measurements gathered by the Dust Aerosol
1078 Observation–Kashi field campaign of April 2019 and the concurrent hourly measurements of
1079 surface $\text{PM}_{2.5}$ and PM_{10} concentrations.

1080

1081 Our DA system includes two main aspects. Firstly, the control variable is the aerosol chemical
1082 composition per size bin corresponding to the WRF-Chem output data. This design allows the
1083 modification of the composition of each aerosol, based on their background error covariances.
1084 The number of control variables could be reduced by intentionally excluding a few aerosol
1085 compositions in a specific case, if these compositions had low concentrations (e.g. chlorine
1086 and sodium in this study). Second, the DA system incorporates the observed AOD by
1087 assimilating the column mean aerosol extinction coefficient. This transfer avoids handling
1088 sensitivity from light attenuation length to the aerosol mass concentration in the tangent linear
1089 operator, which is difficult to accurately estimate and introduces significant errors in the
1090 operator. The tangent linear operator for AOD has two variants that can incorporate
1091 nephelometer and aethalometer measurements at the surface.

1092

1093 The most abundant aerosol at Kashi in April 2019 was dust. The WRF-Chem model captured
1094 the main dust episodes, but lowered the monthly mean concentrations of $PM_{2.5}$ and PM_{10} by
1095 17% and 41%, respectively. The model failed to capture the peak concentrations from a dust
1096 storm on 24 April. The aerosol scattering/absorption coefficients and AOD in the background
1097 data showed strong low biases and weak correlations with the observed levels. The DA
1098 systems effectively assimilate the surface particle concentrations, aerosol scattering
1099 coefficients, and AOD. Some deficiencies in the DA analysis were related to the forward
1100 model bias in transferring the aerosol mass concentrations to the aerosol optical parameter.
1101 Simultaneous assimilation of the $PM_{2.5}$ and PM_{10} concentrations improved the model aerosol
1102 concentrations, with significant increases in the coarse particles; meanwhile, the analyzed
1103 AOD was 42% lower than observed levels. The assimilation of AOD significantly improved
1104 the AOD but overestimated the surface PM_{10} concentration by 68%. Assimilating the aerosol
1105 scattering coefficient improved the scattering coefficient in the analysis but overestimated the
1106 surface PM_{10} concentration by 37%. It therefore seems that WRF-Chem underestimated the
1107 particle extinction efficiency. As a compensation, the DA system overestimated the aerosol
1108 concentration to fit the observed optical values, yielding overly high particle concentrations.
1109

1110 A notable problem was the assimilation of the absorption coefficient, which greatly
1111 overestimated the monthly mean values by a factor of four in PM_{10} . The aerosol absorption
1112 coefficient was improved but was still 16% lower than observed values. The failure of DA
1113 analysis when assimilating the absorption coefficient is associated with many factors,
1114 including the biases of the model in aerosol particle mixture and aged dust, the uncertainties
1115 in the imaginary part of dust refractive index, the uncertain background error of BC and the
1116 likely low bias in anthropogenic emissions. The most effective DA is the simultaneous
1117 assimilation of surface particle concentration and AOD, which provides the best overall DA
1118 analysis.
1119

1120 Our design of control variables allowed the DA system to adjust the aerosol chemical
1121 compositions individually. However, the analyzed anthropogenic aerosol chemical fractions
1122 were almost equivalent to the background chemical fractions. The reason is that the
1123 hydrophilic aerosols have equivalent or comparable refractive indices and hygroscopic
1124 parameters in the forward operator; they therefore have comparable adjoint operator values
1125 when assimilating the aerosol optical data. It may be possible to separate the chemical
1126 compositions based on their background errors. The model anthropogenic aerosols were low
1127 at Kashi, probably owing to the low biases in the anthropogenic emissions. The low
1128 background concentrations led to low background errors and hence few increments for all
1129 chemical compositions. As a result, the chemical fractions of the anthropogenic aerosols
1130 remained close to their background values.
1131

1132 When assimilating surface particles and AOD, the instantaneous clear-sky ADRF (shortwave
1133 plus longwave) at Kashi were -10.4 Wm^{-2} at the TOA, $+20.8 \text{ Wm}^{-2}$ within the atmosphere,
1134 and -31.2 Wm^{-2} at the surface, respectively. Since the DA analyses still lowered the AOD
1135 value and overestimated SSA, the aerosol radiative forcing values assimilating the
1136 observations were underestimated in the atmosphere and at the surface.

1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180

The limitations that necessitate further research include:

(1) The desired binning strategy should link the circulation flow and particle emission sources. A better hybrid DA coupled with the ensemble Kalman filter will be more effective for estimating the aerosol background error.

(2) The observational error could be elaborated further. The PM_{10} included the anthropogenic coarse particles, which should be separated from the dust originating from the desert (Jin et al., 2019). We set the observation errors for PM_x and AOD to the conventional values. The observational errors of the nephelometer and aethalometer were slightly arbitrary in this study, necessitating further consideration.

(3) The anthropogenic aerosols' background errors are needed to harmonize for better assimilation of the aerosol absorption coefficient or absorption AOD.

(4) The DA system was based on four-size bin MOSAIC aerosols, but it can be extended to work with eight-size bin MOSAIC aerosols in WRF-Chem. When assimilating aerosol optical data, the DA quality is strongly dependent on the forward model. The responses of our DA analysis to the bias and uncertainty in the forward aerosol optical model in WRF-Chem need further investigation.

Author contributions

WC developed the DA system, preformed the analyses and wrote the paper. ZL led the field campaign and revised the paper. YZ and KL implemented the observations and the data quality control. YZ helped to design the new adjoint operator. JC verified the DA system.

Competing interests

The authors declare that they have no conflict of interest.

Code/Data availability

The official GSI code is available at <https://dtcenter.org/community-code/gridpoint-statistical-interpolation-gsi/download>. The revised GSI code is available at https://github.com/wenyuan-chang/GSI_WRF-Chem_MOSAIC. The aerosol measurements at Kashi belong to the Sun-sky radiometer Observation NETwork (SONET) which is accessible at <http://www.sonet.ac.cn/en/index.php>.

Acknowledgments

This work is supported by the National Key Research and Development Program of China (Grant number 2016YFE0201400).

References

Adebisi, A. A., Kok, J. F.: Climate models miss most of the coarse dust in the atmosphere, *Sci. Adv.*, 6, 1-10, doi:10.1126/sciadv.aaz9507, 2020.

Bannister, R. N.: A review of operational methods of variational and ensemble-variational data assimilation, *Q.J.R. Meteorol. Soc.*, 143, 607-633, doi:10.1002/qj.2982, 2017.

1181 Bao, Y., Zhu, L., Guan, Q., Guan, Y., Lu, Q., Petropoulos, G. P., Che H., Ali, g., Dong, Y.,
1182 Tang, Z., Gu, Y., Tang, W., and Hou, Y.: Assessing the impact of Chinese FY-3/MERSI
1183 AOD data assimilation on air quality forecasts: sand dust events in northeast China, *Atmos.*
1184 *Environ.*, 205, 78-89, doi:10.1016/j.atmosenv.2019.02.026, 2019.

1185

1186 Barnard, J. C., Fast, J. D., Paredes-Miranda, G., Arnott, W. P., and Laskin, A.: Technical
1187 Note: Evaluation of the WRF-Chem “Aerosol Chemical to Aerosol Optical Properties”
1188 Module using data from the MILAGRO campaign, *Atmos. Chem. Phys.*, 19, 7325-7340,
1189 doi:10.5194/acp-10-7325-2010, 2010.

1190

1191 Bond, T. C., and Bergstrom, R. W.: Light absorption by carbonaceous particles: an
1192 investigative review, *Aerosol Sci. Tech.*, 40, 27-67, doi:10.1080/02786820500421521, 2006.

1193

1194 Calil, P. H. R., Doney, S. C., Yumimoto, K., Eguchi, K., and Takemura, T.: Episodic
1195 upwelling and dust deposition as bloom triggers in low-nutrient, low-chlorophyll regions, *J.*
1196 *Geophys. Res.*, 116, C06030, doi:10.1029/2010jc006704, 2011.

1197

1198 Chen, Y., and Bond, T. C.: Light absorption by organic carbon from wood combustion,
1199 *Atmos. Chem. Phys.*, 10, 1773-1787, doi:10.5194/acp-10-1773-2010, 2010.

1200

1201 Chen, D., Liu, Z., Schwartz, C. S., Lin, H.-C., Cetola, J. D., Gu, Y., and Xue, L.: The impact
1202 of aerosol optical depth assimilation on aerosol forecasts and radiative effects during a wild
1203 fire event over the United States, *Geosci. Model Dev.*, 7, 2709-2715, doi:10.5194/gmd-7-
1204 2709-2014, 2014.

1205

1206 Chen, L., Gao, Y., Zhang, M., Fu, J. S., Zhu, J., Liao, H., Li, J., Huang, K., Ge, B., Wang, X.,
1207 Lam, Y. F., Lin, C.-Y., Itahashi, S., Nagashima, T., Kajino, M., Yamaji, K., Wang, Z., and
1208 Kurokawa, J.-i.: MICS-Asia III: multi-model comparison and evaluation of aerosol over East
1209 Asia, *Atmos. Chem. Phys.*, 19, 11911-11937, doi:10.5194/acp-19-11911-2019, 2019.

1210

1211 Chen, S.-H., and Sun, W.-Y.: A one-dimensional time dependent cloud model, *J. Meteor. Soc.*
1212 *Japan*, 80, 1, 99-118, doi:10.2151/jmsj.80.99, 2002.

1213

1214 Chen, D., Liu, Z., Ban, J., Zhao, P., and Chen, M.: Retrospective analysis of 2015-2017
1215 wintertime PM_{2.5} in China: response to emission regulations and the role of meteorology,
1216 *Atmos. Chem. Phys.*, 19, 7409-7427, doi:10.5194/acp-19-7409-2019, 2019.

1217

1218 Cheng, T., Wang, H., Xu, Y., Li, H., and Tian, L.: Climatology of aerosol optical properties in
1219 northern China, *Atmos. Environ.*, 40, 1495-1509, doi:10.1016/j.atmosenv.2005.10.047, 2006.

