https://doi.org/10.5194/essd-2026-677
Preprint. Discussion started: 23 September 2026

Open Access Earth System Science Data Discussions

A dataset of high–spatiotemporal–resolution dust and non–dust aerosol mass concentration profiles from ground–based remote sensing in central China (2018–2024)

Zhuang Wang[1,2,3], Chengxiao Liu[1,2,3], Yanfeng Huo[1,2,3], Xinfeng Lin[1,2,3], Shaowei Yan[1,2,3], Xiaoyun Sun[2,3], Jiaqi Chen[1,2,3], Nan Ge[1,2,3], Guanyin Yang[1,2,3], Zhenzhen Hua[1,2,3], Hao Zhang[2,3], Chune Shi[2,3], Yong Zhu[1], Yizhi Zhu[7], Congzi Xia[8], Kaidi Zhang[1,2,3], Xintong Chen[2,3], Yujia Chen[2,3*], Chengzhi Xing[4*], Cheng Liu[4,5,6*]

[1]Hefei Meteorological Quantum Technology Innovation Research Center, Hefei, Anhui 23008
[2]Anhui Province Key Laboratory of Atmospheric Science and Satellite Remote Sensing, Anhui Institute of Meteorological Sciences, Hefei 230031, China
[3]Shouxian National Climatology Observatory, Huaihe River Basin Typical Farm Eco-meteorological Experiment Field of CMA, Shouxian 232200, China
[4]Key Lab of Environmental Optics and Technology, Anhui Institute of Optics and Fine Mechanics, Hefei Institutes of Physical Science, Chinese Academy of Sciences, Hefei 230031, China
[5]Department of Precision Machinery and Precision Instrumentation, University of Science and Technology of China, Hefei, 230026, China.
[6]Key Laboratory of Precision Scientific Instrumentation of Anhui Higher Education Institutes, University of Science and Technology of China, Hefei, 230026, China.
[7]School of Environmental Science and Engineering, Suzhou University of Science and Technology, Suzhou 215009, China
[8]Institute of Big Data for Vocational Education, Guangdong Polytechnic of Science and Technology, Zhuhai 519000, China

Correspondence to: Yujia Chen (chenyj18@mail.ustc.edu.cn), Chengzhi Xing (xingcz@aiofm.ac.cn), Cheng Liu (chliu81@ustc.edu.cn)

Abstract. The atmospheric aerosol composition and vertical structure over eastern China are jointly influenced by mineral dust transport, anthropogenic emissions, and boundary layer processes; however, long–term vertically resolved observations that distinguish dust from non–dust aerosol mass remain limited. Here, we present a quality–controlled dataset of dust, non–dust, and total aerosol mass concentration profiles derived from continuous ground–based polarization lidar observations at the Shouxian National Climate Observatory (32.434° N, 116.793° E) in central China from January 2018 to December 2024. The site lies in a regional transport corridor between the Beijing–Tianjin–Hebei and Yangtze River Delta regions and is therefore well–suited for monitoring aerosol exchange between two major source–receptor systems in eastern China. Raw lidar measurements were screened for clouds, precipitation, and instrumental artifacts, harmonized to 15 min and 7.5 m resolution, and processed with the Fernald inversion and Polarization lidar photometer networking (POLIPHON) framework to retrieve vertically resolved dust and non–dust aerosol mass concentrations from approximately 0.3 to 6 km. Monthly extinction–to–mass conversion factors were constrained using collocated CE–318 sun–photometer observations. The product was evaluated against independent column, profile, and near–surface references, including CE–318 aerosol optical depth, CALIPSO aerosol extinction and depolarization profiles, and surface particulate matter measurements from a colocated Grimm monitor. The comparisons indicate that the dataset reproduces the main temporal variability of column aerosol loading, captures the principal

https://doi.org/10.5194/essd-2026-677
Preprint. Discussion started: 23 September 2026

Open Access Earth System Science Data Discussions

vertical structure of aerosol extinction coefficient, and is statistically consistent with near–surface particulate matter variability. A pointwise retrieval quality index (RQI) ranging from 0 to 100 is provided to facilitate data screening and reuse, with RQI values of 75–100 recommended for quantitative analyses and event studies, 50–75 for routine statistical analyses, and lower values mainly for qualitative interpretation or analyses based on temporal or vertical averaging. This dataset provides a multi–year, high–resolution observational record of aerosol type–resolved mass profiles in central China and is intended to support community reuse in satellite validation, model evaluation, aerosol transport studies, boundary layer research, and data assimilation applications. The dataset is available for free at Mendeley Data (https://doi.org/10.17632/kfvpbb3vcs.1; Wang et al., 2026).

1 Introduction

Atmospheric aerosols influence radiative transfer, cloud microphysics, air quality, and climate, and their effects depend strongly on the particle composition, vertical distribution, and mixing state (Adachi et al., 2025; Liu et al., 2025). Over central China, aerosol loading is shaped by the combined effects of mineral dust transport, anthropogenic emissions, monsoonal circulation, and boundary layer dynamics (Wang et al., 2020; Wang et al., 2023b). In such environments, column–integrated or surface–only measurements are often insufficient to determine whether aerosol mass is confined to the planetary boundary layer (PBL), transported aloft, or composed of dust and non–dust particles with distinct optical and mass properties (Ryan et al., 2025; Jing et al., 2025; Julaha et al., 2025; Teri et al., 2025). Long–term vertically resolved observations are therefore needed to quantify aerosol transport pathways, evaluate the vertical structure represented by satellite and model products, and provide process–level constraints for studies of aerosol meteorology interactions (Patel et al., 2024; Yorks et al., 2023; Wang et al., 2023a).

Widely used aerosol observation systems provide only a part of this information. Sun–photometer networks and other passive remote sensing products offer long records of column–integrated optical properties (Holben et al., 1998; Ozdemir et al., 2020); however, they do not resolve aerosol layering. Surface particulate matter monitoring networks provide dense temporal coverage near the ground (Wang et al., 2025; Zhang et al., 2019); however, they cannot detect elevated transport layers or describe the full lower tropospheric structure of aerosol events. Spaceborne lidar has greatly advanced global aerosol profiling; however, orbital sampling and reduced sensitivity near the surface limit its ability to continuously represent sub–daily variability at a fixed site (Huang et al., 2015; Kim et al., 2018b; Wang et al., 2021b). Consequently, long–term ground–based datasets that combine continuous profiling with aerosol type separation remain comparatively limited in central and eastern China.

Previous studies have demonstrated that aerosol pollution over the North China Plain (NCP), including the Beijing–Tianjin–Hebei (BTH) region and Yangtze River Delta (YRD), is dynamically interconnected rather than regionally isolated. Huang et al. (2020) showed that aerosol–boundary layer interactions can amplify transboundary haze transport from the NCP to the YRD over distances of approximately 1000 km. Zhang et al. (2021) documented the wintertime transport of haze particles

https://doi.org/10.5194/essd-2026-677
Preprint. Discussion started: 23 September 2026

Open Access Earth System Science Data Discussions

from the NCP to the YRD under cold frontal condition. Using vertically resolved MAX–DOAS observations, Song et al. (2023) identified bidirectional transport between the NCP and YRD, with YRD–to–NCP transport occurring mainly within the 500–1500 m layer and NCP–to–YRD transport occurring primarily below approximately 1000 m. Recently, Yan et al. (2024) showed that the intervening corridor experiences frequent haze events associated with alternating regional transport and atmospheric stagnation. Together, these studies indicate that the region between the NCP and YRD is not merely a geographical transition zone but an active source–receptor corridor in which aerosol plumes can be transported, accumulated, vertically redistributed, and chemically transformed.

Therefore, this transport corridor is a strategic location for continuous vertically resolved observations that can distinguish near–surface accumulation from elevated transport and capture aerosol plumes moving between northern and eastern China (Yan et al., 2024). The Shouxian National Climate Observatory (SXNCO; 32.434° N, 116.793° E) is situated near the northwestern margin of the YRD along a major transport corridor linking the BTH and YRD regions. The rural surroundings of SXNCO and relatively weak local stationary emissions make it well suited for detecting the regional evolution of transported aerosol plumes, including spring dust intrusions and southward and northward transport of anthropogenic pollution (Zhang et al., 2021). Therefore, a long–term, vertically resolved record at this site can provide a regionally representative benchmark for tracking aerosol structures across a key transition zone in eastern China (Yan et al., 2020).

Here, we present a quality–controlled dataset of dust, non–dust, and total aerosol mass concentration profiles derived from continuous ground–based polarization lidar observations at SXNCO from January 2018 to December 2024. The dataset combines seven years of routine measurements with a consistent retrieval framework based on the Fernald inversion and Polarization lidar photometer networking (POLIPHON) method (Fernald, 1984; Ansmann et al., 2012; Mamouri and Ansmann, 2014; Tesche et al., 2017), together with sun–photometer constraints on extinction–to–mass conversion (Mamouri and Ansmann, 2017; Ansmann et al., 2019; He et al., 2023; He et al., 2025). The released product provides sub–hourly (15 min) and high vertical resolution (7.5 m) profiles from approximately 0.3 to 6 km, allowing the separation of dust from predominantly non–dust aerosols under a unified processing and quality control scheme. By bridging the gap between surface measurements, column observations, and satellite profiling products, this dataset is intended as a community resource that reduces the effort required to obtain harmonized long–term aerosol type profiles for central China. It can support satellite validation, chemical transport model evaluation, data assimilation, studies of aerosol transport and boundary layer coupling, and analyses of extreme aerosol events, while also providing a benchmark for future intercomparisons with other regional observing systems.

This study focused on the generation, evaluation, uncertainty characterization, and potential reuse of the dataset. We describe the observation site, instruments, and retrieval procedures; assess the product against independent column, profile, and near–surface references; and document the major uncertainty sources, application limits, and pointwise retrieval quality index (RQI). In doing so, we aim to provide a transparent and reusable observational product for future Earth system studies in eastern China and for broader community applications that require long–term vertically resolved aerosol mass information.

https://doi.org/10.5194/essd-2026-677
Preprint. Discussion started: 23 September 2026

2 Experimental setup

2.1 Description of the monitoring site

The polarization lidar used to generate this dataset was deployed at the SXNCO (32.434° N, 116.793° E). Situated near the northwestern boundary of the YRD, the SXNCO is strategically positioned along the primary atmospheric pollutant transport corridor connecting the BTH and YRD regions (Fig.1a). This specific geographical location makes it a critical observational node for capturing and quantifying regional aerosol transport processes between two of eastern China's most economically significant and heavily populated zones. Specifically, this location makes SXNCO an effective receptor site for observing the southward export of pollution from northern China, the northward influence of eastern China's emissions, and the passage of spring dust plumes across central China.

To illustrate the regional pollution background, Figure 1a shows the spatial distribution of the Moderate Resolution Imaging Spectroradiometer (MODIS) mean aerosol optical depth (AOD) over eastern China from 2018 to 2024. Both the BTH and YRD regions exhibited persistently high AOD loadings, with the SXNCO site located squarely within the transition zone bridging these heavily polluted areas. This macro–scale spatial pattern confirms the representativeness of the site for monitoring transboundary aerosol plumes. On a local scale, the lidar system was installed within an open, homogeneous agricultural field (primarily rice–wheat rotation croplands), as shown in the aerial photograph (Fig. 1b). Because of this rural setting, the site lacks substantial stationary anthropogenic emission sources (e.g., large–scale industrial facilities) and dense traffic networks. Consequently, the local primary emissions were minimal. This is a crucial attribute for the quality of the dataset: the spatiotemporal evolution of aerosol optical properties and mass concentrations observed here is predominantly driven by regional–scale atmospheric transport rather than abrupt local emission artifacts. This stable background environment effectively isolates regional transport signals, providing a natural advantage for the high–precision retrieval and long–term statistical analysis of dust and non–dust (anthropogenic) aerosol profiles. Therefore, the site preferentially records the regional–scale evolution of transported coarse and fine particles, which is highly advantageous for constructing a long–term dataset intended for community use in transport studies, satellite evaluation, and model benchmarking.

2.2 Description of the polarization lidar

The polarization lidar system was deployed at SXNCO in late 2017 and subsequently transitioned into continuous operation. To ensure long–term stability under complex field conditions, the instrument was housed within a dedicated, environmentally controlled cabin equipped with an uninterruptible power supply and a thermostatic dehumidification system. Operating continuously in a vertically pointing zenith mode through a transparent roof skylight, the lidar monitored tropospheric aerosols 24 h a day.

