



Impacts of direct CALIPSO Level 1B attenuated-backscatter assimilation
on three-dimensional aerosol analyses and forecasts over northern China

Wei You^{a,b*}, Zengliang Zang^{a,b}, Daichun Wang^a, Wei Zhong^c, Shuang Mei^a, Yiwen Hu^a,
Jianping Guo^d, Dan Niu^e

^aCollege of Meteorology and Oceanography, National University of Defense Technology, Changsha
410073, China

^bHigh Impact Weather Key Laboratory of CMA, Changsha 410073, China

^cCollege of Advanced Interdisciplinary Studies, National University of Defense Technology,
Changsha 410073, China

^dState Key Laboratory of Severe Weather, Chinese Academy of Meteorological Sciences, Beijing
100081, China

^eSchool of Automation, Southeast University, Nanjing 210096, China.

*Corresponding author: Wei You, Email: yw1x_1987@163.com



Abstract

Vertical aerosol distributions remain poorly constrained in atmospheric chemistry models because most surface and satellite observations provide only near-surface or column-integrated information. We developed a three-dimensional variational data assimilation algorithm to use vertically resolved laser measurements from the Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP) instrument to improve aerosol simulations. The system directly assimilates 532 nanometre total attenuated backscatter measurements into the Weather Research and Forecasting model coupled with Chemistry, using the four-bin Model for Simulating Aerosol Interactions and Chemistry aerosol scheme. We evaluated the system using two nighttime satellite observations over northern China in December 2019, representing moderate and heavy pollution conditions. Assimilation improved the simulated vertical distribution of aerosols along the satellite tracks. Compared with a simulation without assimilation, the aerosol optical depth correlation coefficient increased from 0.51 to 0.58, the root-mean-square error decreased from 0.28 to 0.24, and the mean bias improved from -0.11 to 0.01 . For the heavy-pollution case, the initial mean surface concentration of fine particles increased from 85 to 107 micrograms per cubic metre, compared with an observed value of 151 micrograms per cubic metre. The improvement in surface fine-particle forecasts lasted approximately 14 hours. These results demonstrate the potential of satellite lidar measurements to improve three-dimensional aerosol analyses, while highlighting the need for broader evaluation because of limited satellite sampling, cloud effects, and observation uncertainty.



1 Keywords: CALIPSO, attenuated backscatter, data assimilation, 3DVAR, WRF-Chem

2



1. Introduction

Air quality models are essential for simulating the spatiotemporal distribution of aerosols and predicting their temporal evolution. These models offer valuable insights into the interactions between aerosols and the Earth's atmospheric system, enabling a more comprehensive understanding of the environmental impacts of anthropogenic activities (Menon et al., 2002; Grell et al., 2005; Jing et al., 2016; Ma et al., 2019; Yu et al., 2020). Nonetheless, achieving precise aerosol simulation remains a challenging task due to the inherent errors arising from emission factors, initial conditions, and the multifaceted interactions between chemical processes and meteorological phenomena. Aerosol data assimilation employs an optimal estimation method to combine measurements from various sources with numerical models, aiming to minimize errors in the initial aerosol fields of the model (Bocquet et al., 2015; Hu et al., 2022). In the early studies on aerosol DA conducted for offline chemical transport models, the assimilated state variables were commonly aligned with specific observed aerosol variables, such as particulate matter with aerodynamic diameter less than 2.5 micrometers ($PM_{2.5}$) or that with aerodynamic diameter less than 10 micrometers (PM_{10}) (Chai et al., 2006; Niu et al., 2008; Tombette et al., 2009). Thanks to the continuous advancement of atmospheric chemistry models and improvements in computational efficiency, the Coupled Chemistry Meteorology Models (CCMMs) have been undergoing significant development (Bocquet et al., 2015). The WRF-Chem model, renowned for its advanced online coupling of meteorology and atmospheric chemistry (Grell et al., 2005), has gained widespread usage in diverse environmental fields, particularly in air quality modeling. It offers a range of aerosol schemes, including the Goddard Chemistry Aerosol Radiation and Transport (GOCART) scheme proposed by Chin et al. (2000), and the Model for Simulating Aerosol Interactions and Chemistry (MOSAIC) aerosol scheme developed by Zaveri et al. (2008). These advanced aerosol schemes employ a large number of model variables, enabling a thorough representation of the mass and/or size distributions of aerosol chemical species. As a result, they are more suitable for accurately simulating complex real-world aerosol processes. Consequently, DA technology concerning more state variables for a detailed description of aerosol chemical species and size was developed for WRF-Chem and other CCMMs. The Grid-



1 point Statistical Interpolation (GSI) system, initially designed for meteorological applications,
2 has undergone further development to incorporate aerosol data assimilation. It is built upon the
3 widely adopted 3DVAR algorithm (Liu et al., 2011). By employing the GSI system, numerous
4 studies have successfully assimilated PM_{2.5} or PM₁₀ data, leading to enhanced aerosol analyses
5 and subsequent forecasting capabilities (Pagowski et al., 2010; Schwartz et al., 2012; Jiang et
6 al., 2013; Feng et al., 2018). However, the GSI system was initially developed based on the
7 GOCART aerosol scheme, with a primary focus on improving the simulation of dust aerosols.
8 The MOSAIC scheme, an advanced size-resolved sectional model, efficiently classifies
9 aerosols into 8 distinct species and characterizes the size distribution of each species using
10 either 4 or 8 discrete size bins. This size-resolved aerosol scheme proves to be more suitable
11 for simulating complex air pollution scenarios in northern China (Chen et al., 2019).
12 Consequently, researchers have successfully developed an aerosol DA system by utilizing this
13 aerosol scheme in conjunction with the 3DVAR algorithm (Li et al., 2013; Sun et al., 2020;
14 Wang et al., 2020), demonstrating that the application of DA yielded positive effects on both
15 the analysis field and subsequent 24-hour PM_{2.5} forecasts.

16 In terms of observations, conventional in situ measurements are commonly used to
17 measure and assimilate PM_{2.5} or PM₁₀, as well as significant gaseous pollutants like ozone (O₃)
18 and nitrogen dioxide (NO₂) concentrations. Nevertheless, these measurements only provide
19 horizontal aerosol information near the surface and have a limited distribution in rural regions.
20 In comparison, satellite- and lidar-derived aerosol optical properties have many advantages
21 over conventional observations and have high application value in the aerosol DA field. In the
22 field of aerosol optical parameters data assimilation, the representation of the relationship
23 between assimilation state variables and aerosol optical properties relies on a nonlinear
24 observation operator. Therefore, establishing a reasonable observation operator is both critical
25 and challenging (Wang et al., 2022). Recently, researchers have made significant progress in
26 assimilating the optical properties data of aerosols. Among these data, the first parameter to be
27 assimilated is the AOD retrieved from satellite observations. Liu et al. (2011) developed the
28 Community Radiative Transfer Model (CRTM) AOD module and integrated it into the GSI
29 framework to enhance the assimilation of satellite derived AOD data. This innovative approach



1 proved effective in assimilating MODIS AOD data and was subsequently widely applied to
2 assimilate AOD retrievals from various platforms, leading to promising outcomes. (Schwartz
3 et al., 2012; Saide et al., 2014; Ha et al., 2019; Xia et al., 2020; Hu et al., 2026). In addition to
4 the utilization of variational methods, ensemble-based DA approaches have also been
5 employed to assimilate satellite AODs (Yumimoto et al., 2016; Dai et al., 2019; Tsikerdekis et
6 al., 2021; Cheng et al., 2022; Pang et al., 2023; Hong et al., 2026). Nevertheless, it is worth
7 noting that satellite-derived AOD products primarily provide column-integrated aerosol
8 properties, limiting their ability to optimize the aerosol profiles in individual layers.

9 To overcome this limitation, continuous assimilation studies focus on aerosol profiles
10 such as extinction or backscatter coefficients obtained from in-situ lidar measurements (Cheng
11 et al., 2019; Liang et al., 2020; Gao et al., 2025). Compared to ground-based lidar systems, the
12 CALIPSO satellite provides a unique advantage of systematically capturing high-resolution
13 vertical cross-sections on a global scale. However, assimilating CALIPSO data to improve
14 aerosol simulations for CCMMs has seldom been performed (Sekiyama et al., 2010, 2012). To
15 achieve this goal, we devised a 3DVAR aerosol DA algorithm using the MOSAIC scheme
16 (Wang et al., 2022), capable of assimilating AOD, extinction coefficients, backscatter
17 coefficients, and mass concentrations. While our previous study (Ye et al., 2021; Zang et al.,
18 2022) successfully assimilated CALIPSO Level 2 extinction profiles alongside surface PM_{2.5}
19 and Himawari-8 AOD, the current study presents a major advancement: the first known attempt
20 to directly assimilate CALIPSO Level 1B attenuated backscatter data using the 3DVAR
21 method and the MOSAIC mechanism. In comparison to CALIPSO Level 2 aerosol products,
22 the Level 1B attenuated backscatter coefficient profiles are free from contamination by
23 retrieval algorithms and possess a significantly larger data volume (Stromatas et al., 2012).
24 Direct assimilation of Level 1B data obviates superfluous and artifactitious errors arising from
25 data retrieval processes, showing greater potential to enhance aerosol analyses and forecasts.
26 This study extends the previous DA algorithm by incorporating an attenuated backscatter
27 observation operator and its newly established adjoint module.

28 The scientific relevance of vertically constraining aerosol distributions extends beyond
29 improvement of air-quality forecasts. Aerosol altitude determines transport pathways and



1 residence times and influences the vertical distribution of radiative heating, atmospheric
2 stability, cloud interactions, and the relationship between column aerosol optical depth and
3 surface particulate matter. Surface observations and passive satellite retrievals cannot uniquely
4 resolve this vertical structure. Active lidar measurements therefore provide complementary
5 information needed to characterize the three-dimensional aerosol state. Although the present
6 experiments focus on winter pollution over northern China, the challenges examined here—
7 including sparse profile sampling, cloud and surface contamination, nonlinear optical
8 observation operators, and the loss of local assimilation increments through transport—are
9 common to aerosol prediction systems in other regions. This study investigates how direct
10 assimilation of Level 1B attenuated backscatter constrains aerosol vertical structure and how
11 this observational information subsequently affects column and surface aerosol fields.

12 The remaining parts of this paper proceed as follows: Section 2 presents the DA algorithm,
13 model configuration, and observation operator. Given the pivotal role of the observation
14 operator in the DA algorithm, it is discussed in detail. Furthermore, this section presents
15 detailed information regarding the observational data used for assimilation and evaluation. It
16 includes an explanation of the preprocessing techniques employed, quality control measures
17 implemented, and the experimental design adopted. Section 3 primarily focuses on showcasing
18 the effects of data assimilation. Specifically, it highlights the improvements in aerosol
19 simulations achieved by DA. Section 4 contains the conclusion and discussions.

20 **2. Data and methodology**

21 *2.1 The model configuration and aerosol scheme*

22 WRF-Chem v4.0 was adopted to perform the simulation and assimilation experiments.
23 The model operates within a two-configured domain, with an outer nested grid resolution of
24 27 km and an inner nested grid with a horizontal spacing of 9 km. It covers the vertical
25 atmosphere from the surface up to the 50 hPa pressure level. The chosen pollution case took
26 place in the Northern China region, so the evaluation and simulation specifically targeted the
27 inner domain (D02). Consequently, only the D02 map is displayed in Figure 1. The basic model



1 configurations are presented in Table 1.