1220

1221 Cheng, X., Liu, Y., Xu, X., You, W., Zang, Z., Gao, L., Chen, Y., Su, D., and Yan, P.: Lidar
1222 data assimilation method based on CRTM and WRF-Chem models and its application in
1223 PM_{2.5} forecasts in Beijing, *Sci. Total Environ.*, 682, 541-552,
1224 doi:10.1016/j.scitotenv.2019.05.186, 2019.

1225
1226 Descombes, G., Auligné, T., Vandenberghe, F., Barker, D. M., and Barré, J.: Generalized
1227 background error covariance matrix model (GEN_BE v2.0), *Geosci. Model Dev.*, 8, 669-696,
1228 doi:10.5194/gdm-8-669-2015, 2015.
1229
1230 Di Biagio, C., Formenti, P., Balkanski, Y., Caponi, L., Cazaunau, M., Pangui, E., Journet, E.,
1231 Nowak, S., Andreae, M. O., Kandler, K., Saeed, T., Piketh, S., Seibert, D., Williams, E., and
1232 Doussin, J.-F.: Complex refractive indices and single-scattering albedo of global dust aerosols
1233 in the shortwave spectrum and relationship to size and iron content, *Atmos. Chem. Phys.*, 19,
1234 15503-15531, doi:10.5194/acp-19-15503-2019, 2019.
1235
1236 Dubovik, O., Sinyuk, A., Lapyonok, T., Holben, B. N., Mishchenko, M., Yang, P., Eck, T. F.,
1237 Volten, H., Munoz, O., Veihelmann, B., van der Zande, W. J., Leon, J.-F., Sorokin, M., and
1238 Slutsker, I.: Application of spheroid models to account for aerosol particle nonsphericity in
1239 remote sensing of desert dust, *J. Geophys. Res.*, 111, doi:10.1029/2005JD006619, 2006.
1240
1241 Fan, J., Shang, Y., Chen, Q., Wang, S., Zhang, X., Zhang, L., Zhang, Y., Xu, X., and Jiang,
1242 P.: Investigation of the “dust reservoir effect” of the Tarim Basin using WRF-GOCART
1243 model, *Arab. J. Geosci.*, 13, 214, doi:10.1007/s12517-020-5154-x, 2020.
1244
1245 Fast, J. D., Gustafson, W. I., Easter, R. C., Zaveri, R. A., Barnard, J. C., Chapman, E. G.,
1246 Grell, G. A. Peckham, S. E.: Evolution of ozone, particulates, and aerosol direct radiative
1247 forcing in the vicinity of Houston using a fully coupled meteorology-chemistry-aerosol
1248 model, *J. Geophys. Res.-Atmos.*, 111, D21305, doi:10.1029/2005jd006721, 2006.
1249
1250 Feng, S., Jiang, F., Jiang, Z., Wang, H., Cai, Z., and Zhang, L.: Impact of 3DVAR
1251 assimilation of surface PM_{2.5} observations on PM_{2.5} forecasts over China during wintertime,
1252 *Atmos. Environ.*, 187, 34-49, doi:10.1016/j.atmosenv.2018.05.049, 2018.
1253
1254 Ge, J. M., Huang, J. P., Xu, C. P., Qi, Y. L., and Liu, H. Y.: Characteristics of Taklimakan
1255 dust emission and distribution: A satellite and reanalysis field perspective, *J Geophys. Res.*
1256 *Atmos.*, 119, 11772-11783, doi:10.1002/2014jd022280, 2014.
1257
1258 Ge, J. M., Liu, H., Huang, J., and Fu, Q.: Taklimakan desert nocturnal low-level jet:
1259 climatology and dust activity, *Atmos. Chem. Phys.*, 16, 7773-7783, doi:10.5194/acp-16-7773-
1260 2016, 2016.
1261
1262 Ginoux, P., Chin, M., Tegen, I., Prospero, J. M., Holben, B., Dubovik, O., and Lin, S.-J.:
1263 Sources and distributions of dust aerosols simulated with the GOCART model, *J. Geophys.*
1264 *Res.*, 106, D17, 20255-20273, doi: 10.1029/2000JD000053, 2001.
1265
1266 Giorgi, F., and Mearns, L. O.: Introduction to special section: Regional climate modeling
1267 revisited, *J. Geophys. Res.*, 104, D6, 6335-6352, doi:10.1029/98JD02072, 1999.
1268

1269 Grell, G. A., Peckham, S. E., Schmitz, R., McKeen, S. A., Frost, G., Skamarock, W. C., and
1270 Eder, B.: Fully coupled “online” chemistry within the WRF model. *Atmos. Environ.*, 39,
1271 6957-7975, doi:10.1016/j.atmosenv.2005.04.027, 2005.
1272
1273 Guenther, A., Karl, T., Harley, P., Wiedinmyer, C., Palmer, P. I., and Geron, C.: Estimates of
1274 global terrestrial isoprene emissions using MEGAN (Model of Emissions of Gases and
1275 Aerosols from Nature), *Atmos. Chem. Phys.*, 6, 3181-3210, doi:10.5194/acp-6-3181-2006,
1276 2006.
1277
1278 He, J., Zhang, Y., Wang, K., Chen, Y., Leung, L. R., Fan, J., Li, M., Zheng, B., Zhang, Q.,
1279 Duan, F., and He, K.: Multi-year application of WRF-CAM5 over East Asia-Part I:
1280 Comprehensive evaluation and formation regimes of O₃ and PM_{2.5}, *Atmos. Environ.*, 165,
1281 122-142, doi:10.1016/j.atmosenv.2017.06.015, 2017.
1282
1283 Holben, B. N., Eck, T. F., Slutsker, I., Tanré D., Buis, J. P., Setzer, A., Vermote, E., Reagan,
1284 J. A., Kaufman, Y. J., Nakajima, T., Lavenu, F., Jankowiak, I., and Smirnov, A.: AERONET–
1285 A federated instrument network and data archive for aerosol characterization, *Remote Sens.*
1286 *Environ.*, 66, 1-16, doi:10.1016/s0034-4257(98)00031-5, 1998.
1287
1288 Hong, S.-Y., Noh, Y., Dudhia, J.: A new vertical diffusion package with an explicit treatment
1289 of entrainment processes, *Mon. Wea. Res.*, 134, 2318-2341, doi:10.1175/MWR3199.1, 2006.
1290
1291 Huang, J., Wang, T., Wang, W., Li, Z., and Yan, H.: Climate effects of dust aerosols over
1292 East Asian arid and semiarid regions, *J. Geophys. Res.*, 119, 11398-11416,
1293 doi:10.1002/2014jd021796, 2014.
1294
1295 Hong, J., Mao, F., Min, Q., Pan, Z., Wang, W., Zhang, T., and Gong, W.: Improved PM_{2.5}
1296 predictions of WRF-Chem via the integration of Himawari-8 satellite data and ground
1297 observations, *Environ. Pollut.*, 263, 114451, doi:10.1016/j.envpol.2020.114451, 2020.
1298
1299 Iacono, M. J., Delamere, J. S., Mlawer, E. J., Shephard, M. W., Clough, S. A., and Collins,
1300 W. D.: Radiative forcing by long-lived greenhouse gases: Calculations with the AER radiative
1301 transfer models, *J. Geophys. Res.*, 113, D13103, doi:10.1029/2008JD009944, 2008.
1302
1303 Jia, R., Liu, Y., Chen, B., Zhang, Z., and Huang, J.: Source and transportation of summer dust
1304 over the Tibetan Plateau, *Atmos. Environ.*, 123, 210-219,
1305 doi:10.1016/j.atmosenv.2015.10.038, 2015.
1306
1307 Jiang, Z., Liu, Z., Wang, T., Schwartz, C. S., Lin, H.-C., and Jiang, F.: Probing into the
1308 impact of 3DVAR assimilation of surface PM₁₀ observations over China using process
1309 analysis, *J. Geophys. Res.*, 118, 6738-6749, doi:10.1002/jgrd.50495, 2013.
1310

1311 Jin, J., Lin, H. X., Segers, A., Xie, Y., and Heemink, A.: Machine learning for observation
1312 bias correction with application to dust storm data assimilation, *Atmos. Chem. Phys.*, 19,
1313 10009-10026, doi:10.5194/acp-19-10009-2019.

1314

1315 Kok, J. F.: A scaling theory for the size distribution of emitted dust aerosols suggests climate
1316 models underestimate the size of the global dust cycle, *Proc. Natl. Acad. Sci. U.S.A.*, 108,
1317 1016-1021, doi:10.1073/pnas.1014798108, 2011.

1318

1319 Kok, J. F., Ridley, D. A., Zhou, Q., Miller, R. L., Zhao, C., Heald, C. L., Ward, D. S., Albani,
1320 S., and Haustein, K.: Smaller desert dust cooling effect estimated from analysis of dust size
1321 and abundance, *Nat. Geosci.*, doi:10.1038/ngeo2912, 2017.

1322

1323 Kumar, R., Monache, L. D., Bresch, J., Saide, P. E., Tang, Y., Liu, Z., de Silva, A. M.,
1324 Alessandrini, S., Pfister, G., Edwards, D., Lee, P., and Djalaova, I.: Toward improving short-
1325 term predictions of fine particulate matter over the United States via assimilation of satellite
1326 aerosol optical depth retrievals, *J. Geophys. Res.*, 124, 2753-2773,
1327 doi:10.1029/2018jd029009, 2019.

1328

1329 Li, L., Li, Z., Chang, W., Ou, Y., Goloub, P., Li, C., Li, K., Hu, Q., Wang, J., and Wendisch,
1330 M.: Solar radiative forcing of aerosol particles near the Taklimakan desert during the Dust
1331 Aerosol Observation-Kashi campaign in Spring 2019, *Atmos. Chem. Phys.*, 20, 10845-10864,
1332 doi:10.5194/acp-2020-60, 2020.

1333

1334 Li, M., Liu, H., Geng, G., Hong, C., Liu, F., Song, Y., Tong, D., Zheng, B., Cui, H., Man, H.,
1335 Zhang, Q., and He, K.: Anthropogenic emission inventories in China: a review, *Natl. Sci.*
1336 *Rev.*, 4, 834-866, doi:10.1093/nsr/nwx150, 2017.