The hardware architecture comprises three primary modules: emission, receiving, and signal detection and acquisition systems. The emission system featured an Nd:YAG solid–state laser (model: Inlite II–20) that simultaneously emitted laser pulses at three wavelengths: 355, 532, and 1064 nm. The single–pulse energies were 40, 125, and 250 mJ. With a pulse repetition

https://doi.org/10.5194/essd-2026-677
Preprint. Discussion started: 23 September 2026

frequency of 20 Hz and a strictly controlled beam divergence angle of < 0.75 mrad, the system ensures sufficient detection energy and excellent beam directionality. On the receiver side, the backscattered signals were collected using a 254 mm Cassegrain telescope with a field of view of 0.3 mrad. The collected light passes through a sophisticated spectroscopic system before being routed to the specific detection channels. The 355 nm, 532 nm parallel polarization (532p), and 532 nm perpendicular polarization (532s) channels utilize high–sensitivity photomultiplier tubes (PMTs; Hamamatsu 9880U) for photoelectric conversion, whereas the 1064 nm channel employs an avalanche photodiode (APD; Licel APD 3.0). Signal acquisition relied on an analogue detection mode powered by an ADLINK PCI–9846H/512 high–speed data acquisition card, achieving a vertical spatial resolution of 7.5 m. Crucially, the polarization detection capability at 532 nm enables the precise retrieval of the particle depolarization ratio (PDR), which is an essential parameter for discriminating non–spherical dust particles from near–spherical non–dust (anthropogenic) aerosols (Freudenthaler et al., 2009; Wang et al., 2021a; Wang et al., 2021b).

A structured trial phase was conducted to optimize the temporal resolution of the data acquisition. During the initial operation from January to June 2018, the temporal resolution was iteratively tested at 5, 7, 10, and 15 min intervals to achieve an optimal balance between the signal–to–noise ratio (SNR) and detection frequency. Since June 2018, the system has maintained a stable long–term operation with a temporal resolution uniformly fixed at 5 min. Figure 1c shows the monthly effective data coverage of the lidar from January 2018 to December 2024, calculated after the stringent exclusion of profiles contaminated by precipitation, snow, low clouds, and instrument malfunctions. Over this continuous seven–year period, the system maintained a data capture rate with no significant long–term operational gaps. Furthermore, Figures 1d–h present the statistical distribution of effective detection profiles within the 0–6 km tropospheric altitude range, aggregated over the entire observation period across the four seasons. The substantial volume of effective samples accumulated at each altitude level not only corroborates the excellent vertical penetration capability of the lidar system but also underscores the rigorous spatiotemporal continuity of SXNCO observations. This high data coverage provides a solid statistical foundation for the subsequent precise retrieval of aerosol mass concentration profiles and the generation of a high–resolution, long–term aerosol dataset.

2.3 Auxiliary equipment

Several collocated or nearby observations were used to constrain the retrievals and evaluate the final product. The auxiliary datasets serve complementary roles: sun–photometer observations provide column optical and microphysical constraints for extinction–to–mass conversion, CALIPSO provides an independent reference for vertical optical structure, and in situ particle measurements test the realism of the lowest lidar–derived mass concentrations.

2.3.1 Sun photometer (CE–318) observations

An automatic tracking sun photometer (CE–318; Cimel Electronique, France) was used to characterize the optical and microphysical properties of the total column aerosols. The instrument was mounted on the roof of the same air–conditioned control cabin housing the polarization lidar (approximately 5 m above ground level), with a horizontal separation of 4 m. This

https://doi.org/10.5194/essd-2026-677
Preprint. Discussion started: 23 September 2026

physical colocation ensured high spatial consistency of the sampled air masses between the active and passive sensors. The effective observation period for CE–318 spanned from January 2016 to June 2023. Under clear sky conditions, it provides high–precision AOD measurements alongside columnar volume size distributions derived using the standard AERONET retrieval framework (Holben et al., 1998). Crucially, these long–term photometric records served a dual purpose in this dataset: they were utilized both for the independent validation of lidar–derived AOD and for deriving the localized, monthly coarse– and fine–mode extinction–to–mass conversion factors required by the POLIPHON mass conversion.

2.3.2 CALIPSO aerosol profile products

To independently evaluate the ground–based lidar retrieval accuracy for the aerosol extinction coefficient (EXT) and PDR profiles in the middle and lower troposphere, we used Level 2 aerosol profile products from the Cloud–Aerosol Lidar with Orthogonal Polarization (CALIOP) onboard the CALIPSO satellite (Kim et al., 2018a; Winker et al., 2009a). Because the CALIPSO mission formally concluded its scientific observations in August 2023, we extracted and matched spaceborne overpasses within a strictly defined spatial window around the SXNCO site from January 2018 to June 2023. CALIPSO provides EXT and PDR measurements at 532 nm, which directly correspond to the operating wavelength of the ground–based polarization lidar.

2.3.3 In situ particulate matter measurements

To verify the near–surface aerosol mass concentration retrievals, continuous in–situ measurements were obtained using a Grimm 180 optical aerosol spectrometer (Grimm Aerosol Technik, Germany). Operating on the principle of optical light scattering, this analyzer provides real–time, high–precision monitoring of particulate matter (PM_1, $PM_{2.5}$, and PM_{10}) mass concentrations in the ambient air. The Grimm 180 was also strictly colocated with the lidar system at a distance of approximately 4 m. To ensure representative and unobstructed sampling, ambient air was drawn through an all–weather 24–h sampling inlet tube that vertically penetrated the cabin roof. Operational from 2017 to the present, this instrument supplied hourly reference PM measurements that were synchronized with the lidar temporal resolution. These data were directly utilized for linear regression comparisons with the PM mass concentrations retrieved from the lowest usable lidar altitude bin (~300 m). This continuous comparison quantitatively assessed the consistency of the lidar optical–to–mass conversion algorithm within the PBL.

2.4 Retrieval of dust and non–dust mass concentrations

In this study, dust and non–dust aerosol mass concentrations were derived from the polarization lidar data using the POLIPHON method (Ansmann et al., 2019; Haarig et al., 2019; Wang et al., 2021a). The final dataset was generated through four main steps: (1) raw signal preprocessing and quality control, (2) retrieval of aerosol optical properties at 532 nm, (3) separation of dust and non–dust optical contributions, and (4) conversion of EXT profiles to dust, non–dust, and total mass concentration profiles using sun–photometer constraints.

https://doi.org/10.5194/essd-2026-677
Preprint. Discussion started: 23 September 2026

2.4.1 Lidar data preprocessing and quality control

Raw measurements were recorded at a temporal resolution of 5 min and a vertical resolution of 7.5 m (after June 2018). Before optical retrieval, every three consecutive raw profiles were averaged to obtain a working resolution of 15 min, which improved the SNR while preserving sub–hourly variability. The data time resolution before June 2018 was constantly adjusted within the range of 5–15 min. We also resampled the data for 15 min. The uncertainty in the retrieved aerosol optical properties was quantified using standard error propagation theory. These estimates were mainly determined by the SNR of the averaged polarization lidar backscatter profiles, following the approach of Heese et al. (Heese et al., 2010). Measurements with an SNR below 1 were discarded, and observations affected by precipitation (rain or snow) or by the presence of low–level clouds were excluded from further analysis. Additional periods affected by obvious instrumental malfunctions were removed from the final dataset.

2.4.2 Retrieval of aerosol optical parameters at 532 nm

After raw backscatter signal quality control, the aerosol backscatter coefficient at 532 nm was retrieved using the Fernald method (Fernald, 1984), which solves the single wavelength elastic lidar equation. Inversion requires an assumed lidar ratio (LR), which is defined as the ratio of aerosol extinction to backscatter coefficient. Based on previous observations and literature values representative of the region (Wang et al., 2021b; Wang et al., 2023b), a lidar ratio of 53 sr was used in the initial inversion, together with a clean reference layer aloft, where the aerosol loading was assumed to be negligible.

Because the system includes a polarized 532 nm channel, the volume depolarization ratio (VDR) was calculated as the ratio of the range–corrected, background–subtracted perpendicular and parallel return signals (Freudenthaler et al., 2009):

$$\delta_v(z) = \eta \frac{P^{\perp}(z)}{P^{\parallel}(z)} \tag{1}$$

where $P^{\perp}$ and $P^{\parallel}$ are the perpendicular and parallel return signals, respectively, at height z after background subtraction and range correction, and η is the system gain ratio. To accurately characterize the non–spherical properties of aerosol particles, the contribution of air molecules (Rayleigh scattering) must be removed from the VDR to obtain the PDR (δ_p):

$$\delta_p = \frac{\beta_m(\delta_v - \delta_m) + \beta_p\delta_v(1+\delta_m)}{\beta_m(\delta_m - \delta_v) + \beta_p(1+\delta_m)} \tag{2}$$

where δ_m is the molecular depolarization ratio, and β_p and β_m are the particle and molecular backscatter coefficients, respectively. The resulting product provides 15 min and 7.5 m profiles of EXT and PDR for subsequent aerosol typing.

https://doi.org/10.5194/essd-2026-677
Preprint. Discussion started: 23 September 2026

2.4.3 Separation of dust and non–dust aerosol components using POLIPHON

The POLIPHON method utilizes significant differences in particle shape (non–spherical vs. near–spherical) between dust and non–dust aerosols to achieve component separation (Sugimoto et al., 2003; Shimizu, 2004; Tesche et al., 2009; Ansmann et al., 2012; Ansmann et al., 2019; He et al., 2023; He et al., 2025). This method assumes that dust and non–dust aerosols are mixed externally. In the context of this dataset, the term "non–dust" mainly refers to fine–mode anthropogenic aerosols under local and regional conditions, although other weakly depolarizing spherical particles may also contribute.

In the first step, the dust backscatter coefficient was estimated from the retrieved aerosol backscatter coefficient and PDR using the characteristic depolarization values for dust and non–dust aerosols:

$$\beta_d = \beta_p \frac{(\delta_p - \delta_{nd})(1+\delta_d)}{(\delta_d - \delta_{nd})(1+\delta_p)} \quad (3)$$

where δ_d and δ_{nd} are the typical PDR values for dust and non–dust aerosols, respectively. To avoid non–physical solutions, β_d is set to 0 when $\delta_p <= \delta_{nd}$ and to β_p when $\delta_p >= \delta_d$. The non–dust backscatter coefficient (β_{nd}) is then calculated as:

$$\beta_{nd} = \beta_p - \beta_d \quad (4)$$

The dust and non–dust EXT profiles were subsequently obtained from their respective backscatter coefficients and component–specific lidar ratios:

$$\alpha_d = \beta_d \times S_d \quad (5)$$

$$\alpha_{nd} = \beta_{nd} \times S_{nd} \quad (6)$$

where S_d and S_{nd} are the lidar ratios of dust and non–dust aerosols, respectively. Based on our previous polarization Raman lidar observations (Wang et al., 2021b; Wang et al., 2023b), we used S_{nd} = 54.4 ± 11.1 sr, S_d = 43.0 ± 8.2 sr, δ_{nd} = 0.043 ± 0.021, and δ_d = 0.287 ± 0.043.

In the second step, the EXT component was converted to mass concentration following Ansmann et al. (Ansmann et al., 2012) and Mamouri et al. (Mamouri and Ansmann, 2014). The dust mass concentration (M_d), non–dust mass concentration (M_{nd}), and total aerosol mass concentration (M_p) were calculated from the EXT profiles using aerosol density and extinction–to–mass conversion factors:

$$M_d = \rho_d \times \overline{(v_c / \tau_c)} \times \alpha_d \quad (7)$$

$$M_{nd} = \rho_{nd} \times \overline{(v_f / \tau_f)} \times \alpha_{nd} \quad (8)$$

$$M_p = M_d + M_{nd} \quad (9)$$

https://doi.org/10.5194/essd-2026-677
Preprint. Discussion started: 23 September 2026

where ρ_d and ρ_{nd} are the mass densities of the dust and non–dust aerosols, respectively. We used ρ_d = 2.6±0.6 g/cm^3 and ρ_{nd} = 1.5±0.3 g/cm^3. The conversion factors linking EXT to particle mass concentration are constrained by CE–318 observations of coarse– and fine–mode volume concentrations (v) and AOD (τ), where the subscripts c and f represent coarse– and fine–mode particles, respectively. To make the long–term product stable and reproducible, monthly mean conversion factors were calculated from all valid CE–318 observations from January 2018 to June 2023 (Fig. 2a). This choice smoothed short–term extremes but provided representative climatological conversion parameters for the site. Table 1 summarizes the input parameters used to generate the datasets.