2 The MOSAIC_4bin chemical mechanism was selected to simulate the three-dimensional
 3 evolution of aerosols. This scheme represents the aerosol size distribution using four discrete
 4 bins spanning 0.039–0.156 μm , 0.156–0.625 μm , 0.625–2.5 μm , and 2.5–10 μm , thereby
 5 resolving the size-resolved behavior of individual species. Previous studies have demonstrated
 6 that MOSAIC is capable of accurately simulating aerosol pollution in China (Chen et al., 2016;
 7 2019). Consequently, aerosols are represented by 32 state variables — 8 aerosol species in each
 8 of the 4 size bins — all of which serve as control variables and are optimized jointly within the
 9 3DVAR assimilation system.

10 2.2 Data assimilation algorithm and observational operator

11 We implemented a DA algorithm based on the 3DVAR method to assimilate the
 12 attenuated backscatter measured by the lidar. The 3DVAR method mathematically minimizes
 13 a cost function, whose incremental form is given by:

$$14 \quad J(\delta\mathbf{x}) = \frac{1}{2}\delta\mathbf{x}^T \mathbf{B}^{-1} \delta\mathbf{x} + \frac{1}{2}(\mathbf{H}\delta\mathbf{x} - \mathbf{d})^T \mathbf{R}^{-1}(\mathbf{H}\delta\mathbf{x} - \mathbf{d}). \quad (1)$$

15 Here, the incremental analysis variable, denoted as $\delta\mathbf{x}$, is defined as the subtraction of the
 16 analysis variable $\mathbf{x}$ from its corresponding background variable $\mathbf{x}^b$. The background error
 17 covariance matrix, represented by $\mathbf{B}$, plays a crucial role in the process of data assimilation by
 18 controlling the distance and magnitude of propagating observational information onto the
 19 model grid points. The $\mathbf{B}$ matrix was constructed using the National Meteorological Center
 20 (NMC) method, where the background error statistics were derived from the differences
 21 between 24-hour and 48-hour forecasts valid at the same time. The cross-correlations between
 22 the 32 MOSAIC state variables across different species and size bins are explicitly considered
 23 in the $\mathbf{B}$ matrix. Furthermore, horizontal and vertical localizations are implicitly applied
 24 through the static correlation length scales defined in $\mathbf{B}$, ensuring observation information is
 25 spread reasonably to neighboring grid points. The $\mathbf{d} = \mathbf{y} - \mathbf{H}\mathbf{x}^b$ represents the observation
 26 innovation, where $\mathbf{y}$ is the observation vector which represents various aerosol observations.
 27 The observation operator, denoted as $\mathbf{H}$, is the linear approximation that maps the aerosol



variables in the MOSAIC_4bin scheme to the corresponding observational variables. The observation error covariance matrix is represented by $\mathbf{R}$. The mathematical formulation utilized by the 3DVAR method affords the ability to assimilate different types of aerosol measurements, either separately or simultaneously, resulting in significant advantages. The fundamental components of $\mathbf{B}$ and $\mathbf{R}$ are the key elements that drive the effectiveness of the DA algorithm. Our previous studies have extensively discussed and analyzed these covariances, shedding light on their intricate details and implications within the context of the DA algorithm (Wang et al., 2022).

Establishing an observation operator is necessary in order to obtain the minimum solution for the previously described cost function. Notably, the state variables utilized in the DA algorithm are derived from the MOSAIC_4bin scheme, which characterizes aerosol chemical species mass concentration, and are distinct from attenuated backscatter observations. Nevertheless, the observation operator can map the model control variables to the observational variables. It is crucial to acknowledge that while a simple linear operator is adequate for conventional observations such as $\text{PM}_{2.5}$ mass concentrations, the optical property observations necessitate a complex and nonlinear operator. Furthermore, it is worth noting that the tangent linear and adjoint operators are obtained through code transformations of the observation operator that adhere to strict coding rules. The computation procedure for the operator associated with attenuated backscatter is presented below:

Initially, the calculation of two optical parameters, namely the aerosol mean wet radius (r) and the complex refractive index (m), relies on the aerosol state variables. Based on the aerosol density of each state variable, the individual volumes for each species can be calculated. The sum of the individual volumes, together with the water volume, provides the total aerosol volume. Dividing this total volume by the aerosol number yields the mean volume. Additionally, the assumption is made that aerosol particles possess a spherical shape, enabling the calculation of the mean wet radius r from the mean volume using the formula for the volume of a sphere. By employing volume averaging, the mean complex refractive index m is calculated by dividing the sum of the products of the individual species volume and their respective complex refractive indices by the total volume. Subsequently, the computation of aerosol optical efficiencies is performed.



1 According to Mie scattering theory, the efficiencies depend on both the complex refractive
 2 index of the aerosol particles and the particle size parameter (x). Once the complex refractive
 3 index (m) and mean wet radius (r) have been calculated, the conversion of r to x is
 4 accomplished using the equation $x=2\pi r/\lambda$, where λ represents the wavelength of the incident
 5 radiation. The Mie calculation is a widely accepted method for directly computing aerosol
 6 optical efficiencies by utilizing the two optical parameters, x and m . It is well-established that
 7 this process is highly intricate and characterized by nonlinearity. This computation consumes
 8 a considerable amount of time and presents significant challenges in writing adjoint program
 9 codes. In order to enhance efficiency in the context of assimilation, we employed polynomial
 10 expansions to approximately calculate extinction and backscatter efficiencies. As documented
 11 by Fast et al. (2006), this method has proven to be highly effective in simplifying the
 12 computation process while maintaining accuracy. The expression for the extinction efficiency
 13 Q_{ext} can be represented as follows:

$$14 \quad Q_{\text{ext}} = \exp\left[\sum_{i=1}^M A_i T_i(s)\right], \quad (2)$$

15 where s is represented as the logarithm of the wet radius r . The notation $T_i(s)$ denotes the
 16 Chebyshev polynomial of order i , where i represents the degree of the polynomial. The symbol
 17 M is used to denote the total number of terms present in the polynomial expansion. By
 18 employing a predefined set of coefficients, the computation of the expansion coefficients A_i ,
 19 which are reliant on the complex refractive index m , is achieved through the application of
 20 bilinear interpolation technique. For more detailed guidelines on calculating A_i , please refer to
 21 the comprehensive instructions provided in Fast et al. (2006). By substituting the determined
 22 coefficients A_i and s into the equation, we can derive Q_{ext} , and similarly, the backscatter
 23 efficiency can be computed.

24 After obtaining the extinction and backscatter efficiencies, the subsequent step involves
 25 the computation of their corresponding coefficients. It is noteworthy that the determination of
 26 optical parameters necessitates the calculation of the cumulative sum of all values within the
 27 four size bins. Therefore, the extinction coefficient b_{ext} can be expressed as follows:

$$28 \quad b_{\text{ext}} = \sum_{i=1}^{4\text{bins}} N_i \pi r_i^2 Q_{\text{ext}}(r_i, m_i), \quad (3)$$



1 Likewise, the backscatter coefficient can be expressed using a similar expression like Eq.
 2 3, where the subscript $i=1, 2, 3, 4$ represents the size bin, and N_i is the aerosol number at the i -
 3 th size bin. Once the aerosol backscatter coefficient, aerosol extinction coefficients, or AOD,
 4 have been assimilated, the observation operator would be deemed complete at this stage. The
 5 focus of this study is on the assimilation of attenuated backscatter coefficients. Therefore, the
 6 operator should proceed by computing the attenuated backscatter coefficients from the
 7 previously obtained aerosol extinction and backscatter efficiencies

8 The ultimate stage of the operator encompasses the transformation of aerosol backscatter and
 9 extinction coefficients into attenuated backscatter. This can be achieved by applying the lidar
 10 equation (Fernald, 1984), which enables the computation of attenuated backscatter as follows:

$$11 \quad \beta'(z) = (\beta_m(z) + \beta_a(z))T^2(z), \quad (4)$$

12 where the term $T^2(z)$ represents the two-way transmittance resulting from atmospheric
 13 extinction, which is expressed as:

$$14 \quad T^2(z) = \exp(-2 \int_0^z (\alpha_a(r) + \alpha_m(r))dr), \quad (5)$$

15 where the coefficients α_a (α_m) and β_a (β_m) refer to the aerosol (molecular) extinction and
 16 backscatter coefficients, respectively, and their values are dependent on the altitude z . The
 17 attenuated backscatter, denoted by $\beta'(z)$, can be expressed as the product of the two-way
 18 optical transmittance and the backscatter coefficient. As indicated in the equation (4), the
 19 attenuated backscatter detected by the lidar comprises both molecular and aerosol components,
 20 with the molecular terms remaining constant and unaffected by aerosols during the assimilation
 21 process.

22 In line with Wang et al. (2014) approach, the molecular α_m and β_m were computed
 23 from WRF-Chem outputs through the application of Rayleigh scattering theory. More
 24 specifically, the molecular β_m was derived through the following calculation:

$$25 \quad \beta_m = \frac{P}{\sigma_b T} S_{Ray}, \quad (6)$$

26 where P denotes pressure, T denotes temperature, and σ_b represents the Boltzmann constant,
 27 while the Rayleigh scattering cross section S_{Ray} is formulated as follows:



$$S_{Ray} = 4.678 \times 10^{-29} \cdot \lambda^{-(3.916+0.074\lambda+0.05/\lambda)}, \quad (7)$$

where λ denotes the wavelength of the incident light. The molecular extinction coefficient α_m can be computed as follows:

$$\alpha_m = \frac{8\pi}{3} \beta_m. \quad (8)$$

The transmittance term is calculated by the integration of both the aerosol α_a , and the molecular α_m over the entire trajectory of the laser beam. Therefore, the value of transmittance is influenced by the distance between the volume that backscatters the laser and the lidar. The main purpose of this investigation is to assimilate the CALIPSO attenuated backscatter coefficient profiles. In this context, the transmittance terms denote the attenuation from the upper atmosphere to the lower target height, in contrast with ground-based systems, where the transmittance term is calculated from the ground up to the height of the upper target (Mamouri et al., 2009; Rogers et al., 2011). The successful derivation of the attenuated backscatter coefficients, resembling those observed in CALIPSO, indicates the completion of the calculation process for the operator at this stage. The overall process reveals that the lidar-derived attenuated backscatter observation operator incorporates multiple observation operators or calculation segments for various aerosol optical properties, including extinction coefficients, backscatter coefficients, and AOD. Therefore, it can be regarded as a comprehensive operator, capable of assimilating not only attenuated backscatter but also these other optical properties.

2.3 Observations and Quality Control

The CALIPSO satellite is equipped with CALIOP, a three-channel lidar system measuring at 532 nm (parallel and perpendicular) and 1064 nm (Winker et al., 2007). After the backscatter signals undergo range scaling, energy normalization, and gain normalization, the resulting attenuated backscatter profiles are used to generate several products. The CALIPSO Level 1B data products provide high-resolution altitude profiles of total attenuated backscatter coefficients. Because Level 2 products rely on Level 1B retrievals, directly assimilating Level 1B data avoids retrieval algorithm artifacts. The high quality of attenuated backscatter data has



1 been verified by numerous studies (Mamouri et al., 2009; Rogers et al., 2011), highlighting the
2 significant importance of its direct assimilation.

3 The CALIPSO Level 1B 532 nm total attenuated backscatter profile products (version
4 4.10) were assimilated in this study. Each profile has 583 elements in the vertical dimension
5 for the -2.0 to 40 km altitude range. Different horizontal and vertical resolutions are applied
6 for different altitude regimes to reduce the data volume. The provided "full resolution" (30 m
7 vertical and 0.33 km horizontal) spans from -0.5 to 8.2 km. To facilitate the monitoring of
8 potential delayed recovery from the surface return, the region ranging from -2.0 to -0.5 km is
9 provided with a horizontal spacing of 0.33 km and a vertical spacing of 300 m. The data is
10 presented at a horizontal spacing of 1 km and a vertical resolution of 60 m, covering an altitude
11 range from 8.2 to 20.2 km. Notably, in the stratospheric region between 20.2 and 30.1 km, the
12 vertical resolution is reduced to 180 meters. At altitudes exceeding 30 kilometers, characterized
13 by negligible aerosol concentration, a coarse spatial resolution proves sufficient for the
14 intended purpose. As aerosols remain in the lower troposphere, the attenuated backscatter data
15 below 5 km are assimilated. The uncertainty of the attenuated backscatter was assumed to be
16 20% as in Winker et al. (2007). As the magnitude of aerosol attenuated backscatter is 10^{-3} , the
17 observation error of attenuated backscatter in this study is estimated to be $\sim 2 \times 10^{-4}$.