1337

1338 Li, Z., Zang, Z., Li, Q. B., Chao, Y., Chen, D., Ye, Z., Liu, Y., and Liou, K. N.: A three-
1339 dimensional variational data assimilation system for multiple aerosol species with
1340 WRF/Chem and an application to PM_{2.5} prediction, *Atmos. Chem. Phys.*, 13, 4265-4278,
1341 doi:10.5194/acp-13-4265-2013, 2013.

1342

1343 Li, Z. Q., Xu, H., Li, K. T., Li, D. H., Xie, Y. S., Li, L., Zhang, Y., Gu, X. F., Zhao, W., Tian,
1344 Q. J., Deng, R. R., Su, X. L., Huang, B., Qiao, Y. L., Cui, W. Y., Hu, Y., Gong, C. L., Wang,
1345 Y. Q., Wang, X. F., Wang, J. P., Du, W. B., Pan, Z. Q., Li, Z. Z., and Bu, D.: Comprehensive
1346 study of optical, physical, chemical, and radiative properties of total columnar atmospheric
1347 aerosols over China An overview of sun-sky radiometer observation network (SONET)
1348 measurements, *Bull. Amer. Meteorol. Soc.*, 739-755, doi:10.1175/BMAS-D-17-0133.1, 2018.

1349

1350 Liu, Z., Liu, Q., Lin, H.-C., Schwartz, C. S., Lee, Y.-H., and Wang, T.: Three-dimensional
1351 variational assimilation of MODIS aerosol optical depth: Implementation and application to a
1352 dust storm over East Asia, *J. Geophys. Res.*, 116, D23206, doi:10.1029/2011JD016159, 2011.

1353

1354 Ma, C., Wang, T., Mizzi, A. P., Anderson, J. L., Zhuang, B., Xie, M., and Wu, R.:
1355 Multiconstituent data assimilation with WRF-Chem/DART: Potential for adjusting
1356 anthropogenic emissions and improving air quality forecasts over eastern China, *J. Geophys.*
1357 *Res. Atmos.*, 124, 7393-7412, doi:10.1029/2019JD030421, 2019.
1358

1359 Ma, C., Wang, T., Jiang, Z., Wu, H., Zhao, M., Zhuang, B., Li, S., Xie, M., Li, M., Liu, J.,
1360 and Wu, R.: Importance of bias correction in data assimilation of multiple observations over
1361 eastern China using WRF-Chem/DART, *J. Geophys. Res. Atmos.*, 125, e2019JD031465,
1362 doi:10.1029/2019JD031465, 2020.
1363

1364 Malm, W., C., and Hand, J. L.: An examination of the physical and optical properties of
1365 aerosols collected in the IMPROVE program, *Atmos. Environ.*, 41, 16, 3407-3427,
1366 doi:10.1016/j.atmosenv.2006.12.012, 2007.
1367

1368 Massart, S., Pajot, B., Piacentini, A., and Pannekoucke, O.: On the merits of using a 3D-
1369 FGAT assimilation scheme with an outer loop for atmospheric situations governed by
1370 transport, *Mon. Weather Rev.*, 138, 12, 4509-4522, doi:10.1175/2010MWR3237.1, 2010.
1371

1372 Meng, L, Yang, X., Zhao, T., He, Q., Lu, H., Mamtimin, A., Huo, W., Yang, F., and Liu, C.:
1373 Modeling study on three-dimensional distribution of dust aerosols during a dust storm over
1374 the Tarim Basin, Northwest China, *Atmos. Res.*, 218, 285-295,
1375 doi:10.1016/j.atmosres.2018.12.006, 2019.
1376

1377 Okada, K., Heintzenberg, J., Kai, K., and Qin, Y.: Shape of atmospheric mineral particles
1378 collected in three Chinese arid-regions, *Geophys. Res. Lett.*, 28, 16, 3123-3126,
1379 doi:10.1029/2000GL012798, 2001.
1380

1381 Okada, K., and Kai, K.: Atmospheric mineral particles collected at Qira in the Taklamakan
1382 Desert, China, *Atmos. Environ.*, 38, 6927-6935, doi:10.1016/j.atmosenv.2004.03.078, 2004.
1383

1384 Pagowski, M., Grell, G. A., McKeen, S. A., Peckham, S. E., and Devenyi, D.: Three-
1385 dimensional variational data assimilation of ozone and fine particulate matter observations:
1386 some results using the Weather Research and Forecasting – Chemistry model and Grid-point
1387 Statistical Interpolation, *Q. J. R., Meteorol. Soc.*, 136, 2014-2024, doi:10.1002/qj.700, 2010.
1388

1389 Pagowski, M., Liu, Z., Grell, G. A., Hu, M., Lin, H.-C., and Schwartz, C. S.: Implementation
1390 of aerosol assimilation in Gridpoint Statistical Interpolation (v.3.2) and WRF-Chem (v.3.4.1),
1391 *Geosci. Model Dev.*, 7, 1621-1627, doi:10.5194/gmd-7-1621-2014, 2014.
1392

1393 Pang, J., Liu, Z., Wang, X., Bresch, J., Ban, J., Chen, D., and Kim, J.: Assimilating AOD
1394 retrievals from GOCI and VIIRS to forecast surface PM_{2.5} episodes over eastern China,
1395 *Atmos. Environ.*, 179, 288-304, doi:10.1016/j.atmosenv.2018.02.011, 2018.
1396

1397 Pang, J., Wang, X., Shao, M., Chen, W, and Chang, M.: Aerosol optical depth assimilation for
1398 a modal aerosol model: Implementation and application in AOD forecasts over East Asia, *Sci.*
1399 *Total Environ.*, 719, 137430, doi:10.1016/j.scitotenv.2020.137430, 2020.
1400
1401 Parrish, D. F., and Derber, J. C.: The National Meteorological Center's spectral statistical-
1402 interpolation analysis system, *Mon. Weaterh Rev.*, 120, 1747-1763, doi: 10.1175/1520-
1403 0493(1992)120<1747:TNMCSS>2.0.CO;2, 1992.
1404
1405 Peng, Z., Liu, Z., Chen, D., and Ban, J.: Improving PM_{2.5} forecast over China by the joint
1406 adjustment of initial conditions and source emissions with an ensemble Kalman filter, *Atmos.*
1407 *Chem. Phys.*, 17, 4837-4855, doi:10.5194/acp-17-4837-2017, 2017.
1408
1409 Peng, Z., Lei, L., Liu, Z., Sun, J., Ding, A., Ban, J., Chen, D., Kou, X., and Chu, K.: The
1410 impact of multi-species surface chemical observation assimilation on air quality forecasts in
1411 China, *Atmos. Chem. Phys.*, 18, 17387-17404, doi:10.5194/acp-18-17387-2018, 2018.
1412
1413 Purser, R. J., Wu, W.-S., Parrish, D. F., and Roberts, N. M.: Numerical aspects of the
1414 application of recursive filters to variational statistical analysis. Part I: spatially homogeneous
1415 and isotropic gaussian covariances, *Mon. Weather Rev.*, 131, 1524-1535, doi: 10.1175//1520-
1416 0493(2003)131<1524:NAOTAO>2.0.CO;2, 2003a.
1417
1418 Purser, R. J., Wu, W.-S., Parrish, D. F., and Roberts, N. M.: Numerical aspects of the
1419 application of recursive filters to variational statistical analysis. Part II: spatially
1420 inhomogeneous and anisotropic general covariances, *Mon. Weather Rev.*, 131, 1536-1548,
1421 doi: 10.1175//2543.1, 2003b.
1422
1423 Saide, P. E., Carmichael, G. R., Liu, Z., Schwartz, C. S., Lin, H. C., da Silva, A. M., and
1424 Hyer, E.: Aerosol optical depth assimilation for a size-resolved sectional model: impacts of
1425 observationally constrained, multi-wavelength and fine mode retrievals on regional scale
1426 analyses and forecasts, *Atmos. Chem. Phys.*, 13, 10425-10444, doi:10.5194/acp-13-10425-
1427 2013, 2013.
1428
1429 Saide, P. E., Kim, J., Song, C. H., Choi, M., Cheng, Y., and Carmichael, G. R.: Assimilation
1430 of next generation geostationary aerosol optical depth retrievals to improve air quality
1431 simulations, *Geophys. Res. Lett.*, 41, 9188-9196, doi:10.1002/2014GL062089, 2014.
1432
1433 Saide, P. E., Gao, M., Lu, Z., Goldberg, D., Streets, D. G., Woo, J.-H., Beyersdorf, A., Corr,
1434 C. A., Thornhill, K. L., Anderson, B., Hair, J. W., Nehrir, A. R., Diskin, G. S., Jimenez, J. L.,
1435 Nault, B. A., Campuzano-Jost, P., Dibb, J., Heim, E., Lamb, K. D., Schwarz, J. P., Perring,
1436 A. E., Kim, J., Choi, M., Holben, B., Pfister, G., Hodzic, A., Carmichael, G. R., Emmons, L.,
1437 and Crawford, J. H.: Understanding and improving model representation of aerosol optical
1438 properties for a Chinese haze event measured during KORUS-AQ, *Atmos. Chem. Phys.*, 20,
1439 6455-6478, doi:10.5194/acp-20-6455-2020, 2020.
1440

1441 Schwartz, C. S., Liu, Z., Lin, H.-C., and McKeen, S. A.: Simultaneous three-dimensional
1442 variational assimilation of surface fine particulate matter and MODIS aerosol optical depth, *J.*
1443 *Geophys. Res.*, 117, D13202, doi:10.1029/2011JD017383, 2012.
1444

1445 Sič, B., Amraoui, L. E., Piacentini, A., Marécal, V., Emili, E., Cariolle, D., Prather, M., and
1446 Attié, J.-L.: Aerosol data assimilation in the chemical transport model MOCAGE during the
1447 TRAQA/ChArMEx campaign: aerosol optical depth, *Atmos. Chem. Phys.*, 9, 5535-5554,
1448 doi:10.5194/amt-9-5535-2016, 2016.
1449

1450 Sorribas, M., Olmo, F. J., Quirantes, A., Lyamani, H., Gil-Ojeda, M., Alados-Arboledas, L.,
1451 and Horvath, H.: Role of spheroidal particles in closure studies for aerosol microphysical-
1452 optical properties, *Q. J. R. Meteorol. Soc.*, 141, 2700-2707, doi:10.1002/qj.2557, 2015.
1453