Table 1. Mean values and one standard deviation of input parameters in dust and non–dust aerosol mass concentration retrieval

Parameter	Value	Reference
Non–dust lidar ratio S_{nd}	54.4±11.1 sr	(Wang et al., 2023)
Asian dust lidar ratio S_d	43.0±8.2 sr	(Wang et al., 2023)
Non–dust particle depolarization ratio δ_{nd}	0.043±0.021	(Wang et al., 2023)
Asian dust particle depolarization ratio δ_d	0.287±0.043	(Wang et al., 2023)
Non–dust mass density ρ_{nd}	1.5±0.3 g/cm^3	(Ansmann et al., 2012)
Asian dust mass density ρ_d	2.6±0.6 g/cm^3	(Ansmann et al., 2012)
Non–dust conversion factor $\overline{v_f / \tau_f}$	Monthly average	Sun–photometer observations
Asian dust conversion factor $\overline{v_c / \tau_c}$	Monthly average	Sun–photometer observations

2.4.4 Uncertainty estimation and retrieval limitations

The uncertainties in the mass concentrations estimated using the POLIPHON method primarily originate from three distinct sources: localized input parameter uncertainties, instrumental lidar retrieval errors, and inherent methodological assumptions.

(1) Input parameter uncertainties: These encompass the uncertainties associated with the assumed typical PDRs (δ_d and δ_{nd}), component–specific lidar ratios (S_d and S_{nd}), mass densities (ρ_d and ρ_{nd}), and extinction–to–mass conversion factors derived from the sun photometer ($\overline{v_c / \tau_c}$ and $\overline{v_f / \tau_f}$). The uncertainties for all these predetermined parameters were explicitly considered in the calculations using their respective standard deviations.

(2) Lidar retrieval errors: The retrieval accuracy of the total particle backscatter coefficient (β_p) and PDR (δ_p) is affected by random signal noise, uncertainties in the assumed lidar ratio during the Fernald inversion, and incomplete overlap corrections in the near–surface layer. Applying the standard law of error propagation based on the SNR, we calculated the relative errors

https://doi.org/10.5194/essd-2026-677
Preprint. Discussion started: 23 September 2026

for β_p and δ_p at each discrete altitude bin. As a crucial quality control step, data points exhibiting excessive relative errors (e.g., β_p > 20% and δ_p > 15%) were systematically excluded from the final mass concentration retrievals.

(3) Inherent method assumptions: A fundamental premise of the POLIPHON framework is that dust and anthropogenic aerosols exist strictly as external mixtures. Internal mixing frequently occurs in complex real–world atmospheric scenarios (e.g., anthropogenic pollutants coating mineral dust particles) (Sugimoto et al., 2015; Pan et al., 2017). This physical alteration suppresses the apparent PDR of dust, thereby introducing separation error. Furthermore, the reliability of the algorithm diminishes under high relative humidity (RH) conditions. In high–humidity environments, substantial hygroscopic growth alters particle morphology, rendering non–spherical particles more spherical while concurrently modifying particle density and the complex refractive index. This dynamic process fundamentally violates the assumption of stable, component–specific particle shapes (Ansmann et al., 2012). Therefore, users should exercise additional caution under humid or strongly mixed conditions, especially when interpreting single–profile retrievals.

Synthesizing all the aforementioned error sources via standard Gaussian error propagation algorithms (as detailed in Table 2), the total relative uncertainty of the derived dust and anthropogenic aerosol mass concentration profiles in this study was conservatively estimated to range between 15% and 70%. This calculated uncertainty budget aligns well with the overall uncertainty of approximately 40% reported by Ansmann et al. (Ansmann et al., 2011) and the relative differences of less than 30% observed by Bravo–Aranda et al. (Bravo–Aranda et al., 2015) during direct quantitative comparisons with airborne in situ measurements.

Table 2. Estimated uncertainties in the polarization lidar–derived parameters.

Parameter	Uncertainty (%)	Reference
Volume depolarization ratio δ_v	<5	(Wang et al., 2023)
Particle depolarization ratio δ_p	2–15	(Wang et al., 2023)
Particle backscatter coefficient β_p	2–20	(Wang et al., 2023)
Particle extinction coefficient α_p	<25	(Wang et al., 2023)
Dust backscatter coefficient β_d	10–30	Mamouri and Ansmann (2016, 2017)
Dust extinction coefficient α_d	15–25	Mamouri and Ansmann (2016, 2017)
Non–dust backscatter coefficient β_{nd}	10–30	Mamouri and Ansmann (2016, 2017)
Non–dust extinction coefficient α_{nd}	20–40	Mamouri and Ansmann (2016, 2017)

2.5 Pointwise retrieval quality index (RQI)

Section 2.4.4 characterizes the principal sources of uncertainty and provides an overall uncertainty envelope for the retrieved mass concentration profiles. We do not, however, assign a formal pointwise uncertainty estimate to each 15 min × 7.5 m grid

https://doi.org/10.5194/essd-2026-677
Preprint. Discussion started: 23 September 2026

cell. A defensible pointwise uncertainty estimate would require time– and height–resolved probability distributions and covariance information for all prescribed retrieval parameters, including EXT, PDR, particle densities, and extinction–to–mass conversion factors, together with structural uncertainties associated with the external mixture assumption, aerosol mixing state, and hygroscopic growth. Several of these contributions are systematic, non–Gaussian, or not independently observable for individual retrievals. Propagating only the available random errors into a pointwise standard deviation would therefore imply a degree of probabilistic precision that is not supported by the observations.

To provide grid cell specific quality information without conflating it with a formal uncertainty estimate, the released dataset includes a pointwise RQI. The RQI is a dimensionless, non–probabilistic indicator of retrieval conditions ranging from 0 to 100 and should not be interpreted as an error percentage. A single RQI is assigned to the dust, non–dust, and total aerosol mass concentration products at each time–height grid cell, providing a consistent quality descriptor for the three mutually related mass products.

The RQI is based on three quantities that characterize conditions relevant to the robustness of the POLIPHON retrieval: relative humidity (RH; obtained from the fifth–generation ECMWF atmospheric reanalysis, ERA5), the relative error of VDR, and the relative error of EXT. The relative importance assigned to RH is additionally allowed to vary with the retrieved dust/non–dust composition. This modification reflects the fact that humidity–induced changes in particle morphology and optical properties are particularly relevant to component partitioning when dust and non–dust aerosols contribute simultaneously to the observed backscatter. By contrast, when the aerosol backscatter is strongly dominated by either component, the dust/non–dust partition is less ambiguous and the relative influence of RH in the common quality index is reduced.

All fields are mapped to the final mass concentration time–height grid before the RQI is calculated. For each valid grid cell, RH and the two lidar–derived relative errors are first transformed into dimensionless subscores between 0 and 1, with more favorable retrieval conditions approaching 1 and progressively less favorable conditions approaching 0.

For RH, no penalty is applied when RH ≤ 70 %. Between 70 % and 98 %, the RH subscore decreases smoothly according to a cubic smooth–step function and is set to zero at RH ≥ 98 %:

$$q_{RH} = \begin{cases} 1, RH \le 70\% \\ 1 - 3u^2 + 2u^3, 70 < RH \le 98\% \\ 0, RH > 98\% \end{cases} \quad (10)$$

$$u = (RH - 70)/28 \quad (11)$$

This formulation represents the increasing susceptibility of the retrieval to hygroscopic growth and humidity dependent changes in particle morphology while avoiding an artificial discontinuity across the transition range. The RH ≥ 98 % condition is retained as a conservative common low–quality limit irrespective of aerosol composition.

For the lidar–derived error terms, let ε_{VDR} and ε_{EXT} denote the relative errors of VDR and EXT, respectively, expressed as percentages. Their subscores are defined using monotonic hill–type functions:

$$q_{VDR} = [1 + (\varepsilon_{VDR} / 3)^{1.5}]^{-1} \quad (12)$$

https://doi.org/10.5194/essd-2026-677
Preprint. Discussion started: 23 September 2026

$$q_{EXT} = [1 + (\varepsilon_{EXT} / 14)^{1.5}]^{-1} \quad (13)$$

Thus, larger VDR or EXT errors always yield lower subscores. The characteristic scales of 3 % for VDR and 14 % for EXT, together with the RH transition limits, were parameterized from the 2018–2024 error distributions used during method development. They are therefore dataset specific quality parameters rather than universal uncertainty thresholds for polarization lidar observations.

To account for the dependence of the humidity penalty on aerosol composition, a dust backscatter fraction is defined from the POLIPHON separated dust and non–dust backscatter coefficients:

$$f_d = \beta_d / \beta_p \ , \ 0 \le f_d \le 1 \quad (14)$$

A continuous dust/non–dust coexistence factor is then defined as

$$m = 4 f_d (1 - f_d) \ , \ 0 \le m \le 1 \quad (15)$$

The factor m approaches zero when the retrieved backscatter is dominated by either non–dust aerosol or dust aerosol and reaches unity when dust and non–dust contribute equally to the particle backscatter. Here, m is used only as a continuous indicator of optical component coexistence and should not be interpreted as a direct measure of particle–scale internal mixing. The weights assigned to the three quality subscores are then defined as

$$\omega_{RH} = 0.15 + 0.2m, \omega_{VDR} = 0.40, \omega_{EXT} = 0.45 - 0.2m \quad (16)$$

such that

$$\omega_{RH} + \omega_{VDR} + \omega_{EXT} = 1 \quad (17)$$

The pointwise RQI is finally calculated as a weighted geometric mean:

$$RQI = 100 \cdot q_{RH}^{\omega_{RH}} \cdot q_{VDR}^{0.40} \cdot q_{EXT}^{\omega_{EXT}} \quad (18)$$

This formulation retains the non–compensatory behavior of the geometric mean, such that strongly degraded quality in one component cannot be fully compensated by favorable values of the other components. The VDR error term retains a constant weight because polarization information directly constrains dust/non–dust separation. In contrast, the RH contribution varies with the retrieved aerosol composition. Under strongly dust or non–dust dominated conditions, the RH weight decreases to 0.15 and the EXT weight increases to 0.45, reflecting the greater relative importance of the overall optical retrieval when aerosol type attribution is comparatively unambiguous. Under the strongest dust/non–dust coexistence, the weights become 0.35, 0.40, and 0.25 for RH, VDR, and EXT, respectively. The latter weighting corresponds to the most compositionally ambiguous retrieval conditions. The composition dependent weighting therefore modifies the strength of the RH penalty primarily within the 70–98 % RH transition range, while the conservative zero–quality condition at RH ≥ 98 % is retained for all aerosol regimes.

The RQI is constrained to the range 0–100 and rounded to the nearest integer. The same RQI is assigned to the dust, non–dust, and total aerosol mass concentration products at a given time–height grid cell because all three are derived from the same optical retrieval and atmospheric state. This common–index design is intentional: the RQI characterizes the overall reliability

https://doi.org/10.5194/essd-2026-677
Preprint. Discussion started: 23 September 2026

of the retrieval conditions rather than providing separate probabilistic uncertainties for individual aerosol components. It therefore allows users to apply a consistent quality mask when jointly analyzing dust, non–dust, and total aerosol mass concentrations.

For practical reuse, RQI values of 75–100 indicate high–quality retrieval conditions and are suitable for pointwise quantitative analyses, event studies, and collocated satellite or model comparisons. Values of 50–75 are generally appropriate for routine quantitative and statistical analyses. Values of 25–50 represent moderately favorable retrieval conditions and are preferably interpreted using temporal or vertical averages or in conjunction with independent observations. Values of 0–25 are retained for transparency but are recommended mainly for qualitative screening or supporting context and are not recommended for quantitative interpretation of absolute mass concentration. These categories provide usage guidance only and do not replace the uncertainty range discussed in Sect. 2.4.4. Because RQI–based filtering may preferentially remove humid mixed–aerosol conditions and/or observations with large lidar–derived errors, users applying an RQI threshold should report the retained data fraction and consider potential changes in seasonal, vertical, and aerosol composition sampling.

3 Multi–source evaluation and validation of the dataset

3.1 Comparison of AOD with CE–318

To systematically evaluate the reliability of the lidar–derived AOD, we cross–validated it against measurements from a colocated CE–318 sun photometer. Given the inherent differences in observation geometry, temporal sampling, and operating wavelengths, rigorous data colocation protocols were applied prior to comparison.