18 In this study, CALIPSO Level 1B 532 nm total attenuated backscatter profile products
19 (version 4.10) were assimilated. The two cases were carefully selected to represent distinct but
20 widely encountered wintertime pollution conditions in northern China: the light pollution
21 episode (Case 1, December 6 2019, average near-track $\text{PM}_{2.5} \sim 72 \mu\text{g m}^{-3}$) and the heavy
22 pollution event (Case 2, December 9 2019, average near-track $\text{PM}_{2.5} \sim 151 \mu\text{g m}^{-3}$). This paired
23 design allows us to comprehensively validate the performance of our newly developed
24 attenuated backscatter assimilation algorithm under low-to-moderate and high aerosol loading
25 conditions respectively, covering most typical winter haze scenarios in the study region.
26 Furthermore, both cases are selected under nighttime CALIPSO overpasses ($\sim 18:00$ UTC) to
27 minimize solar background noise, which ensures the quality of Level 1B attenuated backscatter
28 observations and avoids introducing confounding factors from daytime radiation
29 contamination (Noel et al., 2014). While future work will extend testing to more cases across



different seasons, these two representative, well-characterized cases provide robust proof-of-concept validation for this novel direct assimilation approach.

To ensure data quality and computational efficiency, preprocessing and quality control were applied: (1) Data were vertically smoothed using a Gaussian method to reduce random noise; (2) observation data were thinned horizontally by averaging several profiles (e.g., 27 profiles) to match the ~9 km model grid resolution. This practice of thinning high-resolution data to align with the model grid is commonly applied in practical situations such as geostationary-satellite-derived AOD assimilation (Yumimoto et al., 2016; Ha et al., 2019). (3) Data with extreme values larger than 0.01 km^{-1} were removed, and data below 100 m were excluded from assimilation to prevent overestimation caused by strong surface return signals. This step was necessary because surface return signals in CALIPSO data are stronger than aerosol return signals, making it difficult to accurately determine surface effects when deriving aerosol attenuated backscatter coefficients. Consequently, the exclusion of near-surface attenuated backscatter coefficients was necessary to prevent significant overestimation. Please note that the data preprocessing conducted in the current study is considered elementary and warrants further research. After preprocessing, the attenuated backscatter data at 532 nm from CALIPSO are illustrated in Figures 2a and 2b for Cases 1 and 2, respectively. These figures correspond to tracks (a) and (b) as shown in Figure 1.

The verification of DA effects incorporated CALIPSO Level 2 aerosol profile products—including the 532 nm total backscatter coefficient, extinction coefficient profiles, and column aerosol optical depths—alongside attenuated backscatter data and in-situ $\text{PM}_{2.5}$ measurements. CALIPSO Level 1B and Level 2 data were obtained from the NASA Langley Atmospheric Science Data Center (ASDC; <https://asdc.larc.nasa.gov/>), with comprehensive product details available at the official CALIPSO website (<https://www-calipso.larc.nasa.gov/>). $\text{PM}_{2.5}$ data were acquired from the China National Environmental Monitoring Centre (CNEMC; <http://www.cnemc.cn/en/>), which maintains over 2000 monitoring sites across China, 683 of which are located within the D02 region. Because DA spreads observation information to neighboring model grids within a limited area, only $\text{PM}_{2.5}$ measurements from stations situated less than 50 km from the CALIPSO ground track were selected for the evaluation of DA effects in Case 2, resulting in 70 monitoring sites depicted as green circles in Figure 1.



2.4 Experimental design

To explore the impact of DA on aerosol analysis and forecast accuracy, two sets of experiments were conducted for each of the two cases. These experiments consisted of a control experiment, which did not involve DA, and an assimilation experiment, in which DA was applied. Both experiments had the same model configuration but with a difference in aerosol initialization. The control experiment, referred to as “Control,” did not assimilate observations and employed a previous 24 h aerosol forecast following model spinning up as the initialization. The assimilation experiment referred to as “Assimilation,” employed aerosol analyses generated by assimilating CALIPSO attenuated backscatter as the initialization. Both experiments were run for two cases that corresponded to the attenuated backscatter at two different times. Case 1 simulation was for the light aerosol pollution episode taking place from December 6 to 7, 2019, which was initialized at 18:00 UTC on December 6, 2019. The simulation for Case 2 was conducted to study a heavy pollution event that took place from December 9 to 10, 2019, with initialization at 18:00 UTC on December 9, 2019. The assimilation experiment, in contrast to the control experiment, was exclusively assimilated at the initial time of 18:00 UTC on December 6 and 9, 2019, respectively. The meteorological data utilized to drive the WRF-Chem model were obtained from the National Centers for Environmental Prediction (NCEP) reanalysis dataset, while the aerosol emission inventory employed in this study was sourced from the Multi-resolution Emission Inventory for China (MEIC) compiled by Zheng et al. (2018)

To assess the impact of assimilation, three statistical metrics were used: the correlation coefficient (CORR), mean bias (BIAS), and root-mean-squared error (RMSE). These metrics were computed for the purpose of evaluating the level of agreement between the model simulations and observations across all experiments. Better indicators indicate better model performance, which can be ascribed to attenuated backscatter assimilation.

3. Results



1 3.1 Evaluation with attenuated backscatter coefficient

2 The assimilated attenuated backscatter data were used for “sanity checks” to determine
 3 whether the observation operator and DA algorithm worked properly. The attenuated
 4 backscatter coefficient measurements are represented in the form of a vertical cross-section
 5 since each profile is obtained as the CALIPSO satellite passes over North China (Figure 2).
 6 Figure 3 displays the cross-sections of attenuated backscatter acquired from both the control
 7 and assimilation experiments, along with the increment which represents the disparity between
 8 the assimilation outcome and the control outcome. In Figure 3, the top row illustrates the
 9 attenuated backscatter cross-sections obtained from Case 1, while the bottom row showcases
 10 the attenuated backscatter cross-sections for Case 2. The attenuated backscatter simulations
 11 (Figures 3a1, 3b1, 3a2 and 3b2) are in the same order as the measurements (Figure 2). In Case
 12 1 (Figure 3a1), the layer around 37.5° N at approximately 1 km altitude exhibits high values
 13 that demonstrate a favorable agreement with the observations (Figure 2a). This result can help
 14 demonstrate the reliability of the constructed observation operator. However, the control
 15 experiments underestimated attenuated backscatter values in both cases (Figure 3a1, 3a2). The
 16 assimilation experiments conducted through DA yielded considerably positive increments and
 17 slightly negative values (Figure 3c1, 3c2). Consequently, the assimilation results (Figure 3b1
 18 and 3b2) demonstrated a notably higher level of concordance with the observations in
 19 comparison to the control results. The increment peak exceeds 0.003 in Case 1 (Figure 3c1),
 20 which is equivalent to the background values in the control experiment. These results show
 21 that the DA algorithm worked properly and the attenuated backscatter DA hopes to improve
 22 aerosol analysis.

23 In addition, the assimilation increments of attenuated backscatter are driven primarily by
 24 mass increases in the accumulation mode size bins (Bin 2: 0.15625-0.625 μm and Bin 3: 0.625-
 25 2.5 μm) of the MOSAIC scheme, which dominate aerosol light scattering properties under this
 26 winter North China pollution scenario. Further decomposition reveals that the major
 27 contributors to these increments are sulfate, nitrate, ammonium, and organic aerosol
 28 components, which are the dominant anthropogenic fine aerosol species during winter haze



1 events in the study region. In contrast, coarse mode aerosols (Bin 4: 2.5-10 μm) contribute
2 relatively little to the attenuated backscatter increments in these two selected cases, as dust
3 aerosol concentrations are relatively low during these winter pollution episodes.

4 To investigate the DA performance in the vertical dimension explicitly, the simulated and
5 observed attenuated backscatter profiles at three points are shown in Figure 4. The three points
6 A, B, and C were randomly selected, as presented in Figure 1. In contrast to the control results,
7 when attenuated backscatter was assimilated through DA, the values were enhanced in the
8 analysis fields. Consequently, the profiles after assimilation (red lines) became more consistent
9 with the observations (green lines). It is noteworthy that the exclusion of data during the data-
10 quality process has a significant influence on DA performance. For example, the near-surface
11 attenuated backscatters are high at points B and C, while the control experiments underestimate
12 them here (Figure 4b, 4c). However, because the measurements below 100 m are excluded, the
13 DA has little improvement here. The assimilation experiments also underestimate attenuated
14 backscatters as in the control experiment.

15 *3.2 Evaluation with extinction and backscatter coefficients*

16 To avoid repetition, the Case 1 extinction coefficients and Case 2 backscatter coefficients
17 are shown in Figures 5a and 5b, respectively. As evident from the observations, aerosols exhibit
18 a non-uniform vertical distribution (Figure 5a, 5b). In Case 1 (Figure 5a), aerosols are
19 predominantly concentrated below 2 km, while in Figure 5b, they extend from the surface up
20 to an altitude of 3.5 km. Compared with the Level 1B attenuated backscatter data (Figure 2),
21 much of the Level 2 data were not retrieved after strict quality control, represented as the blank
22 region in Figure 5. For instance, attenuated backscatter data contaminated by clouds and
23 significantly affected by background noise were removed. Additionally, the surface has a
24 severe impact on the near-surface attenuated backscatter, which was removed. As a result, there
25 are no available extinction and backscatter coefficients for aerosols near the surface (Figure 5).
26 Extinction coefficients typically exhibit magnitudes on the order of 10^{-1} . The maximum
27 reaches more than 1.0 km^{-1} . Backscatter coefficients are of the order of 10^{-2} to 10^{-3} , except for



1 some high values which are more than 0.01 km^{-1} . The vertical distribution of aerosols can be
 2 derived from the extinction and backscatter coefficients. As exemplified in Figure 5b, the
 3 decrease in aerosol concentration with height holds great significance for both the analysis and
 4 prediction of aerosols.

5 The simulated extinction coefficient cross-sections corresponding to Case 1 are shown in
 6 Figure 6. The extinction coefficients were simulated at 532 nm. Both extinction coefficient
 7 distributions simulated by the control and assimilation experiments are consistent with the
 8 observations for shape and range (Figure 6a, 6b), indicating that the segment computing the
 9 extinction coefficient in the observation operator is reliable. The positive increment in the
 10 extinction coefficient generated by attenuated backscatter assimilation covers a large range
 11 (Figure 6c). However, the increment is relatively small. For example, most background values
 12 are between 0.2 and 0.3. The increment peak reaches about 0.03 km^{-1} (Figure 6c). Compared
 13 with the observations, the assimilation experiment improved the extinction coefficient
 14 distribution, with a small enhancement. It is noted that there is a rather high region of more
 15 than 1.0 km^{-1} around 36.6°N , in the extinction coefficient observations (Figure 5a). However,
 16 this phenomenon is not reflected in the simulation after assimilation. This is partly because
 17 extreme attenuated backscatter values are excluded from assimilation. Therefore, only minor
 18 and insignificant increments are generated, which may lead to an overestimation of extinction
 19 coefficients.