1454 Tang, Y., Pagowski, M., Chai, T., Pan, L., Lee, P., Baker, B., Kumar, R., Monache, L. D.,
1455 Tong, D., and Kim, H.-C.: A case study of aerosol data assimilation with the Community
1456 Multi-scale Air Quality Model over the contiguous United States using 3D-Var and optimal
1457 interpolation methods, *Geosci. Model Dev.*, 10, 4743-4758, doi:10.5194/gmd-10-4743-2017,
1458 2017.
1459

1460 Tewari, M., Chen, F., Wang, W., Dudhai, J., LeMone, M. A., Mitchell, K., Ek, M., Gayno,
1461 G., Wegiel, J., and Cuenca, R. H.: Implementation and verification of the unified NOAA land
1462 surface model in the WRF model. 20th conference on weather analysis and forecasting/16th
1463 conference on numerical weather prediction, pp. 11-15, 2004.
1464

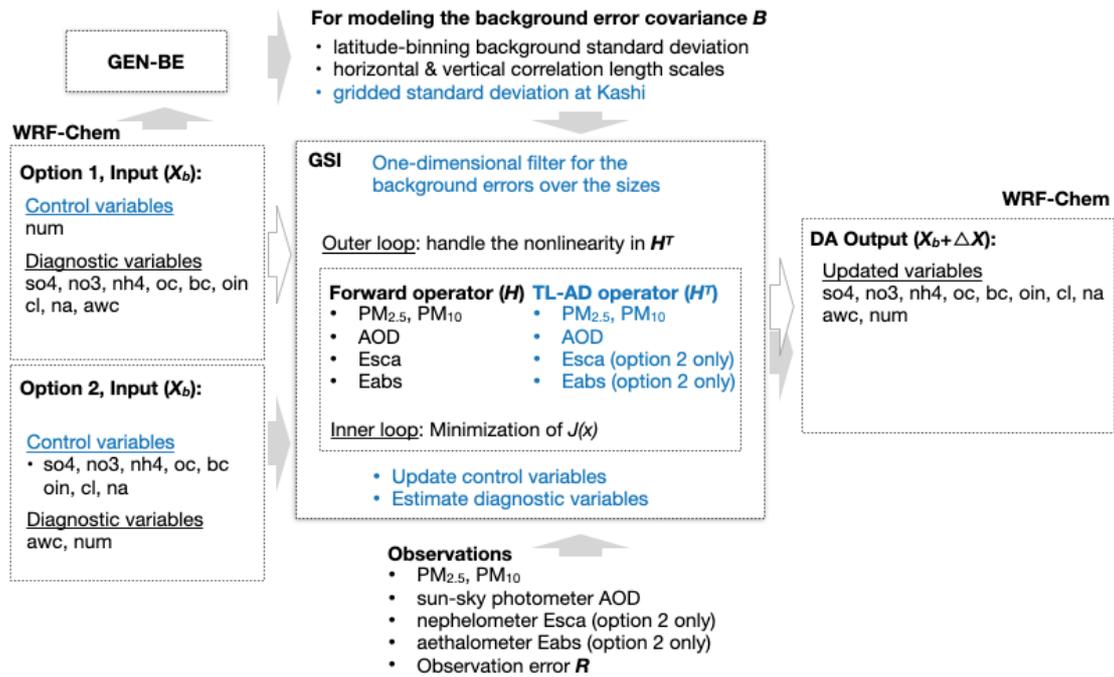
1465 Tian, Y., Wang, Z., Pan, X., Li, J., Yang, T., Wang, D., Liu, X., Liu, H., Zhang, Y., Lei, S.,
1466 Sun, Y., Fu, P., Uno, I., and Wang, Z.: Influence of the morphological change in natural
1467 Asian dust during transport: A modeling study for a typical dust event over northern China,
1468 *Sci. Total Environ.*, 10.1016/j.scitotenv.2020.139791, 2020.
1469

1470 Toon, O. B., Pollack, J. B., and Khare, B. N.: The optical constants of several atmospheric
1471 aerosol species: ammonium sulfate, aluminum oxide, and sodium chloride, *J. Geophys. Res.*,
1472 81, 33, doi:10.1029/JC081i033p05733, 1976.
1473

1474 Wang, K.-Y., Lary, D. J., Shallcross D. E., Hall, S. M., and Pyle, J. A.: A review on the use of
1475 the adjoint method in four-dimensional atmospheric-chemistry data assimilation, *Q. J. R.*
1476 *Meteorol. Soc.*, 127, 2181-2204, doi:10.1002/qj.49712757616, 2001.
1477

1478 Wang, D., You, W., Zang, Z., Pan, X., He, H., and Liang, Y.: A three-dimensional variational
1479 data assimilation system for a size-resolved aerosol model: Implementation and application
1480 for particulate matter and gaseous pollutant forecasts across China, *Sci. China-Earth Sci.*, 63,
1481 1366-1380, doi:10.1007/s11430-019-9601-4, 2020.
1482

1483 Wu, W.-S., Purser, r. J., and Parrish, D. F.: three-dimensional variational analysis with
1484 spatially inhomogeneous covariances, *Mon. Weather Rev.*, 130, 12, 2905-2916,
1485 doi: 10.1175/1520-0493(2002)130<2905:TDVAWS>2.0.CO;2, 2002.
1486
1487 Xia, X., Min, J., Shen, F., Wang, Y., and Yang, C.: Aerosol data assimilation using data from
1488 Fengyun-3A and MODIS: application to a dust storm over East Asia in 2011, *Adv. Atmos.*
1489 *Sci.*, 36, 1-14, doi:10.1007/s00376-018-8075-9, 2019a.
1490
1491 Xia, X., Min, J., Wang, Y., Shen, F., Yang, C., and Sun, Z.: Assimilating Himawari-8 AHI
1492 aerosol observations with a rapid-update data assimilation system, *Atmos. Environ.*, 215,
1493 116866, doi:10.1016/j.atmosenv.2019.116866, 2019b.
1494
1495 Zang, Z., Li, Z., Pan, X., Hao, Z., and You, W.: Aerosol data assimilation and forecasting
1496 experiments using aircraft and surface observations during CalNex, *Tellus B.*, 68,1, 29812,
1497 doi:10.3402/tellusb.v68.29812, 2016.
1498
1499 Zaveri, R. A., and Peters, L. K.: A new lumped structure photochemical mechanism for large-
1500 scale applications, *J. Geophys. Res.*, 104, 30387-30415, doi:10.1029/1999JD900876, 1999.
1501
1502 Zaveri, R. A., Easter, R. C., Fast, J. D., Peters, L. K.: Model for simulating aerosol
1503 interactions and chemistry (MOSAIC), *J. Geophys. Res.*, 113, D13204,
1504 doi:10.1029/2007JD008782, 2008.
1505
1506 Zhao, C., Liu, X., Leung, L. R., Johnson, B., McFarlane, S. A., Gustafson Jr., W. I., Fast, J.
1507 D., and Easter, R.: The spatial distribution of mineral dust and its shortwave radiative forcing
1508 over North Africa: modeling sensitivities to dust emissions and aerosol size treatments,
1509 *Atmos. Chem. Phys.*, 10, 8821-8838, doi:10.5194/acp-10-8821-2010, 2010.
1510
1511 Zhao, J., Ma, X., Wu, S., and Sha, T.: Dust emission and transport in Northwest China: WRF-
1512 Chem simulation and comparisons with multi-sensor observation, *Atmos. Res.*, 241, 104978,
1513 doi:10.1016/j.atmosres.2020.104978, 2020.



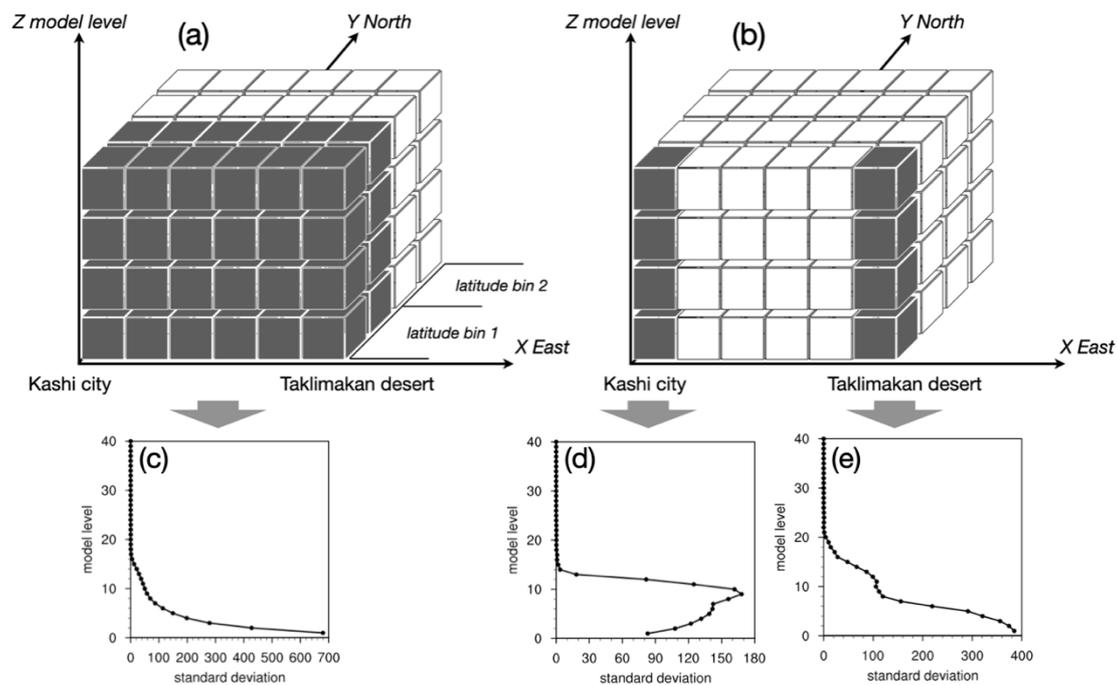
1514

1515

1516 Figure 1. The workflow of aerosol DA in the revised GSI system for the sectional
1517 MOSAIC aerosols in WRF-Chem. The contents in blue are the portions we
1518 developed. The arrows in gray indicate the workflow of option 2 that we did in this
1519 study. Only option 2 can assimilate the aerosol scattering/absorption coefficients.