A temporal matching window of ±1 h was adopted. First, hourly averages of the CE–318 AOD were calculated; subsequently, the corresponding lidar AOD was derived by averaging all valid lidar profiles within a ±1–h window centered on each CE–318 observation time. The lidar AOD was calculated via the vertical integration of the retrieved EXT profiles from the surface to 6 km. To address the near–surface instrumental blind zone caused by incomplete geometric optical overlap, EXT values below 0.3 km were extrapolated using the mean EXT from the lowermost usable layer (0.3–0.35 km). This standard operational procedure effectively mitigates systematic AOD underestimation associated with missing near–surface data. Furthermore, to ensure wavelength consistency, the standard CE–318 AOD products at 500 nm were converted to 532 nm (the lidar operating wavelength) using the Ångström exponent power–law relationship. After collocation, the paired residuals were screened using a 3–sigma criterion to suppress rare outliers caused by mismatches or residual contamination.

Comparative statistical analyses were performed at both hourly and monthly temporal scales (Fig. 2b–d). The hourly comparison yielded moderate agreement, with a linear regression slope of 0.60 and a Pearson correlation coefficient (R) of 0.62 ($P < 0.01$). Notably, this agreement improved substantially at the monthly mean scale, achieving a slope of 0.78 and an R of 0.83 ($P < 0.01$). This statistical behavior, in which higher frequency discrepancies are effectively smoothed by monthly averaging, is characteristic of active versus passive remote sensing comparisons. This demonstrates that hourly non–

https://doi.org/10.5194/essd-2026-677
Preprint. Discussion started: 23 September 2026

stationarity in atmospheric conditions and random retrieval uncertainties are largely mitigated over longer integration periods (Mamouri and Ansmann, 2014).

These observed discrepancies at higher temporal resolutions do not imply instrumental defects but rather stem directly from inherent methodological constraints and distinct observation geometries: 1) Spatiotemporal representativeness errors. While the ±1–h matching window maximizes the sample size and statistical robustness, rapid diurnal variations in PBL dynamics, relative humidity, and advection can significantly alter the column AOD within a 1–h timeframe, amplifying the scatter in the hourly data. Monthly averaging successfully suppressed this high–frequency meteorological noise (Fig. 2d). 2) Column integration limits: The lidar AOD was vertically integrated up to a fixed altitude of 6 km, whereas the CE–318 inherently measured the entire atmospheric column. During events involving high–altitude aerosol transport (e.g., elevated dust plumes above 6 km), the lidar mathematically underestimates the true total column AOD, affecting the regression slope during high AOD events. 3) Blind zone extrapolation biases: Filling the 0–0.3 km blind zone with EXT values from the 0.3–0.35 km layer is operationally consistent for long–term datasets. However, under conditions of severe near–surface pollution or strong vertical EXT gradients, this assumption may introduce systemic biases, thereby influencing the regression intercept and scatter in the low AOD range. 4) Wavelength conversion uncertainty: The Ångström conversion of CE–318 AOD from 500 to 532 nm assumes a strictly power–law relationship. This assumption can slightly break down when coarse–mode particles dominate (e.g., intense dust events) or under severe hygroscopic growth conditions, introducing minor seasonal offsets visible in the intra–annual distribution comparison (Fig. 2c).

In summary, despite the expected variance at the individual high–frequency sample level, the monthly scale results demonstrate robust climatological consistency between the active lidar and passive CE–318 retrievals. This level of statistical agreement is comparable to that of established observational studies (Balmes et al., 2021; Bi et al., 2024; Omar et al., 2013), supporting the applicability of this dataset for long–term aerosol trend analysis and regional climatological assessments.

3.2 Comparison of vertical aerosol structures with CALIPSO

To independently evaluate the credibility of the retrieved vertical structures of the EXT and PDR, we cross–referenced the ground–based polarization lidar dataset with the CALIPSO Version 4 Level 2 aerosol profile product (CAL_LID_L2_05kmAPro–Standard–V4–51). CALIPSO Level 2 profiles were reported on a 5 km horizontal grid with a vertical resolution of 60 m. Because this snapshot–style spatial grid fundamentally differs from the continuous Eulerian time–height representation of ground–based lidar and boundary layer aerosols exhibit strong spatial heterogeneity, rigorous spatiotemporal collocation criteria were enforced. We selected only CALIPSO overpasses within a radial distance of < 90 km from the SXNCO site and temporal window of < 1 h. Between January 2018 and June 2023, 93 valid overpasses were successfully identified.

To reconcile the disparate observation modes and resolutions, we applied symmetric temporal and spatial averaging to construct directly comparable "event mean" profiles. For the ground–based lidar, all valid profiles within the ±1 h window were averaged. Concurrently, for CALIPSO, all 5 km grid profiles falling within a 90 km radius were averaged. Because

https://doi.org/10.5194/essd-2026-677
Preprint. Discussion started: 23 September 2026

spaceborne PDR is inherently noisier than ground–based retrievals and occasionally exhibits non–physical extreme values, we applied an upper quality control threshold of 0.6 to the satellite PDR (Wang et al., 2021b). Ground–based PDR values typically remain well below 0.45, confirming that this threshold effectively suppresses extreme noise without eliminating genuine physical signals. Finally, the event mean profiles were aggregated into annual mean profiles to compute linear regressions and correlation coefficients (Fig. 3), alongside an overall scatter analysis compiling all "annual mean altitude points" from 2018 to 2023 (Fig. 4). In addition, a residual–based 3–sigma filter was applied to the paired altitude bins before regression.

Statistical comparisons revealed a distinct divergence in consistency between aerosol intensity and aerosol shape parameters. The EXT demonstrated stable agreement across all years, with annual correlation coefficients (R) predominantly ranging from 0.88 to 0.93 ($P < 0.01$). The overall scatter analysis for EXT similarly yielded a strong relationship (slope = 0.73, R = 0.82, $P < 0.01$). In contrast, the annual mean PDR profiles exhibited notably lower and more fluctuating correlations. During 2018–2020, PDR correlations remained weak (R = 0.11 to 0.26), although significant improvements were observed in subsequent years (e.g., R = 0.61 in 2021; R = 0.57 in 2023). Overall, the multi–year PDR scatter yielded a slope of 0.31 and R of 0.46 ($P < 0.01$).

This pattern, where the agreement for extinction intensity vastly outperforms the characterization of non–sphericity, is a widely recognized characteristic of ground–to–satellite lidar comparisons (Chang et al., 2024; Floutsi et al., 2023; Midzak et al., 2025). Agreement was substantially weaker for PDR than for EXT, reflecting the combined effects of spatial representativeness differences, retrieval noise, and possible platform–dependent biases. Consequently, the CALIPSO comparison provides only limited quantitative validation of PDR: 1) spatial representativeness mismatch. Ground–based lidar provides a localized point measurement in a pristine agricultural background, effectively isolating regional transport. In contrast, CALIPSO's 90 km spatial average inevitably encompasses adjacent urban and suburban surfaces. Fine PM and complex mixed–state aerosols from urban emissions dramatically alter the PDR. Therefore, aerosols are highly unlikely to remain perfectly homogeneous across a 90 km radius, particularly for the highly sensitive PDR parameters (Pappalardo et al., 2010; Tesche et al., 2013). 2) Spaceborne SNR degradation. The PDR calculation is critically dependent on the backscatter signal from the perpendicular channel. For CALIPSO, the laser pulse must traverse the entire atmosphere twice, resulting in a fundamentally weak perpendicular return signal that is highly susceptible to solar background radiation noise, particularly during daytime overpasses. This extremely low SNR directly induces massive fluctuations and substantial uncertainties in the CALIPSO PDR profiles, severely suppressing the regression slope (Hunt et al., 2009; Winker et al., 2009b). 3) Multiple scattering artifacts. CALIPSO's downward–looking geometry produces a relatively large surface footprint (~100 m). This broad footprint exponentially increases the probability of multiple scattering events occurring within dense aerosol layers or cloud boundaries. Physically, multiple scattering induces additional depolarization of the beam, artificially inflating the PDR values retrieved by the CALIPSO algorithm (Wandinger et al., 2010; Hu et al., 2006). In contrast, ground–based lidar features a remarkably narrow field of view (0.3 mrad), rendering multiple scattering effects mathematically negligible. 4) Eulerian versus snapshot advection. The ground–based lidar continuously observes air mass evolution at fixed coordinates (Eulerian perspective), whereas CALIPSO captures instantaneous snapshots along a polar orbit transect. Given the high spatial inhomogeneity of coarse–mode

https://doi.org/10.5194/essd-2026-677
Preprint. Discussion started: 23 September 2026

dust, even within a strict 1 h/90 km window, the two platforms frequently sample physically distinct air masses (Tesche et al., 2013).

In conclusion, the overall scatter analysis demonstrated robust agreement between the ground–based lidar and CALIPSO regarding the vertical structure of EXT, strongly supporting the climatological validity of this dataset. Although the statistical correlation for PDR was systematically lower, this divergence was primarily driven by the objective geometric, spatial, and instrumental limitations of the spaceborne platform. The ground–based polarization lidar provides statistically consistent vertical aerosol structure information, thereby supplying dependable input parameters for the subsequent separation of dust and non–dust mass concentrations.

3.3 Comparison with surface PM mass concentrations

To evaluate the reliability of the aerosol mass concentrations retrieved using the POLIPHON method, we systematically cross–validated the mass concentrations from the lowest usable lidar altitude bin (averaged over 0.30–0.32 km) against strictly synchronized in situ surface PM measurements. Because the surface instrument samples were collected at ground level, whereas the lidar range gate was several hundred meters above ground, these comparisons were intended to test temporal co–variability rather than exact one–to–one equivalence. Prior to the comparison, both datasets were aggregated at an hourly resolution for analysis. The evaluations were structurally mapped as follows: lidar–derived non–dust aerosols against surface PM_1 and $PM_{2.5}$; lidar–derived dust aerosols against the PM_{10}-PM_1 difference (serving as a dust proxy); and total lidar aerosol mass directly against surface PM_{10}. After collocation, a residual–based 3–sigma filter was used to remove rare outliers.

The analysis revealed statistically significant positive correlations across all four comparison groups, demonstrating that the POLIPHON algorithm effectively captured the long–term temporal evolution of both dust and non–dust mass concentrations (Fig. 5). The correlation coefficients (R) were 0.60, 0.55, 0.60, and 0.62 for non–dust–PM_1, non–dust–$PM_{2.5}$, total–PM_{10}, and dust–PM_{10}-PM_1, respectively (all passing the significance test at $P < 0.01$). These positive correlations confirm the statistical validity of the dataset for tracking near–surface mass variations. However, quantitative evaluation revealed that the linear regression slopes ranged strictly between 0.33 and 0.38, with scatter points systematically biased below the 1:1 reference line. This indicates that the lidar–retrieved mass concentrations at 0.3 km exceeded the surface–measured PM concentrations. This systematic offset and its associated scatter are not indicative of instrumental defects but rather emerge from a complex combination of vertical observation geometries, physical environmental differences, and inherent retrieval equivalency assumptions.

First, the lidar "near surface" layer (0.30–0.32 km) does not strictly represent the surface sampling height. Strong vertical gradients within the PBL, coupled with rapid changes in local near–surface emissions and turbulent mixing intensity, mean that mass concentrations at 0.3 km and at the surface are not always synchronized, thereby introducing representativeness errors. Numerous lidar studies have indicated that aerosols often accumulate near the top of the mixed layer or within nocturnal residual layers, resulting in higher aerosol loads several hundred meters above the surface than those observed at the ground level (Shi et al., 2020; Fan et al., 2019; Quan et al., 2025). Consequently, a slope less than unity in the direct regression of

https://doi.org/10.5194/essd-2026-677
Preprint. Discussion started: 23 September 2026

lidar–derived mass concentrations from the 0.30–0.32 km layer against surface PM is an anticipated representativeness discrepancy. Previous applications of the POLIPHON method in the BTH conducted similar comparisons, revealing that PM in that region tends to concentrate near the top of the PBL rather than at the surface, yielding comparable results (Wang et al., 2020; Wang et al., 2023b).

Second, the POLIPHON mass retrieval structurally propagates uncertainties through localized input parameters. For instance, utilizing extinction–to–mass conversion factors derived from column–integrated AERONET observations and applying them uniformly to specific altitude bins introduces inherent parameter uncertainties. This static column–averaged approach cannot perfectly resolve transient vertical variations in the localized aerosol microphysics. As noted in our uncertainty assessment (Section 2.4.4), the combined retrieval uncertainties in the lower troposphere ranged from 15% to 70%, directly contributing to the observed scatter.