20 For a more straightforward comparison, the simulated extinction coefficient profiles were
 21 evaluated against observations at specific locations A, B, and C. The corresponding results can
 22 be found in Figure 7. In general, the assimilation experiments performed better than the control
 23 experiments considering extinction coefficient simulations, which can be ascribed to
 24 attenuated backscatter DA. Similarly, the DA effects are represented in the form of
 25 enhancements in the extinction coefficient, as the control experiments underestimated the
 26 extinction coefficients. The DA effects vary with points, for instance, the profile after
 27 assimilation is more consistent with the observed profile at point C (Figure 7c) and point B
 28 except that the extinction coefficient is overestimated at the layer of several meters near the 1
 29 km height (Figure 7b), there are no DA effects generated below 1 km at point A, and the effect



1 is slight above 1 km (Figure 7a).

2 Figure 8 shows the simulated backscatter coefficient cross-sections of Case 2 at 532 nm.
 3 Assimilation significantly improves aerosol analyses if evaluated using backscatter
 4 coefficients. Case 2 represents a heavy pollution event, which is illustrated by the high
 5 backscatter coefficient distribution between 39°N and 32°N (Figure 5b), and also by the
 6 averaged surface PM_{2.5} mass concentration of approximately 151 $\mu\text{g m}^{-3}$ over sites near the
 7 track (Figure 11); however, the control experiment seriously underestimated the backscatter
 8 coefficients, failing to depict this event (Figure 8a). With significantly positive increments
 9 generated (Figure 8c), the assimilation experiment produced a backscatter coefficient
 10 distribution that was more consistent with the observation. The analyses were equivalent to the
 11 observed values in terms of value, but with some slight differences, and both distribution
 12 shapes were very similar (Figure 8b). These results also show that the segment for computing
 13 backscatter coefficient in the observation operator is reliable.

14 Similarly, comparisons between the simulated and observed backscatter coefficient
 15 profiles at points A, B, and C are shown in Figure 9. The results show that the analyses in the
 16 assimilation experiments (red lines) are larger than the background values in the control
 17 experiments (blue lines) (Figures 9a, 9b, and 9c). The control experiment substantially
 18 underestimated the backscatter coefficients; however, data assimilation significantly
 19 minimized the discrepancies between the simulations and the observations. Consequently, the
 20 simulated profiles exhibit greater proximity to the observations. The DA effects also differ with
 21 points. The simulated profiles after assimilation are more consistent with the observed profiles
 22 at points B and C, except that the simulations are far less than the observations near the surface.
 23 This phenomenon is caused by data exclusion near the surface. Small DA effects are generated
 24 by spreading downwards the observation information contained in the upper layer. The
 25 backscatter coefficient profile is little changed by assimilation below ~300 m at point A and it
 26 is slightly enlarged above 300 m. Consequently, the profile after assimilation is still far less
 27 than the observation, especially In the layer close to the surface (Figure 9a), which is similar
 28 to the extinction coefficient analysis at A (Figure 7a). The authors think that this small
 29 improvement is reasonable. Considering attenuated backscatter, both simulated profiles are



1 nearly equal to the observation below 1 km at A (Figure 4a). Therefore, little or no DA effects
2 are generated in this region, which is demonstrated by the improvements in extinction and
3 backscatter profiles. In addition, it is noteworthy that extinction or backscatter observations
4 are much larger than the simulations in the lower boundary layer near point A. This may be
5 because the level 2 data are overestimated in this region. For the above region, the attenuated
6 backscatters remain relatively small (less than $2 \times 10^{-3} \text{ km}^{-1}$), while the retrieved extinction
7 coefficient reaches more than 1.0 km^{-1} , which seems to be unreasonable. This needs further
8 validation.

9 3.3 Evaluation with AOD

10 AOD is an independent, column-integrated variable that contains the total aerosol
11 information in the vertical column. The purpose of evaluating AOD is to verify that the vertical
12 profile assimilation not only improves layered distributions but also accurately corrects the
13 total columnar aerosol optical depth. The scatter points between the simulated and CALIPSO
14 AODs are depicted in Figure 10. As the CALIPSO satellite passes over one region, it can only
15 observe the range of a line, implying that AOD is distributed on the ground track shown in
16 Figure 1 with a very small region covered. The amount is much less than that of the
17 geostationary satellite-derived AOD (Yumimoto et al., 2016; Ha et al., 2019; Zang et al., 2022).
18 There were a total of 463 AOD data points for both cases. It can be observed that the majority
19 of the blue data points fall below the 1:1 line, suggesting that the control experiments
20 consistently underestimated AOD values. For those at the horizontal axis, the underestimation
21 is crucial as the red counterparts move upwards by assimilation and attempt to get close to the
22 1:1 line, but remain scattered. After assimilation, the indicators CORR, RMSE, and BIAS were
23 computed for each experiment, revealing notable changes. Specifically, CORR exhibits an
24 increase from 0.51 to 0.58, RMSE shows a decrease from 0.28 to 0.24, and BIAS experiences
25 a reduction from -0.11 to 0.01 . The changes in the three indicators consistently indicate that
26 the AOD simulations were improved by assimilation.



1 3.4 Evaluation with surface PM data

2 Even though the attenuated backscatter data below 100 m were excluded before
 3 assimilation, the DA effects generated by the data above can spread downwards depending on
 4 the background error vertical correlation (Wang et al., 2022), which will improve the surface
 5 PM_{2.5}. As shown in Figure 1, a total of 70 sites were chosen to investigate only the DA effects
 6 on Case 2 PM_{2.5} analyses and forecasts. At the initial time of 18:00 UTC on December 9, 2019,
 7 the control experiment recorded an average PM_{2.5} concentration of 85 $\mu\text{g}/\text{m}^3$, while the
 8 corresponding assimilation experiment exhibited a higher average PM_{2.5} concentration of 107
 9 $\mu\text{g}/\text{m}^3$. The mean observation during this time period was recorded as 151 $\mu\text{g}/\text{m}^3$ (Figure 11).
 10 From the distribution, the simulated PM_{2.5}, over almost all sites in the control experiment, is
 11 underestimated and has great biases from the real heavy pollution band (Figures 12a and 12c).
 12 After assimilation, the simulation is dramatically improved, except at some sites with very high
 13 observations. The simulations over most sites are larger than 100 $\mu\text{g}/\text{m}^3$, which is close to the
 14 real pollution distribution (Figures 12b and 12c). As time progresses, the positive effects of
 15 assimilation on PM_{2.5} analysis can persist, although they gradually diminish and eventually
 16 diminish entirely due to other factors such as emissions and meteorological processes. On
 17 average, the benefits of directly assimilating surface PM_{2.5} concentrations can have a lasting
 18 impact of more than 48 hours in terms of PM_{2.5} forecasts. The assimilation experiment exhibits
 19 enhanced performance in this case due to the substantial improvement in aerosol initial
 20 conditions. However, after integrating 14 h, the assimilation experiment maintains an
 21 equivalent forecasting skill to the control experiment at 08:00 UTC on December 10, 2019 (as
 22 depicted in Figure 11, Figure 12d, 12e), and the trend continues afterwards. Hence, the DA
 23 benefits only last for a short period of ~14 h.

24 In the context of assimilating attenuated backscatters, the DA benefits improve PM_{2.5}
 25 simulations for certain forecasting time. The length of time that DA benefits persist is primarily
 26 influenced by weather conditions and the intensity of emissions. Under weak or moderate
 27 emission intensity, stable weather conditions help DA benefits to continue for a long
 28 forecasting period. For this case study, note that during the forecasting period of 9 to 14 h, the



1 simulations are in a dramatically decreasing process, while the observations change less and
2 remain at high levels. This process leads to poor model performance, at which point the DA
3 benefits rapidly vanish. The dramatic process should be caused by the great change in weather
4 conditions, which can also be inferred from the subsequent reduction of observations (Figure
5 11).

6 For the sake of validation, winds at 10 m that have an important impact on surface $PM_{2.5}$
7 transportation, and distributions at the two times above are also shown in Figure 12. Initially,
8 strong cold air mass from the northern region reached Mongolia and continued moving towards
9 the south. North China is occupied by warm air in front of the cold front, where weak southerly
10 winds of 1–3 m/s occur. Because of the convergence of wind fields, pollutants begin to stack
11 and form a pollution band in North China. At 08:00 UTC on December 10, 2019 (after 14 h),
12 cold air passed over the pollution band, resulting in strong north winds of 6–8 m/s. In such
13 favorable weather conditions, a notable decline in $PM_{2.5}$ concentrations was simulated. The
14 short duration of DA benefits are partly due to weather conditions. Practically, when cold air
15 first arrived, pollution was not immediately alleviated. On the contrary, in the same places such
16 as Tianjin (38.9°N, 117.2°E), pollution was aggravated due to the transportation of pollutants
17 (Figure 12c, 12f). After the complete purification of the air by the cold air, a decline in $PM_{2.5}$
18 concentrations was observed. Furthermore, owing to the narrow spatial coverage of CALIOP
19 ground tracks, aerosol fields undergo "correction" primarily within close proximity to these
20 tracks. Consequently, as the forecast horizon progresses, the influence of advection becomes
21 pronounced, leading to the ingress of aerosol fields unaffected by assimilation into the targeted
22 region. Additionally, the advection process may introduce cleaner air masses, further impacting
23 the evolution of aerosol distributions over time. However, it is noteworthy that the actual
24 decrease in $PM_{2.5}$ concentrations occurred several hours later than what was initially predicted
25 by the simulation, as mentioned earlier. This bias can be ascribed to model uncertainties.

26 4. Conclusions and discussion



1 We have developed a 3DVAR DA algorithm for aerosols, with a focus on satellite lidar-
2 derived attenuated backscatter coefficients. The algorithm is built upon the utilization of the
3 MOSAIC aerosol scheme integrated within the WRF-Chem model. To facilitate the direct
4 assimilation of attenuated backscatter, we have further developed the observation operator.
5 This operator maps the state variables to attenuated backscatter observations, incorporating
6 Mie scattering theory and the lidar equation.

7 A 3DVAR aerosol data assimilation algorithm was developed to directly assimilate 532
8 nm total attenuated backscatter profiles from the CALIPSO Level 1B product into WRF-Chem.
9 The algorithm uses the MOSAIC-4bin aerosol scheme, which represents eight aerosol species
10 in four size bins. An observation operator was developed to map the model aerosol state
11 variables to CALIOP attenuated backscatter. This operator combines Mie-scattering
12 calculations with the lidar equation and includes the contributions of aerosol and molecular
13 scattering. The corresponding tangent-linear and adjoint components were implemented for
14 use in the incremental 3DVAR system.

15 The system was evaluated using two nighttime CALIPSO overpasses over northern China
16 on 6 and 9 December 2019. The two cases represented moderate and heavy aerosol pollution
17 conditions, respectively. The assimilation experiment was compared with a control experiment
18 using the same WRF-Chem configuration but without aerosol data assimilation. The results
19 indicate that assimilating CALIPSO Level 1B attenuated backscatter increased aerosol
20 concentrations in regions where the control simulation underestimated the observations. The
21 resulting attenuated-backscatter fields showed improved agreement with the observed vertical
22 structures, particularly in the elevated aerosol layers. These results provide a diagnostic
23 demonstration that the observation operator and assimilation system can transfer information
24 from vertically resolved lidar observations to the model aerosol state.

25 The assimilation also modified the simulated aerosol extinction and backscatter
26 coefficient profiles. The changes were generally consistent with the observed vertical
27 distribution of aerosol layers, although the magnitude and spatial extent of the increments
28 varied among the selected locations. The largest discrepancies remained near the surface and
29 in regions affected by strong surface returns or incomplete retrieval coverage. These limitations
30 are partly related to the exclusion of observations below 100 m and to the limited vertical



1 propagation of information through the prescribed background-error covariance. Consequently,
2 direct assimilation of the lidar profile does not guarantee accurate correction of the aerosol
3 field throughout the entire boundary layer.