1520 Abbreviations: so4, sulfate; nh4, ammonium; oc, organic carbon; bc, black carbon;
1521 oin, other inorganic matter; awc, aerosol water content; num, aerosol number
1522 concentration; no3, nitrate; cl, chlorine; na, sodium; Esca, aerosol scattering
1523 coefficient; Eabs, aerosol absorption coefficient.

1524

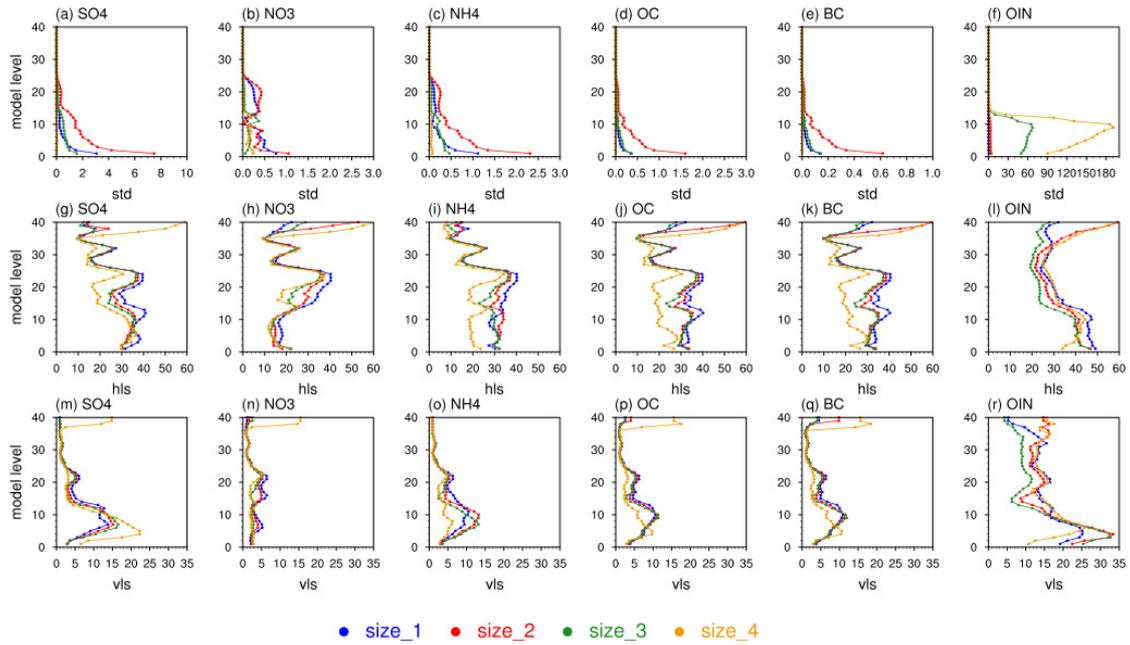


1525

1526

1527 Figure 2. Schematic diagram of the binning strategy for modeling background error
 1528 covariance matrix on (a) the latitude binning data or (b) the gridded data; and the
 1529 vertical profiles of standard deviations ($\mu\text{g kg}^{-1}$) of the coarse OIN component
 1530 concentration at 06:00UTC in April 2019 (c) on average over the latitude bins, (d) at
 1531 Kashi city grid and (e) at the Taklimakan desert grid (i.e., 1.5 degrees east to the
 1532 Kashi city).

1533



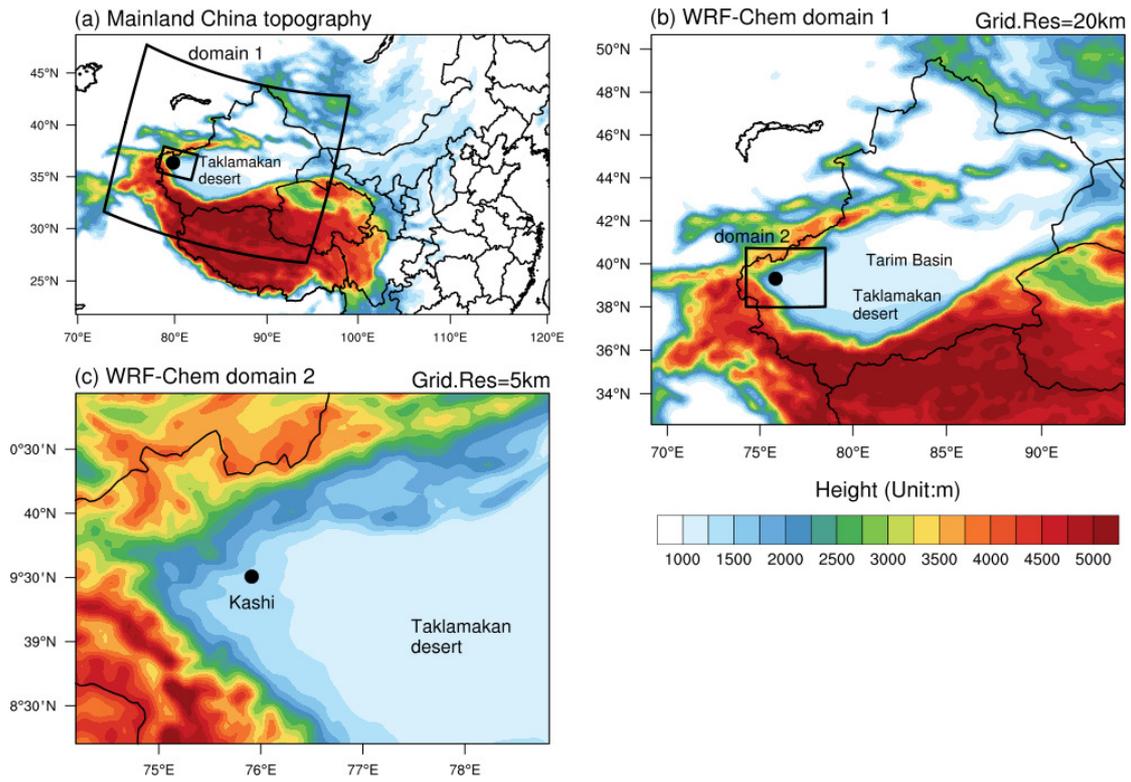
1534

1535

1536 Figure 3. Background error standard deviations at Kashi grid (std, a-f, $\mu\text{g kg}^{-1}$),
 1537 horizontal correlation length scales (hls, g-l, km), and vertical correlation length
 1538 scales (vls, m-r, km) at 00:00 UTC in April 2019 for the sectional sulfate (SO4),
 1539 nitrate (NO3), ammonium (NH4), organic aerosol (OC), black carbon (BC), and other
 1540 inorganic aerosols (OIN, including dust) in the model domain 2. The horizontal and
 1541 vertical correlation length were computed based on the latitude bins with a half degree
 1542 width.