Finally, there is no strict physical equivalence between optically separated "aerosol types" and aerodynamically separated "size cuts." Non–dust aerosols do not map directly to PM_1 or $PM_{2.5}$. Furthermore, using PM_{10}-PM_1 as an aerodynamic proxy for coarse dust inevitably includes contributions from anthropogenic coarse–mode particles and propagates combined instrumental errors from both the PM_{10} and PM_1 channels. This cross–contamination inherently dampens the regression slope and amplifies the scatter at higher concentrations.

Taken together, the three independent comparisons demonstrate that the released product is basically internally consistent across the column, profile, and near–surface perspectives. This multilevel validation substantially strengthens the dataset for community reuse.

4 Data overview: spatiotemporal evolution and vertical interannual changes of aerosols

4.1 Dust aerosol

Figure 6 summarizes the main characteristics of the dust aerosol mass concentration profiles in the released dataset for the SXNCO site from 2018 to 2024. The annual mean profiles (Fig. 6a) show that the dust aerosol mass is concentrated mainly in the lower and middle troposphere, with most of the signal located below approximately 2 km and with decreasing concentrations toward higher altitudes. Considerable interannual variability was evident near the lowest usable retrieval level (~0.3 km), where the annual mean dust mass concentration was lowest in 2018, reached a maximum in 2019, and then decreased during the later part of the record. These variations indicate that the dataset captures substantial year–to–year differences in dust mass concentrations over central China.

The time–height distribution (Fig. 6b) further shows that dust occurrence is seasonal. Dust mass concentrations are generally enhanced in spring and, to a lesser extent, in winter, whereas summer values remain comparatively low. During individual dust episodes, elevated dust layers can extend to several kilometers above the ground, reaching 4–5 km in some cases. This behavior highlights one of the main advantages of the dataset: it resolves not only changes in dust loading but also the vertical placement of transported layers, which cannot be inferred from surface measurements or column–integrated observations alone.

https://doi.org/10.5194/essd-2026-677
Preprint. Discussion started: 23 September 2026

The multi–year profile record also revealed that long–term changes in dust mass concentration were vertically heterogeneous (Fig. 6c–g). Over the full period, dust concentrations generally decreased in much of the 0.5–3.0 km layer, whereas changes near the lowest retrieval level were weaker and showed a stronger seasonal dependence. In summer and autumn, the profiles indicated broadly coherent decreases throughout most of the observed column. In contrast, spring and winter showed more complex behavior, including cases in which near–surface changes differed from those aloft. These results demonstrate that the long–term evolution of dust over the site cannot be represented adequately by a single altitude or by column averages alone, and illustrate the value of vertically resolved observations for describing seasonal and interannual variability.

4.2 Non–dust aerosol

Figure 7 summarizes the main characteristics of the non–dust aerosol mass concentration profiles in the dataset over the SXNCO site during 2018–2024. The non–dust aerosol mass was concentrated primarily within the PBL, with most of the signal located below approximately 2 km, and concentrations decreased rapidly with height. Near the lowest usable retrieval level (~0.3 km), non–dust mass concentrations remained comparatively high throughout the year, and clear interannual variability was visible in the lower troposphere, including relatively enhanced values in the earlier part of the record and lower values during the later years.

The monthly mean time–height distribution (Fig. 7b) showed a distinct seasonal cycle. The non–dust aerosol mass concentration was generally highest in winter, when enhanced loading was concentrated in the near–surface and lower boundary layer atmosphere, whereas summer values were substantially lower throughout most of the observed columns. This seasonal contrast demonstrates that the dataset resolves not only changes in the aerosol amount but also changes in the vertical confinement of non–dust aerosols. Such information is not available from surface measurements alone and is important for describing periods of shallow accumulation that are distinct from vertically mixed conditions.

The multi–year profile record further indicates that long–term changes in the non–dust aerosol mass concentration are vertically heterogeneous (Fig. 7c–g). In much of the profile above approximately 0.5 km, the dataset shows a general decrease over the observation period, whereas changes near the lowest retrieval level are smaller and more season–dependent. In several seasons, the sign and magnitude of the changes differed between the near–surface layer and the overlying atmosphere. These results indicate that long–term variability in non–dust aerosol over the site cannot be represented adequately by a single–level measurement or by column–integrated quantities alone, and they highlight the value of vertically resolved observations for documenting differences between boundary layer and elevated aerosol behavior.

4.3 Total aerosol

Figure 8 summarizes the main characteristics of the total aerosol mass concentration profiles in the dataset over the SXNCO site from 2018 to 2024. Because the total aerosol mass concentration is defined here as the sum of the retrieved dust and non–dust components, the profiles reflect the combined vertical behavior of both aerosol types under a unified processing framework. The annual mean profiles (Fig. 8a) show that most of the total aerosol mass is concentrated below approximately

https://doi.org/10.5194/essd-2026-677
Preprint. Discussion started: 23 September 2026

2 km, indicating that the lower troposphere contains the largest fraction of the aerosol load over the site. At the lowest usable retrieval level (~0.3 km), the total aerosol mass concentrations were relatively high during 2019–2020 and lower during the later part of the record, showing clear interannual variability in the near–surface atmosphere.

The time–height distribution (Fig. 8b) further shows that the vertical structure of the total aerosol mass concentration was strongly season–dependent. In winter, the enhanced total aerosol was mainly confined to the shallow near–surface and lower boundary layers. In spring, by contrast, high aerosol loadings frequently extended to higher altitudes, and in some cases, reached 3–5 km. These differences indicate that the dataset resolves seasonal changes not only in the aerosol amount but also in the vertical extent of the aerosol column. This is one of the main advantages of this product.

The profiles (Fig. 8c–g) show that interannual changes in the total aerosol mass concentration were vertically heterogeneous. Over much of the profile, especially in the 1.0–2.0 km layer, the total aerosol mass concentration generally decreased during the observation period. In spring, summer, and autumn, these decreases were evident throughout most of the observed columns. In winter, however, the profile structure was more complex, with decreases aloft occurring together with weaker changes or locally increasing tendencies near the surface. These vertically contrasting changes demonstrate that long–term aerosol evolution over the site cannot be adequately described by a single altitude or by column means alone. Therefore, the profile record provides added value for documenting how aerosol variability differs between the near–surface atmosphere and the overlying lower troposphere.

Taken together, the dust, non–dust, and total aerosol profiles presented in this section should be interpreted primarily as observational characteristics of the released data product rather than as a complete attribution of the processes controlling aerosol variability over central China. The dataset is particularly useful for identifying elevated dust transport layers, shallow non–dust accumulation layers, and vertically heterogeneous total aerosol structures, examining seasonal and interannual changes in aerosol vertical distribution, and evaluating whether satellite products and chemical transport models reproduce the observed altitude dependence of aerosol mass concentrations. Because the same retrieval and quality control framework is applied consistently throughout the full record, the product provides an internally consistent benchmark for long–term comparison across aerosol types, seasons, and years.

Nevertheless, users should consider the documented application boundaries when interpreting absolute aerosol mass concentrations. The retrieved dust, non–dust, and total aerosol products are affected by uncertainties related to the prescribed depolarization characteristics, lidar ratios, aerosol densities, extinction–to–mass conversion factors, and assumed particle optical properties. These uncertainties may increase under humid conditions, in overlapping limited near–surface layers, and in situations involving complex aerosol mixing or hygroscopic growth, especially when transported dust is internally mixed with anthropogenic aerosols or when aerosol composition departs from the assumed non–dust characteristics. We recommend using the RQI that is released simultaneously with the dataset for data screening. In addition, temporal gaps may occur because of clouds, precipitation, low cloud screening, or routine instrument interruptions, and these valid data omissions should be considered when calculating monthly or seasonal averages.

https://doi.org/10.5194/essd-2026-677
Preprint. Discussion started: 23 September 2026

5 Illustrative applications of the dataset

To illustrate the practical use of the dataset, we present two representative examples: a dust–dominated transport event (Fig. 9) and a non–dust pollution episode (Fig. 10). These examples are included to demonstrate how the combined profile information can be used to characterize contrasting aerosol regimes rather than to provide complete process attribution. In both figures, the thermodynamic profiles (relative humidity and temperature) and ERA5 wind vectors are shown together with the lidar–derived EXT, PDR, dust mass concentration, and non–dust mass concentration. The black line indicates the ERA5 PBL height. This combined presentation illustrates how the dataset can be used to relate aerosol type and vertical distribution to the lower–tropospheric structure.

Figure 9 shows an event characterized by enhanced PDR, increased dust mass concentration, and the presence of elevated aerosol layers. During this episode, the increase in PDR coincided with enhanced extinction and a pronounced increase in the retrieved dust component, whereas the non–dust component remained comparatively weak. This event also shows that dust layers can first occur aloft and subsequently extend downward into the lower troposphere. This example demonstrates that the dataset can distinguish dust–dominated conditions from background or weakly depolarizing aerosol conditions and can resolve the vertical evolution of transported dust layers at a high temporal resolution. From a data use perspective, this type of event highlights the value of the product for identifying elevated dust transport, describing the height range of dust influence, and comparing the vertical structure of aerosol types with meteorological contexts, such as wind shifts and boundary layer depth. Therefore, this example is relevant for studies of aerosol transport pathways, satellite overpass evaluation, and model assessment of dust layer altitude and timing.

Figure 10 presents an event in which aerosol loading was dominated by non–dust components. In this case, EXT was concentrated mainly within the lower troposphere, PDR remained low, and the retrieved non–dust mass concentration was much larger than the dust mass concentration. The aerosol layer was largely confined to the PBL and its immediate overlying region, demonstrating the ability of the dataset to capture shallow pollution accumulation and its temporal evolution. This example demonstrates that the dataset can distinguish non–dust–dominated aerosol conditions from dust transport events and describe how pollution–related aerosol mass is vertically distributed relative to the PBL. Such cases are useful for analyzing boundary layer confinement, comparing near–surface and elevated aerosol behavior, and evaluating whether satellite and model products reproduce the observed vertical placement of pollution–related aerosols.

Together, these two examples show that the SXNCO polarization lidar dataset can capture contrasting aerosol–type regimes and their vertical structures at a sub–hourly temporal resolution. The event–based information complements the long–term statistics presented in Section 4 and illustrates the broader reuse potential of the dataset for studies that require vertically resolved aerosol–type information. Simultaneously, these examples should be interpreted with the same caution as long–term products. The retrieved dust and non–dust mass concentrations were affected by the uncertainties described in Sect. 2.4.4, including sensitivity to the assumed depolarization characteristics, extinction–to–mass conversion parameters, humidity effects, and aerosol mixing state.

https://doi.org/10.5194/essd-2026-677
Preprint. Discussion started: 23 September 2026

Data availability

The dataset described in this study is archived at Mendeley Data (https://doi.org/10.17632/kfvpbb3vcs.1; Wang et al., 2026). The repository contains quality–controlled profiles of dust, non-dust, and total aerosol mass concentrations derived from ground–based polarization lidar observations at the Shouxian National Climate Observatory (SXNCO) from January 2018 to December 2024. High–resolution products are provided on the released 15 min time grid and 7.5 m vertical grid, together with a pointwise Retrieval Quality Index (RQI) on the same time–height grid. Monthly mean dust, non–dust, and total aerosol mass concentration products are additionally provided for convenient climatological and interannual analyses. The data are currently distributed in Microsoft Excel (.xlsx) format. In the high–resolution files, the first column specifies observation time in Beijing Time (BJT; UTC+08:00), the first row specifies the vertical coordinate in kilometres, and the internal cells contain aerosol mass concentrations (μg m^{-3}) or RQI values at the corresponding time and height. Blank cells denote missing or invalid retrievals and must not be interpreted as zero aerosol concentration. The accompanying README describe the file structure, variable definitions, units, missing–data convention, RQI interpretation, retrieval assumptions, quality–control procedures, and recommended use of the products. All files are openly available under the Creative Commons Attribution 4.0 International (CC BY 4.0) licence.

Conclusion

This study presents a quality–controlled dataset of dust, non–dust, and total aerosol mass concentration profiles derived from continuous ground–based polarization lidar observations at the SXNCO in central China from 2018 to 2024. The dataset provides vertically resolved aerosol mass concentration information at 15 min temporal resolution and 7.5 m vertical resolution from approximately 0.3–6 km using a consistent processing chain that combines lidar signal screening, optical retrieval, aerosol type separation, and extinction–to–mass conversion. Also, a pointwise RQI (0–100) is provided to facilitate data screening and reuse, with higher RQI values recommended for quantitative analyses and lower values primarily for qualitative interpretation or analyses based on temporal or vertical averaging. As a multi–year profile product from a regionally representative receptor site between the BTH and YRD regions, it adds an observational constraint that is not available from column–only or surface–only measurements alone.