4 The CALIPSO Level 2 extinction, backscatter, and AOD products provided additional
5 diagnostic evaluations. After assimilation, the correlation coefficient of AOD increased from
6 0.51 to 0.58, the RMSE decreased from 0.28 to 0.24, and the mean bias changed from -0.11
7 to 0.01 . These changes are consistent with an improvement in the column-integrated aerosol
8 field. However, the Level 2 products are derived from CALIPSO Level 1 measurements and
9 therefore should not be regarded as fully independent validation data. They are used here to
10 evaluate the consistency of the assimilated aerosol profiles with retrieval-based aerosol optical
11 properties. Independent validation was provided by surface $\text{PM}_{2.5}$ observations for the heavy-
12 pollution case.

13 For the selected surface stations within 50 km of the CALIPSO ground track, the mean
14 simulated $\text{PM}_{2.5}$ concentration at the initial time increased from 85 to $107 \mu\text{g m}^{-3}$ after
15 assimilation, compared with an observed mean of $151 \mu\text{g m}^{-3}$. The positive impact persisted
16 for approximately 14 h in the examined case. The relatively short duration of the benefit was
17 associated with the narrow spatial sampling of the CALIPSO track, changing meteorological
18 conditions, and the subsequent transport and dilution of the corrected aerosol field. In particular,
19 the passage of a cold-air system rapidly changed the near-surface wind field and aerosol
20 distribution. Aerosol information outside the immediate influence region of the CALIPSO
21 track was not directly constrained and could subsequently be advected into the evaluation area.
22 These results indicate that CALIPSO profile assimilation can improve the local aerosol
23 analysis and the short-range forecast.

24 The direct use of Level 1B attenuated backscatter has two potential advantages. It avoids
25 the need to assimilate a retrieved aerosol extinction or backscatter product and therefore avoids
26 introducing some assumptions and screening decisions made during the Level 2 retrieval
27 process. It also retains the high vertical resolution and large data volume of the Level 1B
28 measurements. Nevertheless, the present experiments do not include a parallel assimilation of
29 CALIPSO Level 2 products. Therefore, the relative advantage of Level 1B assimilation over
30 Level 2 assimilation cannot be quantified from the current results. The present study



1 demonstrates the feasibility of the Level 1B approach rather than proving that it provides
2 superior forecast performance.

3 Several limitations should be considered. First, the evaluation is based on only two
4 nighttime CALIPSO overpasses during one winter period. The results may therefore depend
5 on the aerosol composition, cloud conditions, synoptic circulation, and CALIOP sampling
6 geometry of these cases. Second, the assimilated total attenuated backscatter is affected by
7 molecular scattering, aerosol extinction, clouds, surface returns, calibration uncertainty, and
8 signal noise. The current preprocessing, including vertical smoothing, horizontal averaging,
9 threshold filtering, and removal of observations below 100 m, is a preliminary treatment. Third,
10 the observation error was represented using a simplified constant relative uncertainty. In
11 practice, the error varies with altitude, signal strength, aerosol loading, cloud contamination,
12 and the representativeness difference between the CALIOP observation and the model grid cell.
13 Forward-model and representativeness errors should therefore be included in future estimates
14 of the observation-error covariance matrix.

15 Future work should extend the evaluation to a larger set of CALIPSO overpasses covering
16 different seasons, aerosol types, cloud conditions, and meteorological regimes. Direct
17 comparisons between Level 1B and Level 2 assimilation should be conducted to quantify the
18 value of avoiding the Level 2 retrieval step. Sensitivity experiments using different
19 observation-error specifications, background-error correlation scales, vertical smoothing
20 methods, and cloud-screening criteria are also needed. Combining CALIPSO profile
21 observations with surface PM measurements, geostationary-satellite AOD, and ground-based
22 lidar observations may provide more complete spatial and vertical constraints than any single
23 observing system. Such multi-source assimilation could reduce the rapid loss of local forecast
24 benefits caused by the narrow CALIPSO ground track.

25 In summary, the results show that direct assimilation of CALIPSO Level 1B attenuated
26 backscatter into a MOSAIC-based WRF-Chem 3DVAR system is technically feasible and can
27 improve the analyzed vertical distribution of aerosols and short-range PM_{2.5} forecasts near the
28 satellite track. The study provides a proof of concept for using Level 1B lidar measurements
29 in aerosol data assimilation. Broader conclusions concerning operational forecast performance



and the superiority of Level 1B over Level 2 assimilation require additional cases, independent validation, and systematic uncertainty experiments.

Author contributions

The construction of the 3DVAR system and the study design were carried out by Wei You and Zengliang Zang. Wei You supervised the development of the project and provided manuscript revisions. Daichun Wang and Yiwen Hu were responsible for preparing the dataset. Daichun Wang and Wei You conducted the simulations, analyzed the results, and drafted the initial version of the manuscript. Shuang Mei, Wei Zhong, Dan Niu, and Jianping Guo revised the manuscript. Throughout the study, all authors were actively engaged in discussions regarding the results and findings.

Data availability

The NCEP FNL reanalysis data were obtained from the UCAR Research Data Archive (<https://rda.ucar.edu/datasets/ds083.2/>; last accessed on September 5, 2022; DOI: 10.5065/D6M043C). Hourly PM2.5 observations were downloaded from the CNEMC website (<http://www.cnemc.cn/en/>; last accessed on September 5). CALIPSO data were obtained from the Atmospheric Science Data Center of the NASA Langley Research Center.

Code availability

Upon reasonable request, please contact the corresponding author at yw1x_1987@163.com for access to the necessary codes for conducting the analyses.

Competing interests

The authors declare that they have no conflict of interest.



1 Acknowledgements

2 This research was supported by the National Natural Science Foundation of China (Grant No.
 3 42475166) and the Outstanding Youth Foundation of the Natural Science Foundation of Hunan
 4 Province (Grant No. 2023JJ20050).

5 References

- 6 Bocquet, M., Elbern, H., Eskes, H., Hirtl, M., Žabkar, R., Carmichael, G.R., Flemming, J.,
 7 Inness, A., Pagowski, M., Pérez Camaño, J.L., Saide, P.E., San Jose, R., Sofiev, M., Vira,
 8 J., Baklanov, A., Carnevale, C., Grell, G., Seigneur, C., 2015. Data assimilation in
 9 atmospheric chemistry models: current status and future prospects for coupled
 10 chemistry meteorology models. *Atmos. Chem. Phys.* 15, 5325–5358.
 11 <https://doi.org/10.5194/acp-15-5325-2015>
- 12 Chai, T., Carmichael, G. R., Sandu, A., Tang, Y., Daescu, D. N., 2006. Chemical data
 13 assimilation of Transport and Chemical Evolution over the Pacific (TRACE - P)
 14 aircraft measurements. *Journal of Geophysical Research: Atmospheres*, 111(D2).
 15 <https://doi.org/10.1029/2005JD005883>
- 16 Chen, D., Liu, Z., Ban, J., Chen, M., 2019. The 2015 and 2016 wintertime air pollution in
 17 China: SO₂; emission changes derived from a WRF-Chem/EnKF coupled data
 18 assimilation system. *Atmos. Chem. Phys.* 19, 8619–8650. [https://doi.org/10.5194/acp-](https://doi.org/10.5194/acp-19-8619-2019)
 19 [19-8619-2019](https://doi.org/10.5194/acp-19-8619-2019)
- 20 Chen, D., Liu, Z., Fast, J., Ban, J., 2016. Simulations of sulfate-nitrate-ammonium (SNA)
 21 aerosols during the extreme haze events over northern China in October 2014. *Atmos.*
 22 *Chem. Phys.* 16, 10707–10724. <https://doi.org/10.5194/acp-16-10707-2016>
- 23 Chen, F., Dudhia, J., 2001. Coupling an advanced land surface-hydrology model with the Penn
 24 State-NCAR MM5 modeling system. Part I: Model implementation and
 25 sensitivity. *Mon. weather Rev.* 129, 569-585. [https://doi.org/10.1175/1520-](https://doi.org/10.1175/1520-0493(2001)129<0569:CAALSH>2.0.CO;2)
 26 [0493\(2001\)129<0569:CAALSH>2.0.CO;2](https://doi.org/10.1175/1520-0493(2001)129<0569:CAALSH>2.0.CO;2)
- 27 Cheng, X., Liu, Y., Xu, X., You, W., Zang, Z., Gao, L., Chen, Y., Su, D., Yan, P., 2019. Lidar



- 1 data assimilation method based on CRTM and WRF-Chem models and its application
2 in PM_{2.5} forecasts in Beijing. *Sci. Total Environ.* 682, 541–552.
3 <https://doi.org/10.1016/j.scitotenv.2019.05.186>
- 4 Cheng, Y., Dai, T., Goto, D., Schutgens, N.A.J., Shi, G., Nakajima, T., 2019. Investigating the
5 assimilation of CALIPSO global aerosol vertical observations using a four-dimensional
6 ensemble Kalman filter. *Atmos. Chem. Phys.* 19, 13445–13467.
7 <https://doi.org/10.5194/acp-19-13445-2019>
- 8 Cheng, Y., Dai, T., Cao, J., Chen, L., Goto, D., Yoshida, M., Nakajima, T., Shi, G.,
9 2022. Improvement of the aerosol forecast and analysis over East Asia with joint
10 assimilation of two geostationary satellite observations. *Geophysical Research*
11 *Letters*, 49, e2022GL099908. <https://doi.org/10.1029/2022GL099908>
- 12 Chin, M., Savoie, D.L., Huebert, B.J., Bandy, A.R., Thornton, D.C., Bates, T.S., Quinn, P.K.,
13 Saltzman, E.S., De Bruyn, W.J., 2000. Atmospheric sulfur cycle simulated in the global
14 model GOCART: Comparison with field observations and regional budgets. *J. Geophys.*
15 *Res.* 105, 24689–24712. <https://doi.org/10.1029/2000JD900385>
- 16 Dai, T., Cheng, Y., Suzuki, K., Goto, D., Kikuchi, M., Schutgens, N.A.J., Yoshida, M., Zhang,
17 P., Husi, L., Shi, G., Nakajima, T., 2019. Hourly Aerosol Assimilation of Himawari - 8
18 AOT Using the Four - Dimensional Local Ensemble Transform Kalman Filter. *J. Adv.*
19 *Model. Earth Syst.* 11, 680 – 711. <https://doi.org/10.1029/2018MS001475>
- 20 Fast, J.D., Gustafson, W.I., Easter, R.C., Zaveri, R.A., Barnard, J.C., Chapman, E.G., Grell,
21 G.A., Peckham, S.E., 2006. Evolution of ozone, particulates, and aerosol direct
22 radiative forcing in the vicinity of Houston using a fully coupled meteorology-
23 chemistry-aerosol model. *J. Geophys. Res.* 111, D21305.
24 <https://doi.org/10.1029/2005JD006721>
- 25 Feng, S., Jiang, F., Jiang, Z., Wang, H., Cai, Z., Zhang, L., 2018. Impact of 3DVAR assimilation
26 of surface PM_{2.5} observations on PM_{2.5} forecasts over China during wintertime. *Atmos.*
27 *Environ.* 187, 34–49. <https://doi.org/10.1016/j.atmosenv.2018.05.049>
- 28 Fernald, F.G., 1984. Analysis of atmospheric lidar observations: some comments. *Appl. Opt.*
29 23, 652. <https://doi.org/10.1364/AO.23.000652>