1543

1544



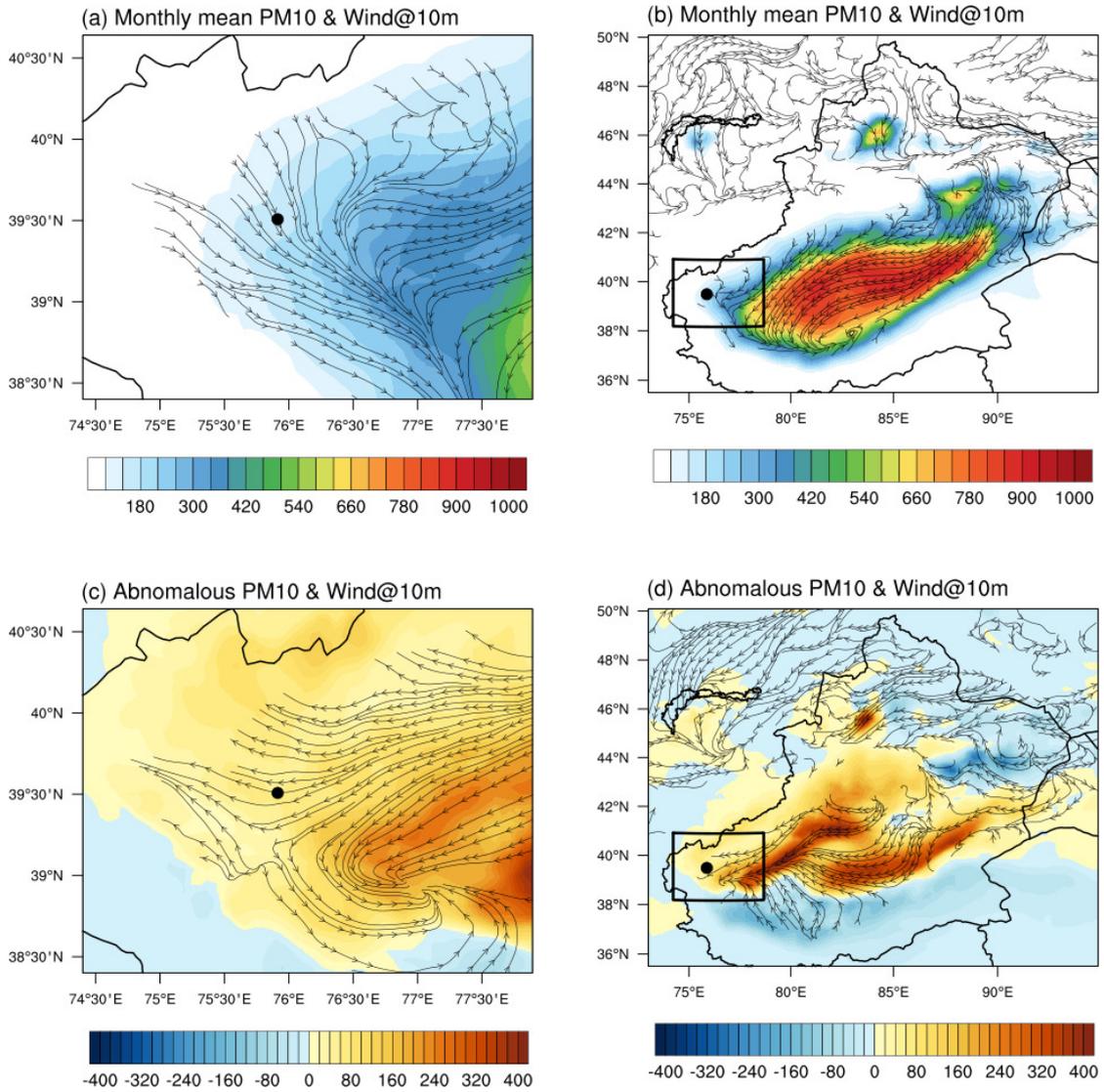
1545

1546

1547 Figure 4. Topography in China (a) and the model domains with the grid resolution of

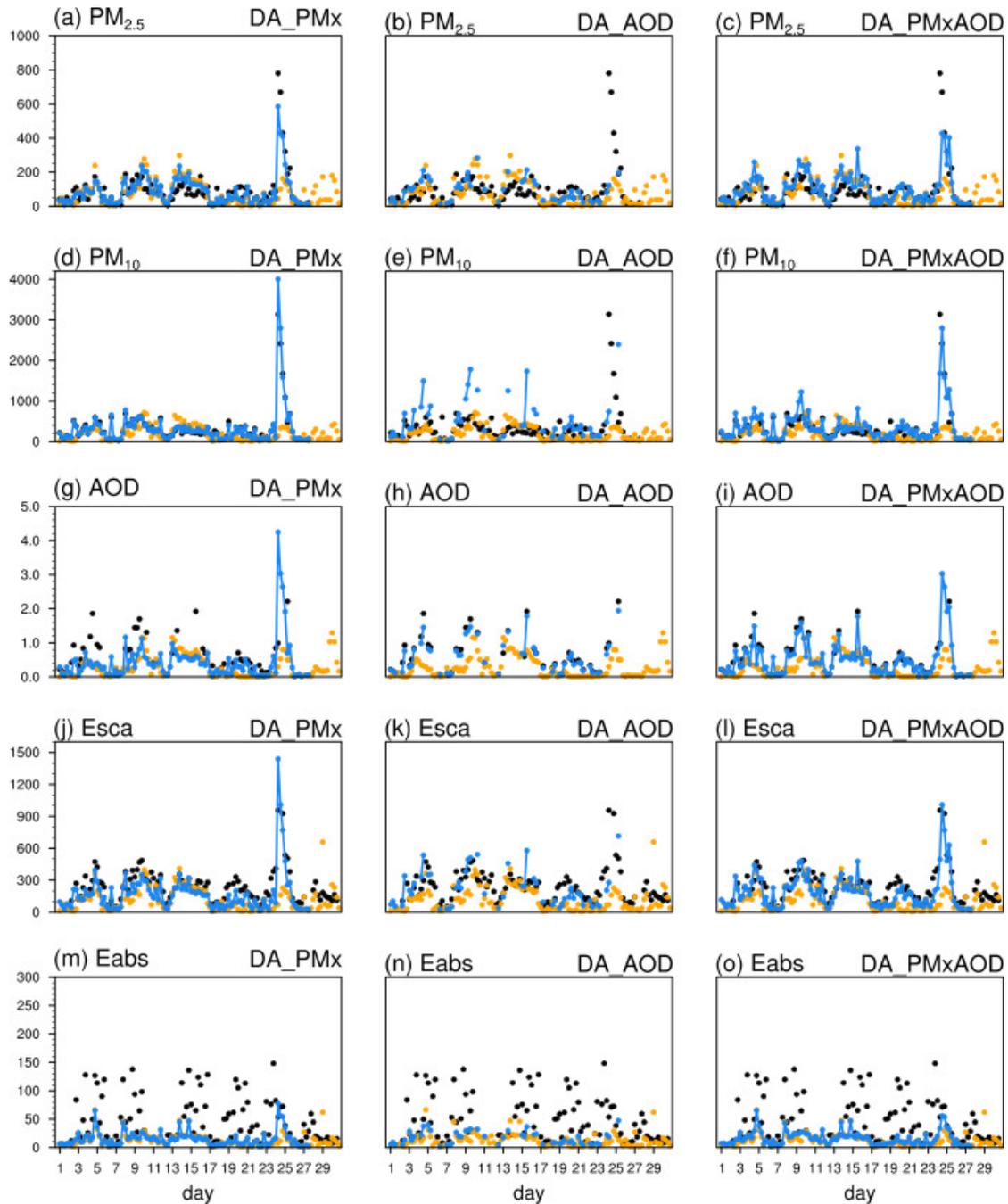
1548 20 km (b) and 5 km (c) in WRF-Chem.

1549



1550
 1551
 1552
 1553
 1554
 1555
 1556
 1557
 1558
 1559
 1560

Figure 5. Monthly mean PM₁₀ concentration ($\mu\text{g m}^{-3}$) and the streamlines of the 10-m wind (m s^{-1}) in April (a, b) and their daily mean anomalies (c, d) during a dust storm on 24 April to the monthly mean values. Only the streamlines at the topographical height lower than 2500 meters are shown for clarity. The rectangles in figures (b) and (d) denote the fine model domain 2, which was the geographical range in the figures (a) and (c). The black points indicate the Kashi city.



1561

1562

1563 Figure 6. Comparison of $PM_{2.5}$ ($\mu g m^{-3}$; a-c), PM_{10} ($\mu g m^{-3}$; d-f), 870 nm AOD (g-i),

1564 635 nm aerosol scattering coefficient (Esca, Mm^{-1} ; j-l), and 660 nm aerosol

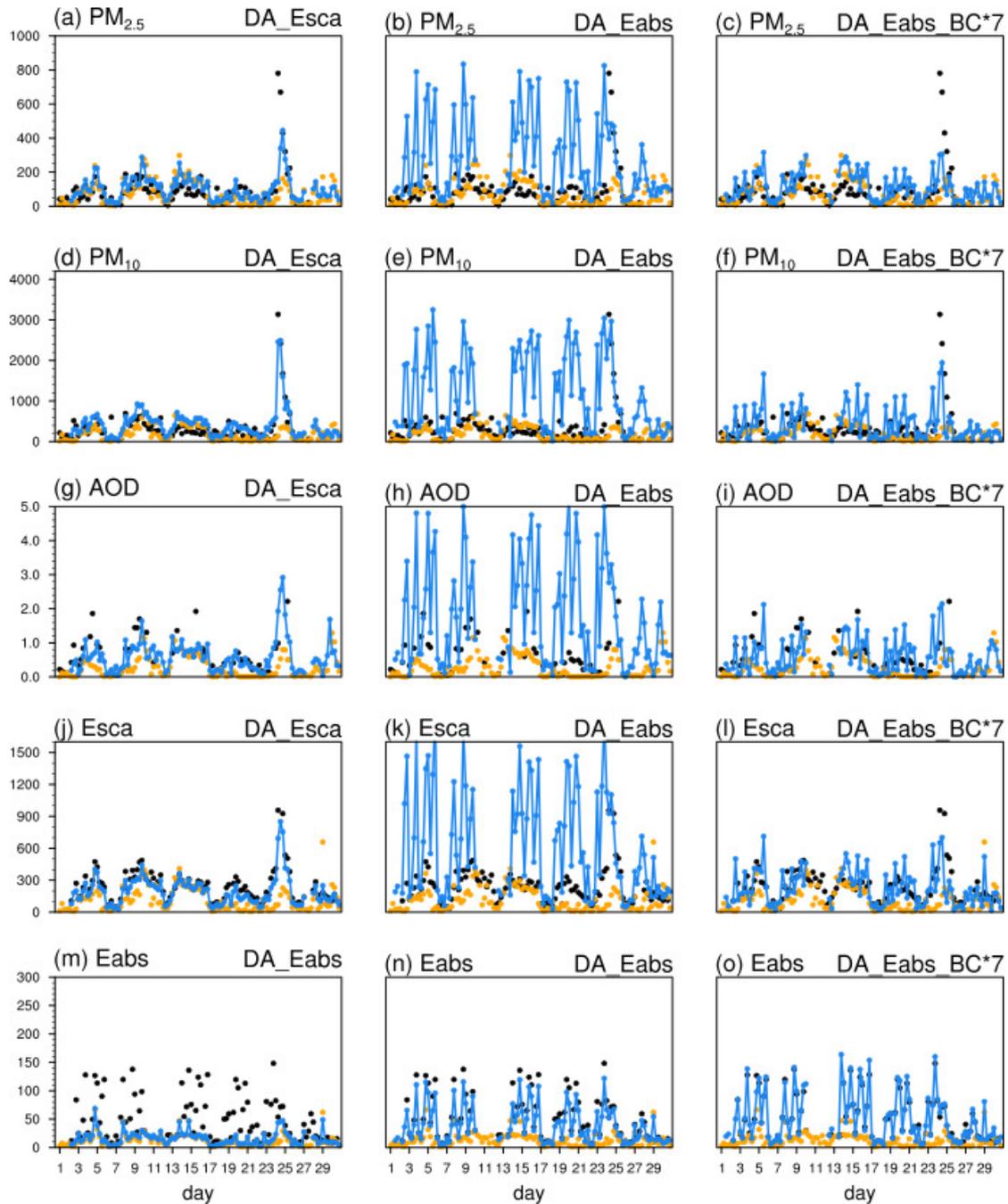
1565 absorption coefficient (Eabs, Mm^{-1} ; m-o) in the observation (black solid point), the

1566 background simulation (orange solid point), and the DA analyses (blue line) when

1567 assimilating the observed $PM_{2.5}$ and PM_{10} (DA_PMx), AOD (DA_AOD), and

1568 simultaneously assimilating PMx and AOD (DA_PMxAOD) at Kashi in April 2019.