The main strength of the data product is its internally consistent description of the aerosol vertical structure and aerosol type over seven years. The dataset resolves the seasonal and interannual variability of dust and non–dust aerosol mass within the lower and middle troposphere and provides information on aerosol layering, elevated transport, and boundary layer confinement under a unified quality control and retrieval framework. Evaluation against colocated sun–photometer observations, CALIPSO aerosol profile products, and near–surface particulate matter measurements indicated that the dataset reproduces the principal temporal and vertical characteristics of aerosol variability and provides a credible basis for subsequent community use. In this sense, the product is intended to reduce the effort required for future users to obtain harmonized, quality–assessed, aerosol–type–resolved profile observations for central China.

https://doi.org/10.5194/essd-2026-677
Preprint. Discussion started: 23 September 2026

The dataset should be used with attention to its documented application boundaries. Mass retrieval depends on the POLIPHON framework and therefore inherits uncertainties associated with the assumed aerosol depolarization ratios, lidar ratios, aerosol densities, and extinction–to–mass conversion factors. Additional uncertainty can arise under humid conditions, in strongly mixed aerosol environments, and near the lower overlap–limited part of the profile. The released product is thus best interpreted as a quality–controlled observational constraint with quantified uncertainty, rather than as an exact gravimetric representation of atmospheric aerosol mass at every height and time. Users should consider these limitations, particularly when analyzing individual profiles, high–humidity episodes, or situations involving complex internal mixing between dust and anthropogenic particles.

Despite these limitations, the dataset has substantial reuse potential. It can serve as a ground–based benchmark for evaluating satellite aerosol products and chemical transport models, especially in studies that require a realistic aerosol vertical structure rather than only column–integrated or near–surface quantities. It can also support process–oriented analyses of aerosol transport, boundary layer coupling, and extreme aerosol events, as well as data assimilation and regional climatological assessments that benefit from long–term vertically resolved constraints.

Looking ahead, the new–generation polarization Raman lidar deployed at the SXNCO in July 2024 provides an important basis for the next phase of dataset development. The upgraded system is equipped with both depolarization and Raman channels at 532 and 355 nm, respectively, enabling more direct constraints on aerosol optical properties and helping reduce retrieval uncertainties associated with assumed lidar ratios and aerosol type parameters. Building on this upgraded observational capability, future work will generate higher–precision aerosol profile products and release a follow–up dataset harmonized with the present 2018–2024 record, thereby extending the continuity of aerosol vertical structure observations in central China.

Author contributions

Yujia Chen, Chengzhi Xing, and Cheng Liu conceived and supervised the study. Zhuang Wang performed the data analysis and drafted the original manuscript. Yujia Chen, Chengzhi Xing, Chengxiao Liu, Yanfeng Huo, Xiaoyun Sun, Hao Zhang, Chune Shi, and Cheng Liu reviewed and edited the manuscript. Yizhi Zhu, Congzi Xia, Kaidi Zhang, Yong Zhu, and Xintong Chen validated the dataset. Xinfeng Lin, Shaowei Yan, Jiaqi Chen, Nan Ge, Guanyin Yang, and Zhenzhen Hua maintained the lidar and assisted with data processing.

Acknowledgments.

We would like to thank the ERA5 data developers for providing free and open source materials.

https://doi.org/10.5194/essd-2026-677
Preprint. Discussion started: 23 September 2026

Open Access Earth System Science Data Discussions

Competing interests

The authors declare that they have no competing interests.

Financial support

This research was supported by grants from the National Natural Science Foundation of China (42507158), China Meteorological Administration Innovation Development Special Project (CXFZ2026J113), the President's Foundation of Hefei institutes of Physical Science, Chinese Academy of Sciences (BJPY2024B09), and China Meteorological Administration "Application of quantum technology in meteorological detection" Youth Innovation Team Project (No.CMA2024QN11). the Innovation and Development Special Fund of China Meteorological Administration under Grant No. CXFZ2026QB04, and the Innovation and Development Special Fund of Anhui Provincial Meteorological Bureau under Grant No. CXB202405.

References

Adachi, K., Yoshida, A., Mori, T., Moteki, N., Ohata, S., Kita, K., Kawai, Y., and Koike, M.: Individual particle compositions and aerosol mixing states at different altitudes over the ocean in East Asia, Atmospheric Chemistry and Physics, 25, 12599-12613, 10.5194/acp-25-12599-2025, 2025.

Ansmann, A., Seifert, P., Tesche, M., and Wandinger, U.: Profiling of fine and coarse particle mass: case studies of Saharan dust and Eyjafjallajökull/Grimsvötn volcanic plumes, Atmospheric Chemistry and Physics, 12, 9399-9415, 10.5194/acp-12-9399-2012, 2012.

Ansmann, A., Mamouri, R.-E., Hofer, J., Baars, H., Althausen, D., and Abdullaev, S. F.: Dust mass, cloud condensation nuclei, and ice-nucleating particle profiling with polarization lidar: updated POLIPHON conversion factors from global AERONET analysis, Atmospheric Measurement Techniques, 12, 4849-4865, 10.5194/amt-12-4849-2019, 2019.

Ansmann, A., Tesche, M., Seifert, P., Groß, S., Freudenthaler, V., Apituley, A., Wilson, K. M., Serikov, I., Linné, H., Heinold, B., Hiebsch, A., Schnell, F., Schmidt, J., Mattis, I., Wandinger, U., and Wiegner, M.: Ash and fine-mode particle mass profiles from EARLINET-AERONET observations over central Europe after the eruptions of the Eyjafjallajökull volcano in 2010, Journal of Geophysical Research, 116, 10.1029/2010jd015567, 2011.

Balmes, K. A., Fu, Q., and Thorsen, T. J.: The Diurnal Variation of the Aerosol Optical Depth at the ARM SGP Site, Earth and Space Science, 8, 10.1029/2021ea001852, 2021.

Bi, J., Li, B., Ma, J., Wang, X., Meng, Z., Shi, J., and Li, Z.: Evaluation of Nocturnal Aerosol Optical Depth Determining From a Lunar Photometry in Lanzhou of Northwest China, Journal of Geophysical Research: Atmospheres, 129, 10.1029/2023jd040387, 2024.

Bravo-Aranda, J. A., Titos, G., Granados-Muñoz, M. J., Guerrero-Rascado, J. L., Navas-Guzmán, F., Valenzuela, A., Lyamani, H., Olmo, F. J., and Andrey, J.: Study of mineral dust entrainment in the planetary boundary layer by
lidar depolarisation technique, Tellus B: Chemical and Physical Meteorology, 67, 10.3402/tellusb.v67.26180, 2015.

Chang, L., Li, J., Ren, J., Xiong, C., and Zhang, L.: An iterative algorithm to simultaneously retrieve aerosol extinction and effective radius profiles using CALIOP, Atmospheric Measurement Techniques, 17, 2637-2648, 10.5194/amt-17-2637-2024, 2024.

Fan, W., Qin, K., Xu, J., Yuan, L., Li, D., Jin, Z., and Zhang, K.: Aerosol vertical distribution and sources estimation at a site of the Yangtze River Delta region of China, AtmRe, 217, 128-136, 10.1016/j.atmosres.2018.11.002, 2019.

Fernald, F. G.: Analysis of atmospheric lidar observations: some comments, Applied optics, 23, 652-653, 1984.

https://doi.org/10.5194/essd-2026-677
Preprint. Discussion started: 23 September 2026

Floutsi, A. A., Baars, H., Engelmann, R., Althausen, D., Ansmann, A., Bohlmann, S., Heese, B., Hofer, J., Kanitz, T., Haarig, M., Ohneiser, K., Radenz, M., Seifert, P., Skupin, A., Yin, Z., Abdullaev, S. F., Komppula, M., Filioglou, M., Giannakaki, E., Stachlewska, I. S., Janicka, L., Bortoli, D., Marinou, E., Amiridis, V., Gialitaki, A., Mamouri, R.-E., Barja, B., and Wandinger, U.: DeLiAn – a growing collection of depolarization ratio, lidar ratio and Ångström exponent for different aerosol types and mixtures from ground-based lidar observations, Atmospheric Measurement Techniques, 16, 2353-2379, 10.5194/amt-16-2353-2023, 2023.

Freudenthaler, V., Esselborn, M., Wiegner, M., Heese, B., Tesche, M., Ansmann, A., Müller, D., Althausen, D., Wirth, M., and Fix, A.: Depolarization ratio profiling at several wavelengths in pure Saharan dust during SAMUM 2006, Tellus B, 61, 165-179, 2009.

Haarig, M., Walser, A., Ansmann, A., Dollner, M., Althausen, D., Sauer, D., Farrell, D., and Weinzierl, B.: Profiles of cloud condensation nuclei, dust mass concentration, and ice-nucleating-particle-relevant aerosol properties in the Saharan Air Layer over Barbados from polarization lidar and airborne in situ measurements, Atmospheric Chemistry and Physics, 19, 13773-13788, 10.5194/acp-19-13773-2019, 2019.

He, Y., Choudhury, G., Tesche, M., Ansmann, A., Yi, F., Müller, D., and Yin, Z.: Extended POLIPHON dust conversion factor dataset for lidar-derived cloud condensation nuclei and ice-nucleating particle concentration profiles, Atmospheric Measurement Techniques, 18, 5669-5685, 10.5194/amt-18-5669-2025, 2025.

He, Y., Yin, Z., Ansmann, A., Liu, F., Wang, L., Jing, D., and Shen, H.: POLIPHON conversion factors for retrieving dust-related cloud condensation nuclei and ice-nucleating particle concentration profiles at oceanic sites, Atmospheric Measurement Techniques, 16, 1951-1970, 10.5194/amt-16-1951-2023, 2023.

Heese, B., Flentje, H., Althausen, D., Ansmann, A., and Frey, S.: Ceilometer lidar comparison: backscatter coefficient retrieval and signal-to-noise ratio determination, Atmospheric Measurement Techniques, 3, 1763-1770, 2010.

Holben, B. N., Eck, T. F., Slutsker, I., Tanre, D., Buis, J., Setzer, A., Vermote, E., Reagan, J., Kaufman, Y., and Nakajima, T.: AERONET—A federated instrument network and data archive for aerosol characterization, Remote sensing of environment, 66, 1-16, 1998.

Hu, Y., Liu, Z., Winker, D., Vaughan, M., Noel, V., Bissonnette, L., Roy, G., and McGill, M.: Simple relation between lidar multiple scattering and depolarization for water clouds, Optics Letters, 31, 1809-1811, 10.1364/OL.31.001809, 2006.

Huang, X., Ding, A. J., Wang, Z. L., Ding, K., Gao, J., Chai, F. H., and Fu, C. B.: Amplified transboundary transport of haze by aerosol–boundary layer interaction in China, Nature Geoscience, 13, 428–434, 10.1038/s41561-020-0583-4, 2020.

Huang, J. P., Liu, J. J., Chen, B., and Nasiri, S. L.: Detection of anthropogenic dust using CALIPSO lidar measurements, Atmospheric Chemistry and Physics, 15, 11653-11665, 10.5194/acp-15-11653-2015, 2015.

Hunt, W. H., Winker, D. M., Vaughan, M. A., Powell, K. A., Lucker, P. L., and Weimer, C.: CALIPSOLidar Description and Performance Assessment, JAtOT, 26, 1214-1228, 10.1175/2009jtecha1223.1, 2009.

Jing, D., He, Y., Yin, Z., Huang, K., Liu, F., and Yi, F.: Evolution of tropospheric aerosols over central China during 2010–2024 as observed by lidar, Atmospheric Chemistry and Physics, 25, 17047-17067, 10.5194/acp-25-17047-2025, 2025.

Julaha, K., Ždímal, V., Mbengue, S., Brus, D., and Zíková, N.: Drone-based vertical profiling of particulate matter size distribution and carbonaceous aerosols: urban vs. rural environment, Atmospheric Chemistry and Physics, 25, 17933-17951, 10.5194/acp-25-17933-2025, 2025.

Kim, M., Omar, A., Tackett, J. L., Vaughan, M. A., Winker, D. M., Trepte, C. R., Hu, Y., Liu, Z., Poole, L. R., and Pitts, M. C.: The CALIPSO version 4 automated aerosol classification and lidar ratio selection algorithm, Atmospheric Measurement Techniques, 11, 6107-6135, 2018a.