- 1 Gao, L., Jiang, L., Sun, W., Yan, P., Qi, B., Ji, C., Tao, F. (2025). Data assimilation of
2 ceilometer aerosol extinction coefficient profile contributes to predictions of the three-
3 dimensional structures of aerosols in East China. *Journal of Geophysical Research:*
4 *Atmospheres*, 130, e2024JD042408. <https://doi.org/10.1029/2024JD042408>
- 5 Grell, G.A., Peckham, S.E., Schmitz, R., McKeen, S.A., Frost, G., Skamarock, W.C., Eder, B.,
6 2005. Fully coupled “online” chemistry within the WRF model. *Atmos. Environ.* 39,
7 6957–6975. <https://doi.org/10.1016/j.atmosenv.2005.04.027>
- 8 Grell, G. A., Freitas, S. R., 2014. A scale and aerosol aware stochastic convective
9 parameterization for weather and air quality modeling. *Atmos. Chem.*
10 *Phys.* 14, 5233– 5250. <https://doi.org/10.5194/acp-14-5233-2014>
- 11 Ha, S., Liu, Z., Sun, W., Lee, Y., Chang, L., 2019. Improving air quality forecasting with the
12 assimilation of GOCI AOD retrievals during the KORUS-AQ period (preprint). *Atmos.*
13 *Chem. Phys.* 20, 6015-6036. <https://doi.org/10.5194/acp-2019-648>
- 14 Hong S. Y., Lim J. O. J., 2006. The WRF single-moment 6-class microphysics scheme
15 (WSM6). *Journal of Korean Meteorological Society.* 42(2): 129-151.
- 16 Hong, J., Chang, W., Han, W., Zhuang, Z., Wang, H., Bi, L., Shen, X., 2026. Data assimilation
17 of aerosol optical depth (AOD) in CMA-MESO/CUACE model to improve weather
18 predictions in China. *Atmospheric Research*, 108780.
19 <https://doi.org/10.1016/j.atmosres.2026.108780>
- 20 Hu, F., Mao, F., Zhang, Y., Hong, J., Zang, L., Zeng, Z., Lin, S., Gong, W., Rosenfeld, D., 2026.
21 Assimilating FY-4B satellite aerosol data to improve PM_{2.5} and surface shortwave
22 radiation prediction. *Atmospheric Research*, 108764.
23 <https://doi.org/10.1016/j.atmosres.2026.108764>
- 24 Hu, Y., Zang, Z., Ma, X., Li, Y., Liang, Y., You, W., Pan, X., Li, Z., 2022. Four-dimensional
25 variational assimilation for SO₂ emission and its application around the COVID-19
26 lockdown in the spring 2020 over China. *Atmos. Chem. Phys.* 22, 13183–13200.
27 <https://doi.org/10.5194/acp-22-13183-2022>
- 28 Iacono, M. J., Delamere, J. S., Mlawer, E. J., Shephard, M. W., Clough, S. A., Collins, W. D.,
29 2008. Radiative forcing by long - lived greenhouse gases: Calculations with the AER



- 1 radiative transfer models. *Journal of Geophysical Research: Atmospheres*, 113(D13).
- 2 <https://doi.org/10.1029/2008JD009944>
- 3 Jiang, Z., Liu, Z., Wang, T., Schwartz, C.S., Lin, H.-C., Jiang, F., 2013. Probing into the impact
- 4 of 3DVAR assimilation of surface PM₁₀ observations over China using process analysis.
- 5 *J. Geophys. Res. Atmos.* 118, 6738–6749. <https://doi.org/10.1002/jgrd.50495>
- 6 Jing, B., Wu, L., Mao, H., Gong, S., He, J., Zou, C., Song, G., Li, X., Wu, Z., 2016.
- 7 Development of a vehicle emission inventory with high temporal-spatial resolution
- 8 based on NRT traffic data and its impact on air pollution in Beijing-Part 1: Development
- 9 and evaluation of vehicle emission inventory. *Atmos. Chem. Phys.* 16, 3161–3170.
- 10 <https://doi.org/10.5194/acp-16-3161-2016>
- 11 Li, Z., Zang, Z., Li, Q.B., Chao, Y., Chen, D., Ye, Z., Liu, Y., Liou, K.N., 2013. A three-
- 12 dimensional variational data assimilation system for multiple aerosol species with
- 13 WRF/Chem and an application to PM_{2.5} prediction. *Atmos. Chem. Phys.* 13, 4265–4278.
- 14 <https://doi.org/10.5194/acp-13-4265-2013>
- 15 Liang, Y., Zang, Z., Liu, D., Yan, P., Hu, Y., Zhou, Y., You, W., 2020. Development of a three-
- 16 dimensional variational assimilation system for lidar profile data based on a size-
- 17 resolved aerosol model in WRF-Chem model v3.9.1 and its application in PM_{2.5}
- 18 forecasts across China. *Geosci. Model Dev.* 13, 6285–6301.
- 19 <https://doi.org/10.5194/gmd-13-6285-2020>
- 20 Liu, Z., Liu, Q., Lin, H.C., Schwartz, C.S., Lee, Y.H., Wang, T., 2011. Three-dimensional
- 21 variational assimilation of MODIS aerosol optical depth: Implementation and
- 22 application to a dust storm over East Asia. *J. Geophys. Res.* 116, D23206.
- 23 <https://doi.org/10.1029/2011JD016159>
- 24 Ma, S., Zhang, X., Gao, C., Tong, D.Q., Xiu, A., Wu, G., Cao, X., Huang, L., Zhao, H., Zhang,
- 25 S., Ibarra-Espinosa, S., Wang, X., Li, X., Dan, M., 2019. Multimodel simulations of a
- 26 springtime dust storm over northeastern China: implications of an evaluation of four
- 27 commonly used air quality models (CMAQ v5.2.1, CAMx v6.50, CHIMERE v2017r4,
- 28 and WRF-Chem v3.9.1). *Geosci. Model Dev.* 12, 4603–4625.
- 29 <https://doi.org/10.5194/gmd-12-4603-2019>
- 30 Mamouri, R.E., Amiridis, V., Papayannis, A., Giannakaki, E., Tsaknakis, G., Balis, D.S., 2009.



- 1 Validation of CALIPSO space-borne-derived aerosol vertical structures using a ground-
2 based lidar in Athens, Greece. *Atmos. Meas. Tech. Discuss.* 2, 561-587.
3 <https://doi.org/10.5194/amtd-2-561-2009>
- 4 Menon, S., Hansen, J., Nazarenko, L., Luo, Y., 2002. Climate Effects of Black Carbon Aerosols
5 in China and India. *Science* 297, 2250–2253. <https://doi.org/10.1126/science.1075159>
- 6 Niu, T., Gong, S. L., Zhu, G. F., Liu, H. L., Hu, X. Q., Zhou, C. H., Wang, Y. Q., 2008. Data
7 assimilation of dust aerosol observations for the CUACE/dust forecasting
8 system. *Atmospheric Chemistry and Physics*, 8(13), 3473-3482.
9 <https://doi.org/10.5194/acp-8-3473-2008>
- 10 Noel, V., Chepfer, H., Hoareau, C., Reverdy, M., Cesana, G., 2014. Effects of solar activity on
11 noise in CALIOP profiles above the South Atlantic Anomaly. *Atmos. Meas. Tech.* 7,
12 1597–1603. <https://doi.org/10.5194/amt-7-1597-2014>
- 13 Pagowski, M., Grell, G.A., McKeen, S.A., Peckham, S.E., Devenyi, D., 2010. Three-
14 dimensional variational data assimilation of ozone and fine particulate matter
15 observations: some results using the Weather Research and Forecasting-Chemistry
16 model and Grid-point Statistical Interpolation. *Q. J. R. Meteorol. Soc.* 136, 2013–2024.
17 <https://doi.org/10.1002/qj.700>
- 18 Pang, M., Jin, J., Segers, A., Jiang, H., Fang, L., Lin, H. X., Liao, H., 2023. Dust storm
19 forecasting through coupling LOTOS-EUROS with localized ensemble Kalman
20 filter. *Atmos. Environ.* 306, 119831. <https://doi.org/10.1016/j.atmosenv.2023.119831>
- 21 Ruggaber, A., Dlugi, R., Nakajima, T., 1994. Modelling radiation quantities and photolysis
22 frequencies in the troposphere, *J. Atmos. Chem.*, 18, 171–210,
23 <https://doi.org/10.1007/BF00696813>
- 24 Rogers, R.R., Hostetler, C.A., Hair, J.W., Ferrare, R.A., Liu, Z., Obland, M.D., Harper, D.B.,
25 Cook, A.L., Powell, K.A., Vaughan, M.A., Winker, D.M., 2011. Assessment of the
26 CALIPSO Lidar 532 nm attenuated backscatter calibration using the NASA LaRC
27 airborne High Spectral Resolution Lidar. *Atmos. Chem. Phys.* 11, 1295–1311.
28 <https://doi.org/10.5194/acp-11-1295-2011>
- 29 Saide, P.E., Kim, J., Song, C.H., Choi, M., Cheng, Y., Carmichael, G.R., 2014. Assimilation of
30 next generation geostationary aerosol optical depth retrievals to improve air quality



- 1 simulations. Geophys. Res. Lett. 41, 9188–9196.
- 2 <https://doi.org/10.1002/2014GL062089>
- 3 Schwartz, C.S., Liu, Z., Lin, H.-C., McKeen, S.A., 2012. Simultaneous three-dimensional
- 4 variational assimilation of surface fine particulate matter and MODIS aerosol optical
- 5 depth. J. Geophys. Res. 117, D13202. <https://doi.org/10.1029/2011JD017383>
- 6 Sekiyama, T. T., Tanaka, T. Y., Miyoshi, T., 2012. A simulation study of the ensemble-based
- 7 data assimilation of satellite-borne lidar aerosol observations. Geosci. Model Dev.
- 8 Discuss. 5(3), 1877–1947, <https://doi.org/10.5194/gmdd-5-1877-2012>
- 9 Sekiyama, T.T., Tanaka, T.Y., Shimizu, A., Miyoshi, T., 2010. Data assimilation of CALIPSO
- 10 aerosol observations. Atmos. Chem. Phys. 10, 39–49. [https://doi.org/10.5194/acp-10-](https://doi.org/10.5194/acp-10-39-2010)
- 11 [39-2010](https://doi.org/10.5194/acp-10-39-2010)
- 12 Sun, W., Liu, Z., Chen, D., Zhao, P., Chen, M., 2020. Development and application of the
- 13 WRFDA-Chem three-dimensional variational (3DVAR) system: aiming to improve air
- 14 quality forecasting and diagnose model deficiencies. Atmos. Chem. Phys. 20, 9311–
- 15 9329. <https://doi.org/10.5194/acp-20-9311-2020>
- 16 Stromatas, S., Turquety, S., Menut, L., Chepfer, H., Pere, J. C., Cesana, G., Bessagnet, B., 2012.
- 17 Lidar signal simulation for the evaluation of aerosols in chemistry transport
- 18 models. Geosci. Model Dev. 5(6), 1543–1564. [https://doi.org/10.5194/gmd-5-1543-](https://doi.org/10.5194/gmd-5-1543-2012)
- 19 [2012](https://doi.org/10.5194/gmd-5-1543-2012)
- 20 Tombette, M., Mallet, V., Sportisse, B., 2009. PM₁₀ data assimilation over Europe with the
- 21 optimal interpolation method. Atmos. Chem. Phys. 9, 57–70.
- 22 <https://doi.org/10.5194/acp-9-57-2009>
- 23 Tsikerdekis, A., Schutgens, N. A. J., Hasekamp, O. P., 2021. Assimilating aerosol optical
- 24 properties related to size and absorption from POLDER/PARASOL with an ensemble
- 25 data assimilation system, Atmos. Chem. Phys., 21, 2637–2674,
- 26 <https://doi.org/10.5194/acp-21-2637-2021>, 2021.
- 27 Wang, D., You, W., Zang, Z., Pan, X., Hu, Y., Liang, Y., 2022. A three-dimensional variational
- 28 data assimilation system for aerosol optical properties based on WRF-Chem v4.0:
- 29 design, development, and application of assimilating Himawari-8 aerosol observations.
- 30 Geosci. Model Dev. 15, 1821–1840. <https://doi.org/10.5194/gmd-15-1821-2022>