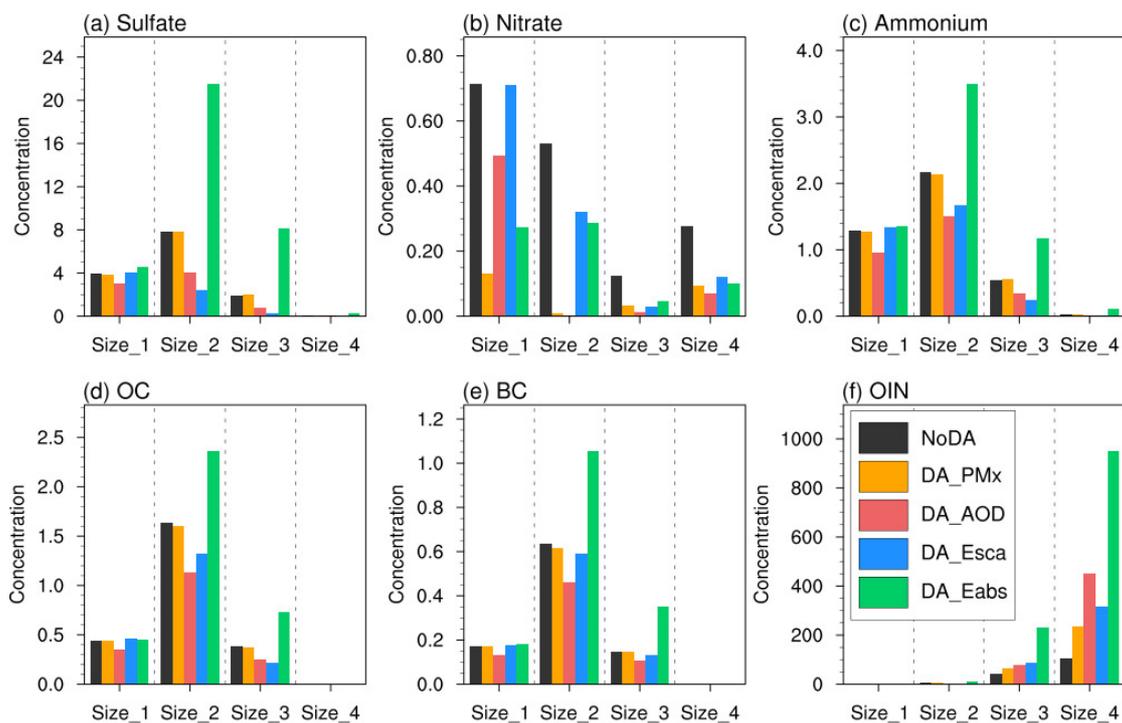
1569



1570
 1571
 1572
 1573
 1574
 1575
 1576
 1577
 1578
 1579

Figure 7. Comparison of $PM_{2.5}$ ($\mu g m^{-3}$; a-c), PM_{10} ($\mu g m^{-3}$; d-f), 870 nm AOD (g-i), 635 nm aerosol scattering coefficient (Esca, Mm^{-1} ; j-l), and 660 nm aerosol absorption coefficient (Eabs, Mm^{-1} ; m-o) in the observation (black solid point), the background simulation (orange solid point), and the DA analyses (blue line) when assimilating the aerosol scattering coefficient (DA_Esca), aerosol absorption coefficient (DA_Eabs), and absorption coefficient with the background error of BC enlarged by a factor of 7 (DA_Eabs_BC*7) at Kashi in April 2019.

1580

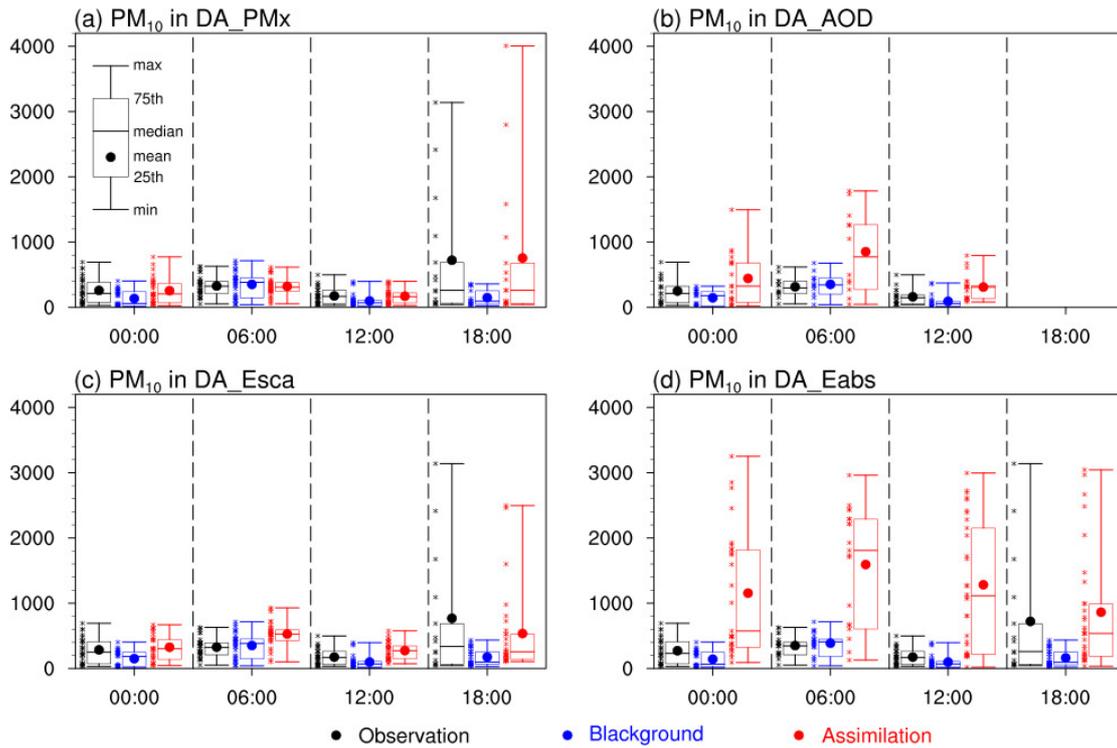


1581

1582

1583 Figure 8. Monthly mean aerosol concentrations ($\mu\text{g m}^{-3}$) per size bin in the
1584 background (NoDA) and the DA analyses when assimilating each individual
1585 observation at Kashi in April 2019.

1586

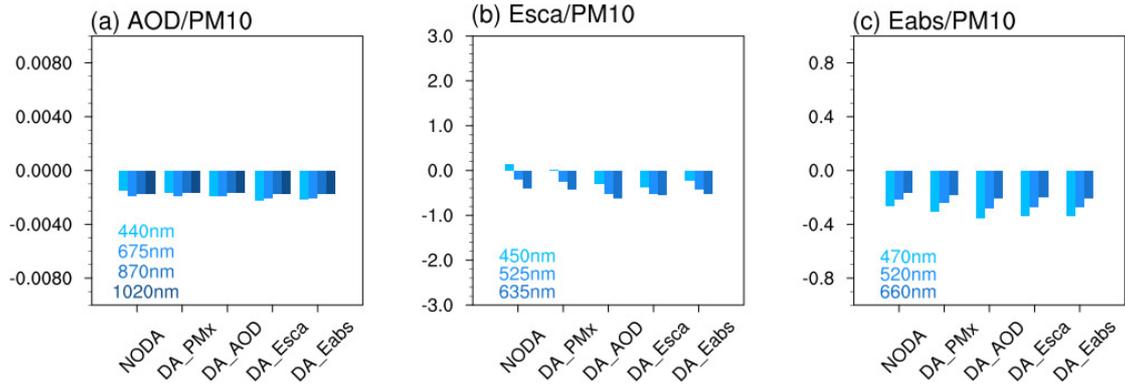


1587

1588

1589 Figure 9. Surface PM_{10} concentrations ($\mu\text{g m}^{-3}$) in the observation (black),
 1590 background simulation (blue) and the DA analyses (red) at 00:00, 06:00, 12:00, 18:00
 1591 UTC in April when assimilating the observations of (a) PM_x , (b) AOD, (c) aerosol
 1592 scattering coefficients (Esca), and (d) aerosol absorption coefficient (Eabs),
 1593 respectively. The DA_AOD had no analysis at 18:00 UTC that was local midnight.
 1594 Kashi is 6 hours ahead of UTC (UTC+6).

1595

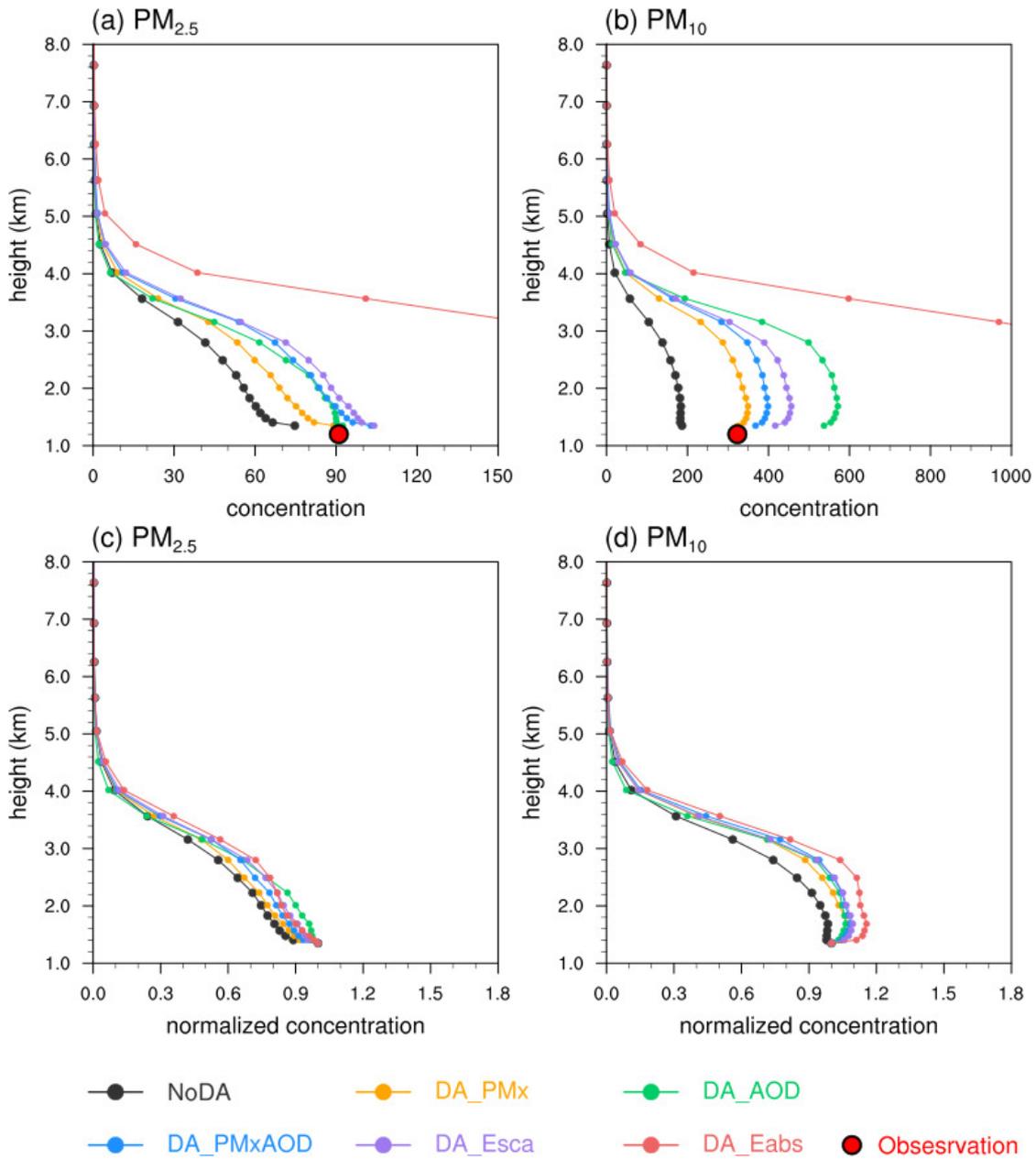


1596

1597

1598 Figure 10. Monthly mean biases in the ratios of AOD to PM₁₀, aerosol scattering
 1599 coefficient (Esca, Mm⁻¹), and aerosol absorbing coefficient (Eabs, Mm⁻¹) at Kashi in
 1600 April 2019.