Kim, M. H., Omar, A. H., Tackett, J. L., Vaughan, M. A., Winker, D. M., Trepte, C. R., Hu, Y., Liu, Z., Poole, L. R., Pitts, M. C., Kar, J., and Magill, B. E.: The CALIPSO Version 4 Automated Aerosol Classification and Lidar Ratio Selection Algorithm, Atmos Meas Tech, 11, 6107-6135, 10.5194/amt-11-6107-2018, 2018b.

Liu, J., Zhang, F., Ren, J., Chen, L., Zhang, A., Wang, Z., Zou, S., Xu, H., and Yue, X.: The evolution of aerosol mixing state derived from a field campaign in Beijing: implications for particle aging timescales in urban atmospheres, Atmospheric Chemistry and Physics, 25, 5075-5086, 10.5194/acp-25-5075-2025, 2025.

Mamouri, R.-E. and Ansmann, A.: Potential of polarization/Raman lidar to separate fine dust, coarse dust, maritime, and anthropogenic aerosol profiles, Atmospheric Measurement Techniques, 10, 3403-3427, 10.5194/amt-10-3403-2017, 2017.

Mamouri, R. E. and Ansmann, A.: Fine and coarse dust separation with polarization lidar, Atmospheric Measurement Techniques, 7, 3717-3735, 10.5194/amt-7-3717-2014, 2014.

https://doi.org/10.5194/essd-2026-677
Preprint. Discussion started: 23 September 2026

760 Midzak, N., Yorks, J. E., and Zhang, J.: Statistics of Smoke Sphericity and Optical Properties Using Spaceborne Lidar Measurements, Remote Sensing, 17, 10.3390/rs17030409, 2025.
Omar, A. H., Winker, D. M., Tackett, J. L., Giles, D. M., Kar, J., Liu, Z., Vaughan, M. A., Powell, K. A., and Trepte, C. R.: CALIOP and AERONET aerosol optical depth comparisons: One size fits none, Journal of Geophysical Research: Atmospheres, 118, 4748-4766, 10.1002/jgrd.50330, 2013.
765 Ozdemir, E., Tuna Tuygun, G., and Elbir, T.: Application of aerosol classification methods based on AERONET version 3 product over eastern Mediterranean and Black Sea, Atmospheric Pollution Research, 10.1016/j.apr.2020.06.008, 2020.
Pan, X., Uno, I., Wang, Z., Nishizawa, T., Sugimoto, N., Yamamoto, S., Kobayashi, H., Sun, Y., Fu, P., Tang, X., and Wang, Z.: Real-time observational evidence of changing Asian dust morphology with the mixing of heavy anthropogenic pollution, Sci. Rep., 7, 335, 10.1038/s41598-017-00444-w, 2017.
770 Pappalardo, G., Wandinger, U., Mona, L., Hiebsch, A., Mattis, I., Amodeo, A., Ansmann, A., Seifert, P., Linné, H., Apituley, A., Alados Arboledas, L., Balis, D., Chaikovsky, A., D'Amico, G., De Tomasi, F., Freudenthaler, V., Giannakaki, E., Giunta, A., Grigorov, I., Iarlori, M., Madonna, F., Mamouri, R.-E., Nasti, L., Papayannis, A., Pietruczuk, A., Pujadas, M., Rizi, V., Rocadenbosch, F., Russo, F., Schnell, F., Spinelli, N., Wang, X., and Wiegner, M.: EARLINET correlative measurements for CALIPSO: First intercomparison results, Journal of Geophysical Research, 115, 10.1029/2009jd012147, 2010.
775 Patel, P. N., Jiang, J. H., Gautam, R., Gadhavi, H., Kalashnikova, O., Garay, M. J., Gao, L., Xu, F., and Omar, A.: A remote sensing algorithm for vertically resolved cloud condensation nuclei number concentrations from airborne and spaceborne lidar observations, Atmospheric Chemistry and Physics, 24, 2861-2883, 10.5194/acp-24-2861-2024, 2024.
Quan, J., Ma, P., Pan, Y., Cheng, S., Wang, Q., Lan, L., Dai, L., Wang, Y., and Liu, Y.: Vertical mixing of pollutants and secondary inorganic aerosol formations over Beijing, China, Atmospheric Environment, 357,
780 10.1016/j.atmosenv.2025.121312, 2025.
Ryan, R. G., Toms-Hardman, L., Smirnov, A., Harrison, D. P., and Schofield, R.: Measurement report: Aerosol vertical profiling over the Southern Great Barrier Reef using lidar and MAX-DOAS measurements, Atmospheric Chemistry and Physics, 25, 11183-11197, 10.5194/acp-25-11183-2025, 2025.
Shi, Y., Liu, B., Chen, S., Gong, W., Ma, Y., Zhang, M., Jin, S., and Jin, Y.: Characteristics of aerosol within the nocturnal
785 residual layer and its effects on surface PM2.5 over China, Atmospheric Environment, 241, 10.1016/j.atmosenv.2020.117841, 2020.
Shimizu, A.: Continuous observations of Asian dust and other aerosols by polarization lidars in China and Japan during ACE-Asia, Journal of Geophysical Research, 109, 10.1029/2002jd003253, 2004.
Sugimoto, N., Nishizawa, T., Shimizu, A., Matsui, I., and Kobayashi, H.: Detection of internally mixed Asian dust with air
790 pollution aerosols using a polarization optical particle counter and a polarization-sensitive two-wavelength lidar, Journal of Quantitative Spectroscopy and Radiative Transfer, 150, 107-113, 10.1016/j.jqsrt.2014.08.003, 2015.
Song, Y., Xing, C., Liu, C., Lin, J., Wu, H., Liu, T., Lin, H., Zhang, C., Tan, W., Ji, X., Liu, H., and Li, Q.: Evaluation of transport processes over North China Plain and Yangtze River Delta using MAX-DOAS observations, Atmospheric Chemistry and Physics, 23, 1803–1824, 10.5194/acp-23-1803-2023, 2023.
795 Sugimoto, N., Uno, I., Nishikawa, M., Shimizu, A., Matsui, I., Dong, X., Chen, Y., and Quan, H.: Record heavy Asian dust in Beijing in 2002: Observations and model analysis of recent events, Geophysical Research Letters, 30, 2003.
Teri, M., Gasteiger, J., Heimerl, K., Dollner, M., Schöberl, M., Seibert, P., Tipka, A., Müller, T., Aryasree, S., Kandler, K., and Weinzierl, B.: Pollution affects Arabian and Saharan dust optical properties in the eastern Mediterranean, Atmospheric Chemistry and Physics, 25, 6633-6662, 10.5194/acp-25-6633-2025, 2025.
800 Tesche, M., Wandinger, U., Ansmann, A., Althausen, D., Müller, D., and Omar, A. H.: Ground-based validation of CALIPSO observations of dust and smoke in the Cape Verde region, Journal of Geophysical Research: Atmospheres, 118, 2889-2902, 10.1002/jgrd.50248, 2013.
Tesche, M., Ansmann, A., Müller, D., Althausen, D., Engelmann, R., Freudenthaler, V., and Groß, S.: Vertically resolved separation of dust and smoke over Cape Verde using multiwavelength Raman and polarization lidars during Saharan Mineral
805 Dust Experiment 2008, Journal of Geophysical Research: Atmospheres, 114, 2009.
Tesche, M., Müller, D., Gross, S., Ansmann, A., Althausen, D., Freudenthaler, V., Weinzierl, B., Veira, A., and Petzold, A.: Optical and microphysical properties of smoke over Cape Verde inferred from multiwavelength lidar measurements, Tellus B: Chemical and Physical Meteorology, 63, 677-694, 10.1111/j.1600-0889.2011.00549.x, 2017.

https://doi.org/10.5194/essd-2026-677
Preprint. Discussion started: 23 September 2026

Wandinger, U., Tesche, M., Seifert, P., Ansmann, A., Müller, D., and Althausen, D.: Size matters: Influence of multiple scattering on CALIPSO light-extinction profiling in desert dust, Geophysical Research Letters, 37, 10.1029/2010gl042815, 2010.

Wang, Y., Bagya Ramesh, C., Giangrande, S. E., Fast, J., Gong, X., Zhang, J., Tolga Odabasi, A., Oliveira, M. V. B., Matthews, A., Mei, F., Shilling, J. E., Tomlinson, J., Wang, D., and Wang, J.: Examining the vertical heterogeneity of aerosols over the Southern Great Plains, Atmospheric Chemistry and Physics, 23, 15671-15691, 10.5194/acp-23-15671-2023, 2023a.

Wang, Z., Liu, C., Dong, Y., Hu, Q., Liu, T., Zhu, Y., and Xing, C.: Profiling of Dust and Urban Haze Mass Concentrations during the 2019 National Day Parade in Beijing by Polarization Raman Lidar, Remote Sensing, 13, 3326, 10.3390/rs13163326, 2021a.

Wang, Z., Liu, C., Hu, Q., Dong, Y., Liu, H., Xing, C., and Tan, W.: Quantify the Contribution of Dust and Anthropogenic Sources to Aerosols in North China by Lidar and Validated with CALIPSO, Remote Sensing, 13, 1811, 10.3390/rs13091811, 2021b.

Wang, Z., Shi, C., Zhang, H., Chen, Y., Chi, X., Xia, C., Wang, S., Zhu, Y., Zhang, K., Chen, X., Xing, C., and Liu, C.: Measurement report: Dust and anthropogenic aerosols' vertical distributions over northern China dense aerosols gathered at the top of the mixing layer, Atmospheric Chemistry and Physics, 23, 14271-14292, 10.5194/acp-23-14271-2023, 2023b.

Wang, Z., Shi, C., Zhang, H., Ji, X., Zhu, Y., Xia, C., Sun, X., Lin, X., Yan, S., Wang, S., Zhou, Y., Xing, C., Chen, Y., and Liu, C.: Opposing trends in the peak and low ozone concentrations in eastern China: anthropogenic and meteorological influences, Atmospheric Chemistry and Physics, 25, 347-366, 10.5194/acp-25-347-2025, 2025.

Wang, Z., Liu, C., Xie, Z., Hu, Q., Andreae, M. O., Dong, Y., Zhao, C., Liu, T., Zhu, Y., Liu, H., Xing, C., Tan, W., Ji, X., Lin, J., and Liu, J.: Elevated dust layers inhibit dissipation of heavy anthropogenic surface air pollution, Atmospheric Chemistry and Physics, 20, 14917-14932, 10.5194/acp-20-14917-2020, 2020.

Wang, Z. Chen, Y.: A dataset of high–spatiotemporal–resolution dust and non–dust aerosol mass concentration profiles from ground–based remote sensing in central China (2018–2024), Mendeley Data, V1, doi: 10.17632/kfvpbb3vcs.1, 2026.

Winker, D. M., Vaughan, M. A., Omar, A., Hu, Y., Powell, K. A., Liu, Z., Hunt, W. H., and Young, S. A.: Overview of the CALIPSO Mission and CALIOP Data Processing Algorithms, JAtOT, 26, 2310-2323, 10.1175/2009jtecha1281.1, 2009a.

Winker, D. M., Rogers, R. R., Trepte, C. R., Lee, K.-P., Vaughan, M. A., Hostetler, C. A., Powell, K. A., Liu, Z., Kuehn, R. E., Hunt, W. H., and Young, S. A.: CALIPSO Lidar Calibration Algorithms. Part I: Nighttime 532-nm Parallel Channel and 532-nm Perpendicular Channel, JAtOT, 26, 2015-2033, 10.1175/2009jtecha1242.1, 2009b.

Yan, F., Su, H., Cheng, Y., Huang, R., Liao, H., Yang, T., Zhu, Y., Zhang, S., Sheng, L., Kou, W., Zeng, X., Xiang, S., Yao, X., Gao, H., and Gao, Y.: Frequent haze events associated with transport and stagnation over the corridor between the North China Plain and Yangtze River Delta, Atmospheric Chemistry and Physics, 24, 2365-2376, 10.5194/acp-24-2365-2024, 2024.

Yan, S., Zhu, B., Huang, Y., Zhu, J., Kang, H., Lu, C., and Zhu, T.: To what extents do urbanization and air pollution affect fog?, Atmospheric Chemistry and Physics, 20, 5559-5572, 10.5194/acp-20-5559-2020, 2020.