- 1 Wang, D., You, W., Zang, Z., Pan, X., He, H., Liang, Y., 2020. A three-dimensional variational
2 data assimilation system for a size-resolved aerosol model: Implementation and
3 application for particulate matter and gaseous pollutant forecasts across China. *Science*
4 *China Earth Sciences*. 63, 1366-1380. <https://doi.org/10.1007/s11430-019-9601-4>
- 5 Wang, Y., Sartelet, K. N., Bocquet, M., and Chazette, P., 2014. Modelling and assimilation of
6 lidar signals over Greater Paris during the MEGAPOLI summer campaign, *Atmos.*
7 *Chem. Phys.*, 14, 3511-3532, <https://doi.org/10.5194/acp-14-3511-2014>
- 8 Winker, D.M., Hunt, W.H., McGill, M.J., 2007. Initial performance assessment of CALIOP.
9 *Geophys. Res. Lett.* 34, L19803. <https://doi.org/10.1029/2007GL030135>
- 10 Xia, X., Min, J., Shen, F., Wang, Y., Xu, D., Yang, C., Zhang, P., 2020. Aerosol data
11 assimilation using data from Fengyun-4A, a next-generation geostationary
12 meteorological satellite. *Atmospheric Environment*, 237, 117695.
13 <https://doi.org/10.1016/j.atmosenv.2020.117695>
- 14 Ye, H., Pan, X., You, W., Zhu, X., Zang, Z., Wang, D., Zhang, X., Hu, Y., Jin, S., 2021. Impact
15 of CALIPSO profile data assimilation on 3-D aerosol improvement in a size-resolved
16 aerosol model. *Atmos. Res.* 264, 105877.
17 <https://doi.org/10.1016/j.atmosres.2021.105877>
- 18 Yu, M., Tang, G., Yang, Y., Li, Q., Wang, Yonghong, Miao, S., Zhang, Y., Wang, Yuesi, 2020.
19 The interaction between urbanization and aerosols during a typical winter haze event in
20 Beijing. *Atmos. Chem. Phys.* 20, 9855–9870. [https://doi.org/10.5194/acp-20-9855-](https://doi.org/10.5194/acp-20-9855-2020)
21 [2020](https://doi.org/10.5194/acp-20-9855-2020)
- 22 Yumimoto, K., Nagao, T.M., Kikuchi, M., Sekiyama, T.T., Murakami, H., Tanaka, T.Y., Ogi,
23 A., Irie, H., Khatri, P., Okumura, H., Arai, K., Morino, I., Uchino, O., Maki, T., 2016.
24 Aerosol data assimilation using data from Himawari - 8, a next - generation
25 geostationary meteorological satellite. *Geophys. Res. Lett.* 43, 5886–5894.
26 <https://doi.org/10.1002/2016GL069298>
- 27 Zang, Z., You, W., Ye, H., Liang, Y., Li, Y., Wang, D., Hu, Y., Yan, P., 2022. 3DVAR Aerosol
28 Data Assimilation and Evaluation Using Surface PM_{2.5}, Himawari-8 AOD and
29 CALIPSO Profile Observations in the North China. *Remote Sens.* 14, 4009.



- 1 <https://doi.org/10.3390/rs14164009>
- 2 Zaveri, R.A., Easter, R.C., Fast, J.D., Peters, L.K., 2008. Model for Simulating Aerosol
- 3 Interactions and Chemistry (MOSAIC). J. Geophys. Res. 113, D13204.
- 4 <https://doi.org/10.1029/2007JD008782>
- 5 Zheng, B., Tong, D., Li, M., Liu, F., Hong, C., Geng, G., Li, H., Li, X., Peng, L., Qi, J., Yan,
- 6 L., Zhang, Y., Zhao, H., Zheng, Y., He, K., Zhang, Q., 2018. Trends in China's
- 7 anthropogenic emissions since 2010 as the consequence of clean air actions. Atmos.
- 8 Chem. Phys. 19, 14095-14111. <https://doi.org/10.5194/acp-2018-374>



Table 1. WRF-Chem model configuration

Physical or chemical process	Option
Microphysics	Single-moment six-class (WSM6, Hong and Lim, 2006)
Longwave radiation	Radiative Transfer Model for General Circulation Model (Iacono et al., 2008)
Shortwave radiation	Radiative Transfer Model for General Circulation Model (Iacono et al., 2008)
Boundary layer scheme	Yonsei University (Hong et al., 2006)
Land surface model	Noah land surface model (Chen & Dudhia, 2001)
Convective parameterization	Grell-Freitas scheme (Grell & Freitas, 2014)
Photolysis scheme	Fast-J (Ruggaber et al., 1994)
Aerosol scheme	Model for Simulating Aerosol Interactions and Chemistry (MOSAIC-4 bin) (Zaveri et al., 2008)

1

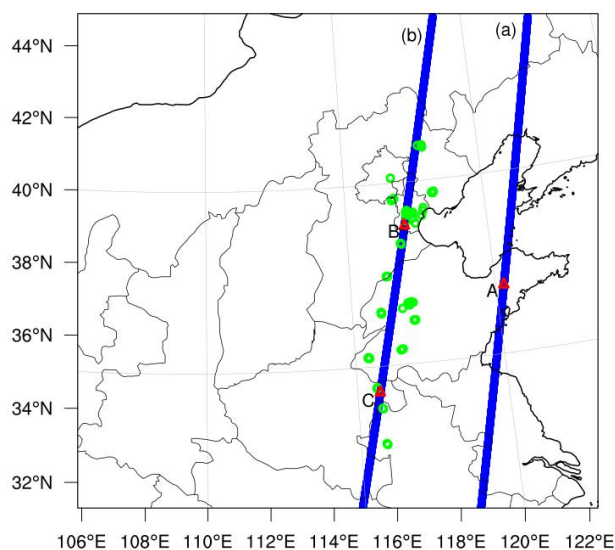


Figure 1. The model inner domain (D02) map covering North China. Two blue lines represent CALIPSO ground tracks marked with (a) and (b), obtained at about 18:00 UTC on December 6 and 9, 2019, respectively. Seventy green circles represent surface PM sites less than 50 km from the track (b). Three red triangles locate at A (36.911°N, 120.287°E) on track (a), B (38.842°N, 117.069°E), and C (34.333°N, 115.747°E) on track (b).

1

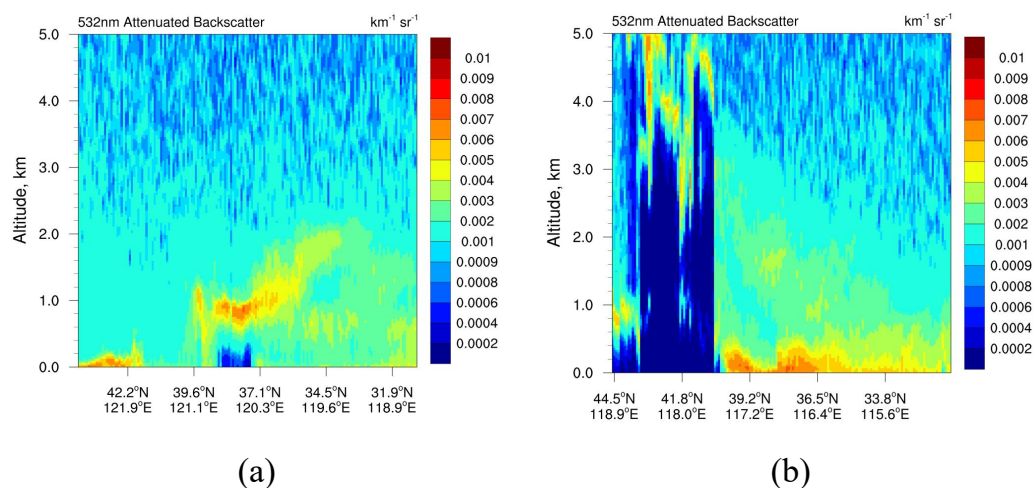


Figure 2. CALIPSO Level 1B 532 nm total attenuated backscatter coefficients across D02, obtained at about 18:00 UTC on December 6, 2019 (a), denoted as Case 1 and about 18:00 UTC on December 9, 2019 (b), denoted as Case 2, which correspond to tracks (a) and (b) in Figure 1, respectively.

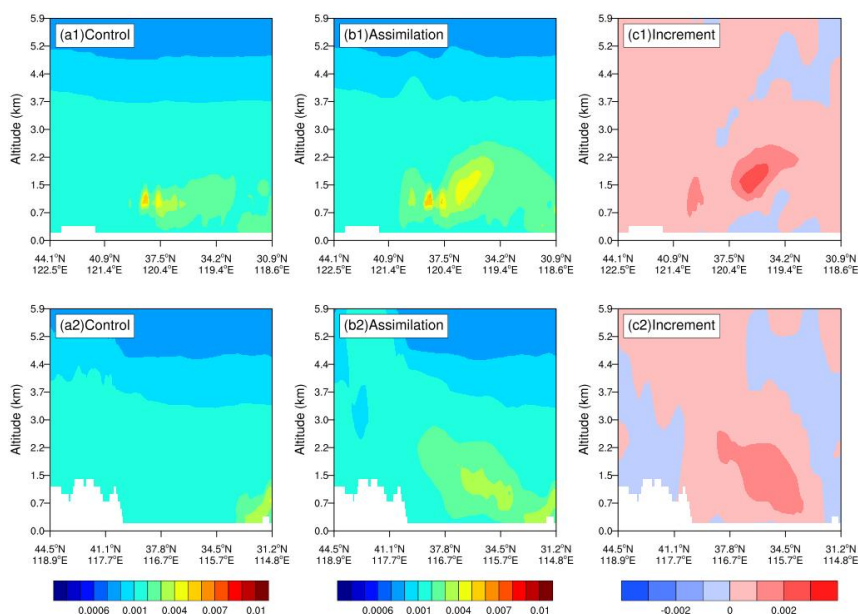


Figure 3. Attenuated backscatter cross-sections at 532 nm simulated by the control experiment (a1, a2), the assimilation experiment (b1, b2), and the DA increment that is assimilation result minus control result (c1, c2), respectively. The upper row corresponds to Case 1. The lower row corresponds to Case 2.

3

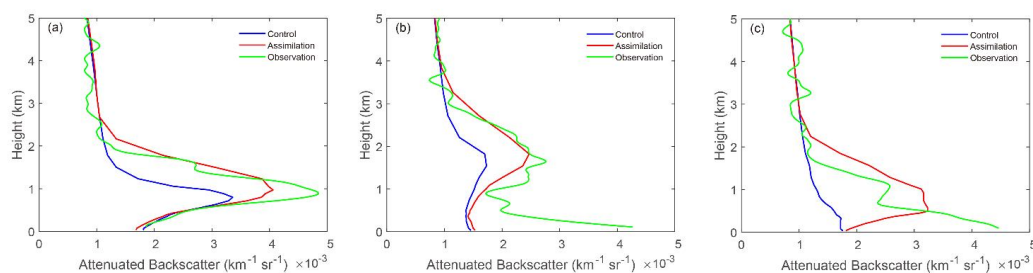


Figure 4. Simulated attenuated backscatter profiles against observations at three A (a), B (b), and C (c) points located in Figure 1. The red, blue, and green lines represent the assimilation experiment, the control experiment, and the observation, respectively (similar below).