1601



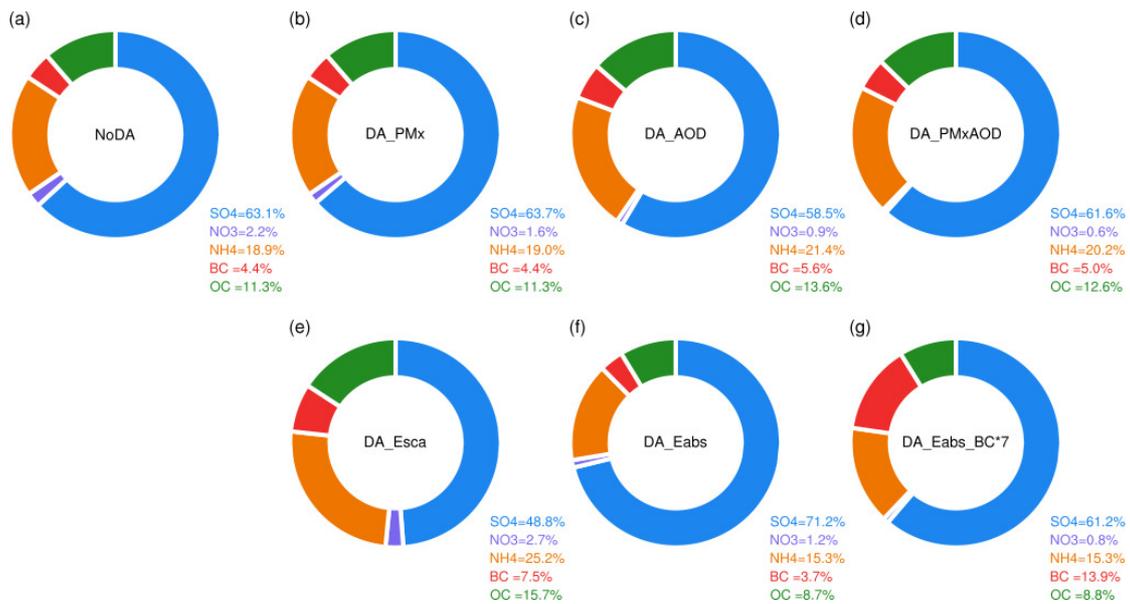
1602

1603

1604 Figure 11. Monthly mean vertical concentration profiles of (a) $PM_{2.5}$ ($\mu\text{g m}^{-3}$), (b)
 1605 PM_{10} ($\mu\text{g m}^{-3}$) and their normalized concentration respect to the surface concentration
 1606 (c, d) at Kashi in April 2019.

1607

1608



1609

1610

1611

1612 Figure 12. Monthly mean chemical composition in percent (%) in the simulated PM₁₀,

1613 excluding the OIN component at Kashi in April 2019.

1614

1615

1616 Table 1. The observed surface particle concentration, aerosol scattering coefficient
 1617 (Esca), aerosol absorption coefficient (Eabs), and AOD used for the DA analysis and
 1618 their observational errors.

1619

	Data time range	Wavelength (nm)	Observation error (e)
PM _{2.5} & PM ₁₀ ($\mu\text{g m}^{-3}$)	Apr 1 – Apr 30		$e = \sqrt{e_1^2 + e_2^2}$ $e_1 = 1.5 + 0.0075 \cdot PM_x$ $e_2 = 0.5 \cdot e_1 \cdot \sqrt{\frac{d}{3000}}$
AOD	Mar 29 – Apr 25	440, 675, 870, 1020	d : grid spacing in meter $e = 0.01/\text{height} \times 10^8$
Esca (Mm^{-1})	Apr 2 – Apr 30	450, 525, 635	$e = 10$
Eabs (Mm^{-1})	Apr 2 – Apr 30	470, 520, 660	$e = 10$

1620

1621

1622 Table 2. The monthly mean values of the PM_{2.5} and PM₁₀ concentrations ($\mu\text{g m}^{-3}$),
 1623 635 nm aerosol scattering coefficient (Esca, Mm^{-1}), 660 nm aerosol absorption
 1624 coefficient (Eabs, Mm^{-1}) and 870 nm AOD in the background and analysis data and
 1625 their correlation values (in brackets) with the observations at 00:00, 06:00, 12:00,
 1626 18:00 UTC at Kashi in April 2019. The underlined number in bold denotes the
 1627 monthly mean value that is not significantly different from the observation, and the
 1628 dashed line denotes an insignificant correlation. Both the statistical tests of the mean
 1629 difference and correlation are conducted at the significance level of 0.05.
 1630

DA experiment	PM _{2.5} ($\mu\text{g m}^{-3}$)	PM ₁₀ ($\mu\text{g m}^{-3}$)	870 nm AOD	635nm Esca (Mm^{-1})	660nm Eabs (Mm^{-1})
Observation	91.0	323.2	0.66	231.5	47.4
Background	<u>75.3</u> (0.28)	190.7 (0.24)	0.24 (0.60)	123.3 (0.36)	12.9 (0.34)
DA_PM _x	<u>89.3</u> (0.89)	<u>329.3</u> (0.99)	0.38 (0.35)	170.4 (0.89)	15.8 (0.42)
DA_AOD	<u>92.6</u> (0.35)	541.7 (0.31)	<u>0.59</u> (0.98)	<u>222.6</u> (0.61)	17.0 (0.26)
DA_PM _x _AOD	<u>103.6</u> (0.61)	<u>372.7</u> (0.86)	<u>0.59</u> (0.98)	<u>192.2</u> (0.86)	16.7 (0.45)
DA_Esca	<u>103.6</u> (0.67)	442.1 (0.93)	<u>0.53</u> (0.62)	<u>192.1</u> (0.97)	16.5 (0.47)
DA_Eabs	298.8 (0.36)	1281.2 (0.34)	1.73 (----)	612.2 (0.54)	<u>40.0</u> (0.98)
DA_Eabs_BC*7	106.7 (0.48)	463.7 (0.45)	<u>0.75</u> (0.50)	<u>226.2</u> (0.52)	<u>51.9</u> (0.90)

1631
 1632

1633 Table 3. The Ångström exponent based on the AOD at 440 nm and 1020 nm and the
 1634 surface single scattering albedo ($SSA_{sr} = \frac{Esca_{525}}{Esca_{525} + Eabs_{520}}$) at Kashi in
 1635 April 2019
 1636

	440-1020 nm Ångström exponent	SSA _{sr}
Observation	0.18	0.78
Background	0.54	0.86
DA_PMx	0.30	0.88
DA_AOD	-0.01	0.88
DA_PMx_AOD	0.17	0.89
DA_Esca	-0.15	0.88
DA_Eabs	-0.01	0.90
DA_Eabs_BC*7	0.33	0.82

1637
 1638

1639 Table 4. The ratios of AOD, aerosol scattering/absorption coefficient to PM₁₀
 1640 concentration (mean ± standard deviation) in the observations, the model background
 1641 data, and the DA analyses.
 1642

	Ratios of 870 nm AOD to PM ₁₀ (μg ⁻¹ m ³)	Ratios of 635 nm aerosol scattering coefficient (Esca) to PM ₁₀ (Mm ⁻¹ μg ⁻¹ m ³)	Ratios of 660 nm aerosol absorption coefficient (Eabs) to PM ₁₀ (Mm ⁻¹ μg ⁻¹ m ³)
Observation	0.0030±0.0020	1.05±0.57	0.25±0.22
Background	0.0013±0.0009	0.65±0.18	0.09±0.05
DA_PMx	0.0013±0.0008	0.61±0.22	0.07±0.05
DA_AOD	0.0013±0.0011	0.51±0.24	0.05±0.04
DA_PMx_AOD	0.0015±0.0010	0.61±0.24	0.06±0.05
DA_Esca	0.0015±0.0010	0.52±0.21	0.05±0.05
DA_Eabs	0.0015±0.0010	0.58±0.37	0.05±0.06
DA_Eabs_BC*7	0.0023±0.0085	0.74±0.51	0.30±0.48

1643
 1644

1645 Table 5. The mean instantaneous clear-sky shortwave (SW), longwave (LW) and the
 1646 net (SW+LW) direct radiative forcing (Wm^{-2}) at the top of atmosphere (TOA), in the
 1647 atmosphere (ATM) and at the surface (SRF) in the background and the simulations
 1648 restarted from the analyses of DA_PMx and DA_PMx_AOD at one hour after the
 1649 analysis times of AOD DA at Kashi in April 2019.

1650

	SW (Wm^{-2})			LW (Wm^{-2})			SW+LW (Wm^{-2})		
	TOA	ATM	SRF	TOA	ATM	SRF	TOA	ATM	SRF
Background	-7.0	+17.0	-24.0	+0.3	-2.9	+3.2	-6.7	+14.1	-20.8
DA_PMx	-8.5	+22.7	-31.2	+0.6	-6.3	+6.9	-7.9	+16.4	-24.3
DA_PMx_A OD	-11.4	+28.6	-40.0	+1.0	-7.8	+8.8	-10.4	+20.8	-31.2

1651

1652