4

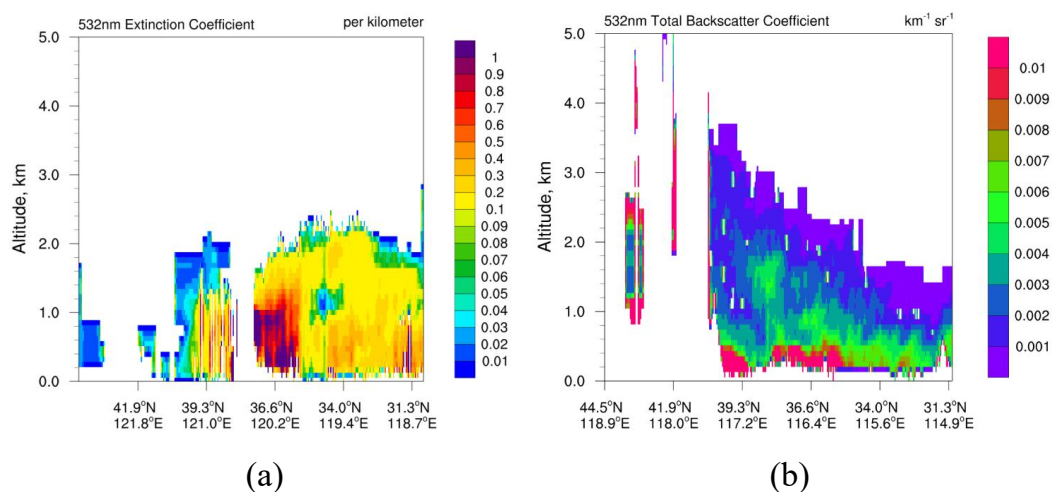


Figure 5. Case 1 CALIPSO Level 2 532 nm extinction coefficients (a) and Case 2 CALIPSO Level 2 532 nm backscatter coefficients (b).

5

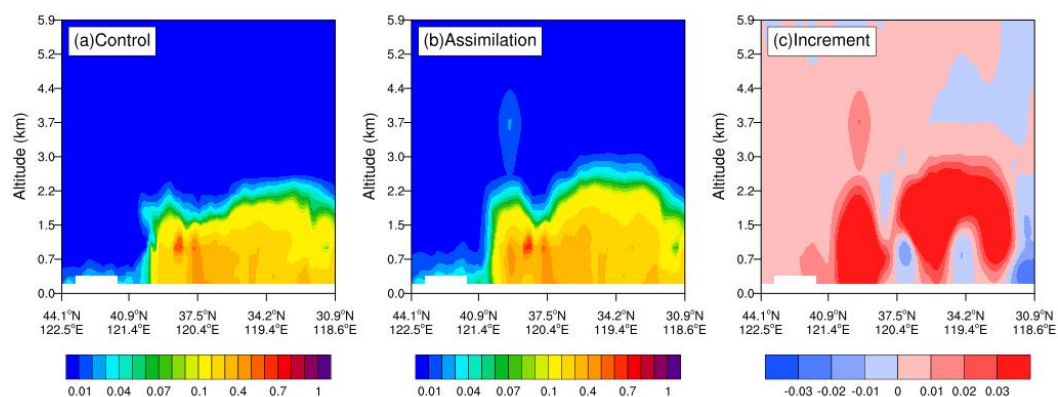


Figure 6. Aerosol extinction coefficient cross-sections at 532 nm simulated by the control experiment (a), the assimilation experiment (b), and the difference between them (c) in Case 1.

6

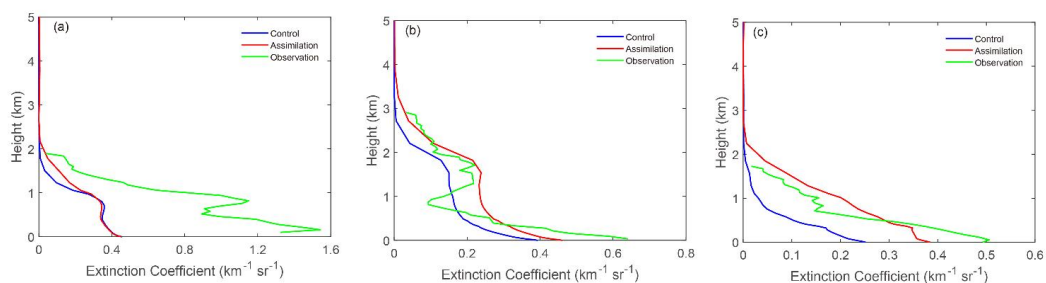


Figure 7. Simulated extinction coefficient profiles against observations at A (a), B (b), and C (c) points.

7

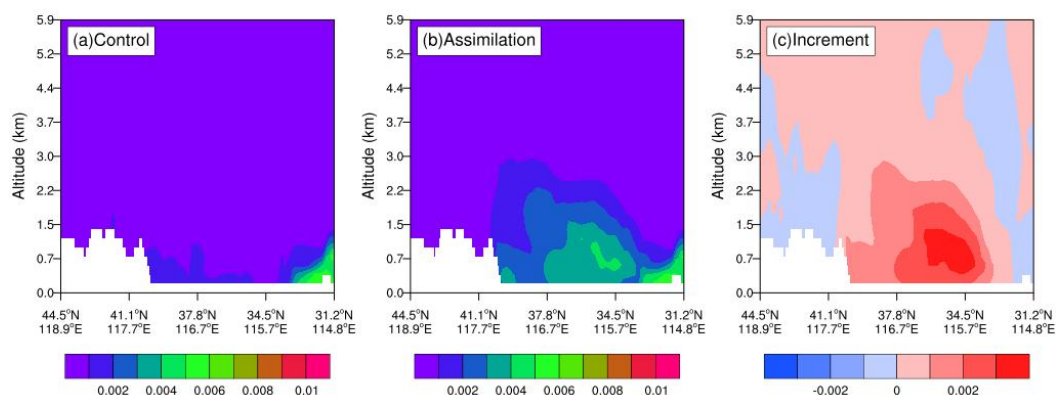


Figure 8. Aerosol backscatter coefficient cross-sections at 532 nm simulated by the control experiment (a), the assimilation experiment (b), and the difference between them (c) in Case 2.

8

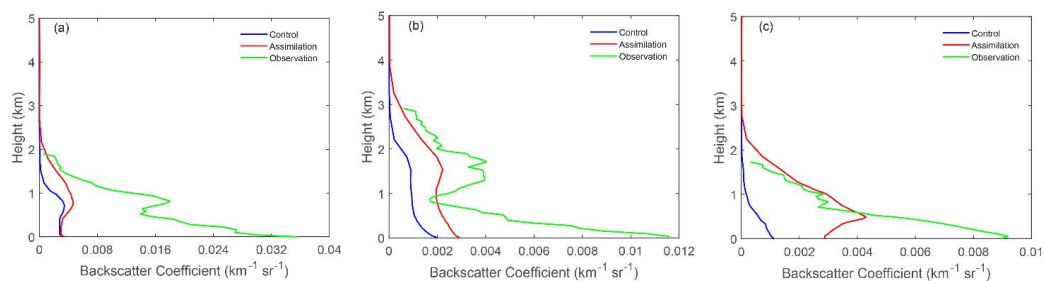


Figure 9. Simulated backscatter coefficient profiles against observations at A (a), B (b), and C (c) points.

9

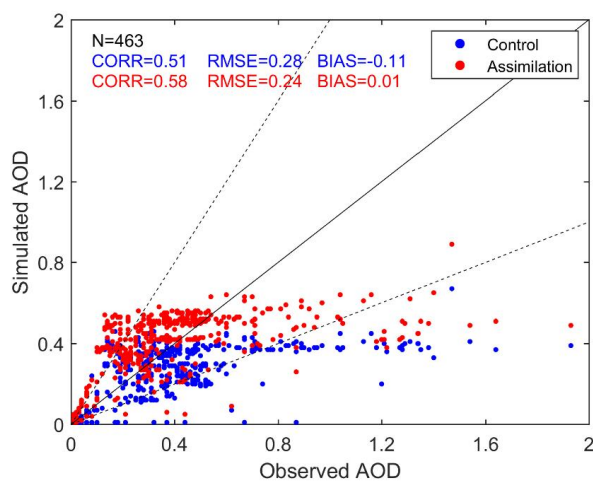


Figure 10. Scatter plot of CALIPSO Level 2 532 nm AODs in both cases against simulated AODs.

The control experiment is illustrated in blue while the assimilation experiment is illustrated in red.

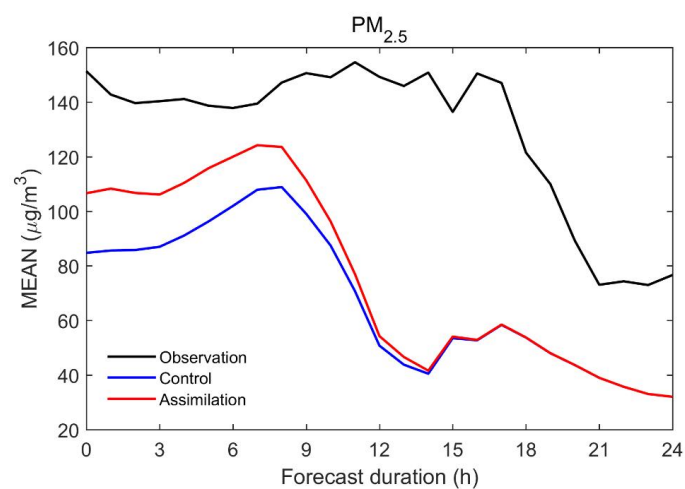


Figure 11. Variations of surface PM_{2.5} mass concentrations averaged by selected sites (70 sites) in the Case 2 forecasting range, the black line represents observations, while the red and blue lines represent the simulations by the assimilation experiment and the control experiment, respectively.

11

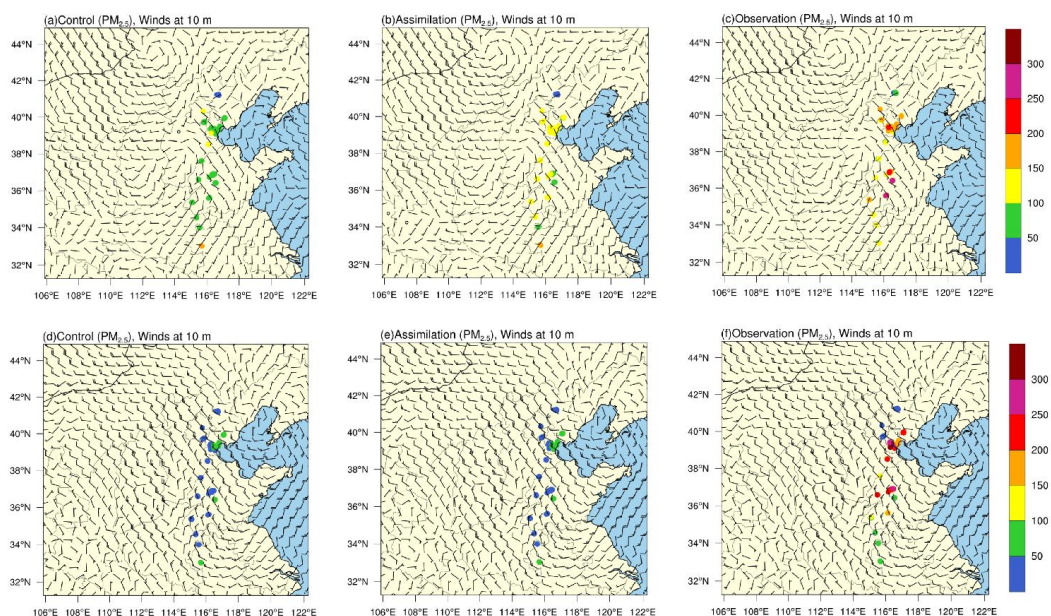


Figure 12. Simulated surface $PM_{2.5}$ distributions ($\mu g m^{-3}$) over selected sites in the control experiment (a, d) and the assimilation experiment (b, e), and the observations (c, f) at two time, respectively; one is the initial time, 18:00 UTC on December 9, 2019 (a, b, c), and the other is at 08:00 UTC on December 10, 2019 (d, e, f). Also, winds at 10 m (m/s) at the two above time are shown, which were produced by WRF-Chem.