Yorks, J. E., Wang, J., McGill, M. J., Follette-Cook, M., Nowottnick, E. P., Reid, J. S., Colarco, P. R., Zhang, J., Kalashnikova, O., Yu, H., Marenco, F., Santanello, J. A., Weckwerth, T. M., Li, Z., Campbell, J. R., Yang, P., Diao, M., Noel, V., Meyer, K. G., Carr, J. L., Garay, M., Christian, K., Bennedetti, A., Ring, A. M., Crawford, A., Pavolonis, M. J., Aquila, V., Kim, J., and Kondragunta, S.: A SmallSat Concept to Resolve Diurnal and Vertical Variations of Aerosols, Clouds, and Boundary Layer Height, Bulletin of the American Meteorological Society, 104, E815-E836, 10.1175/bams-d-21-0179.1, 2023.

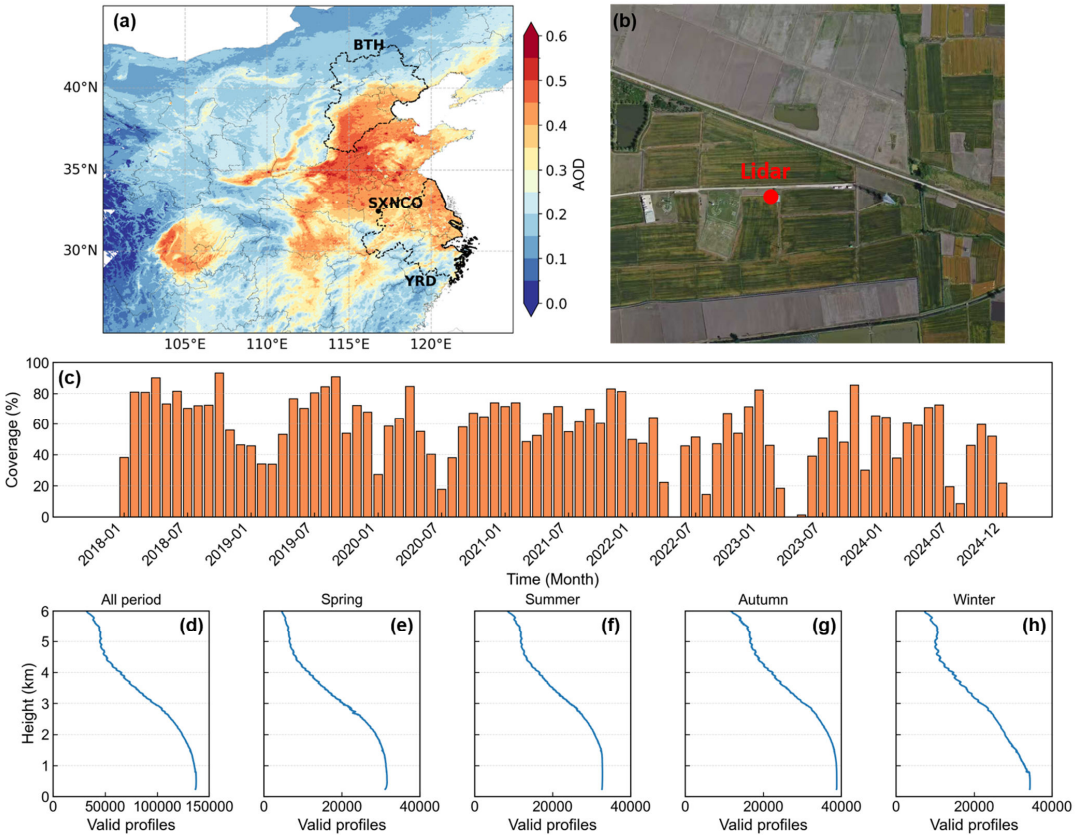
Zhang, Q., Zheng, Y., Tong, D., Shao, M., Wang, S., Zhang, Y., Xu, X., Wang, J., He, H., Liu, W., Ding, Y., Lei, Y., Li, J., Wang, Z., Zhang, X., Wang, Y., Cheng, J., Liu, Y., Shi, Q., Yan, L., Geng, G., Hong, C., Li, M., Liu, F., Zheng, B., Cao, J., Ding, A., Gao, J., Fu, Q., Huo, J., Liu, B., Liu, Z., Yang, F., He, K., and Hao, J.: Drivers of improved PM2.5 air quality in China from 2013 to 2017, Proceedings of the National Academy of Sciences of the United States of America, 116, 24463-24469, 10.1073/pnas.1907956116, 2019.

Zhang, Y., Wei, F., Luo, H., Su, Y., Feng, Y., Huang, Y., Yang, J., Wu, Y., and Han, Y.: Aerosol Physical-Optical Properties and PBLH Under Different Air Pollution Levels From Ground Lidar and Satellite Observations Over Shouxian Area, China, Journal of Geophysical Research: Atmospheres, 126, 10.1029/2021jd035236, 2021.

Zhang, J., Yuan, Q., Liu, L., Wang, Y., Zhang, Y., Xu, L., Pang, Y., Zhu, Y., Niu, H., Shao, L., Yang, S., Liu, H., Pan, X., Shi, Z., Hu, M., Fu, P., and Li, W.: Trans-regional transport of haze particles from the North China Plain to Yangtze River Delta during winter, Journal of Geophysical Research: Atmospheres, 126, e2020JD033778, 10.1029/2020JD033778, 2021.

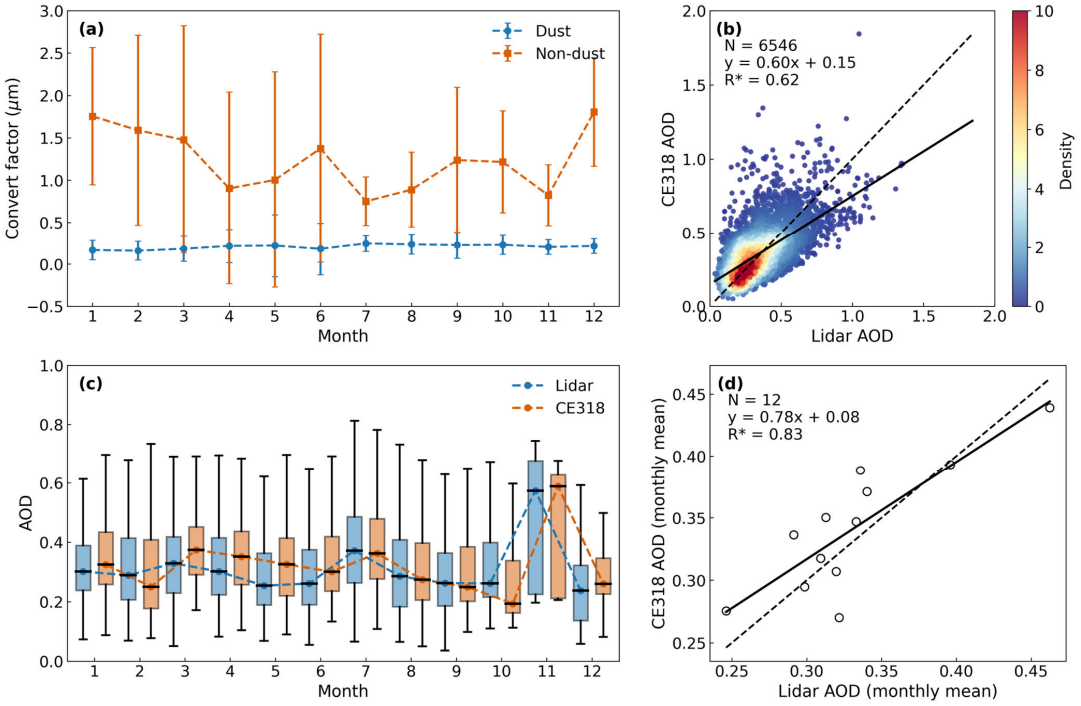
https://doi.org/10.5194/essd-2026-677
Preprint. Discussion started: 23 September 2026

Open Access Earth System Science Data Discussions



860 **Figure 1: Overview of lidar data coverage and associated locations. (a) MODIS–derived annual average AOD from 2018 to 2024 across Eastern China, with the boundaries of the BTH, YRD, and SXNCO regions outlined on the map. (b) Aerial photograph of the SXNCO showing the location of the lidar system (red dot) in the fields. (c) Monthly coverage percentage of the lidar data from 2018 to 2024, with coverage computed for each month. Statistics of the number of valid lidar profiles within the 0–6 km range: (d) entire period (2018–2024), (e) spring, (f) summer, (g) autumn, and (h) winter.**

https://doi.org/10.5194/essd-2026-677
Preprint. Discussion started: 23 September 2026



865 **Figure 2: Comparison of AOD from Lidar and CE–318. (a) Monthly variation of the conversion factors derived from CE–318 for dust and non–dust aerosols; error bars represent ±1 standard deviation. (b) Scatter plot of hourly Lidar AOD versus CE–318 AOD, with density contours, the linear regression line (black solid line), and 1:1 line (black dashed line). The sample size and fitting parameters are presented in the upper–left corner. (c) Boxplots of monthly AOD values for Lidar and CE–318; the box–and–whisker plots show the 5th, 25th, 50th, 75th, and 95th percentiles. (d) Monthly mean comparison of Lidar and CE–318 AOD with a scatter**
870 **plot showing the linear regression line (solid) and 1:1 line (dashed); the sample size and fitting parameters are shown in the upper–left corner. An asterisk (*) next to R indicates significance at P < 0.01.**

https://doi.org/10.5194/essd-2026-677
Preprint. Discussion started: 23 September 2026

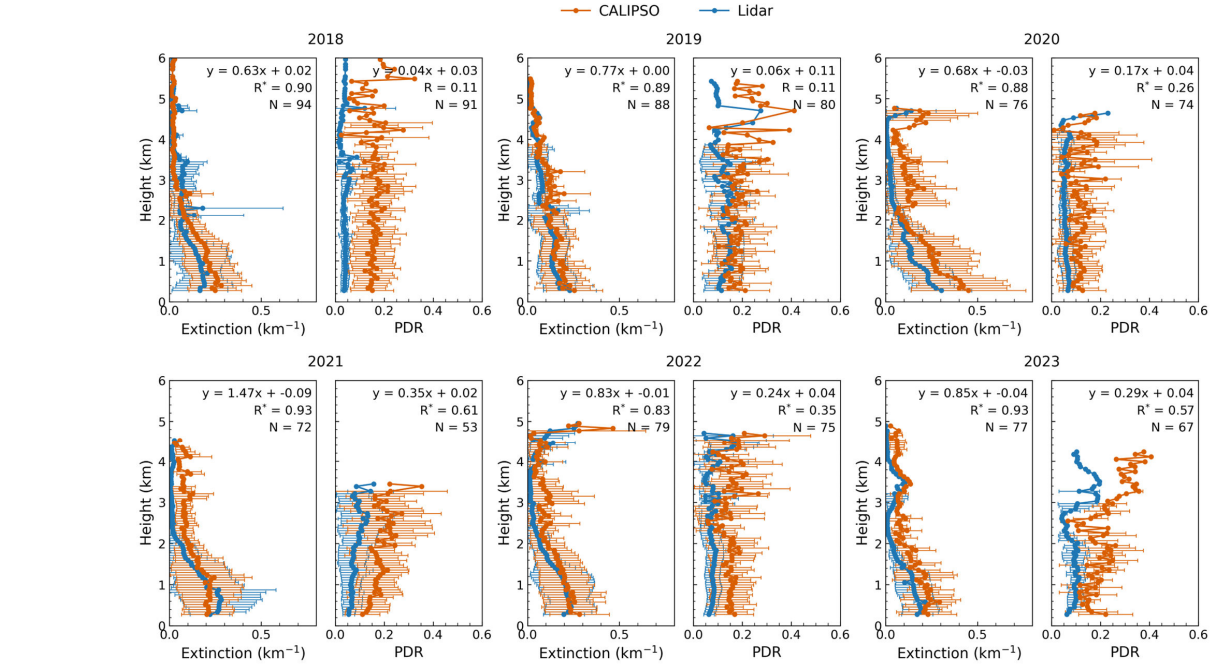


Figure 3: Comparison of annual mean EXT and PDR profiles retrieved by CALIPSO and ground–based lidar for 2018–2023. Each row represents a year, with EXT (left) and PDR (right) profiles for CALIPSO (orange) and lidar (blue). Error bars represent ±1 standard deviation. The fitting parameters are presented in the top right corner of each plot. An asterisk (*) next to R indicates significance at P < 0.01.

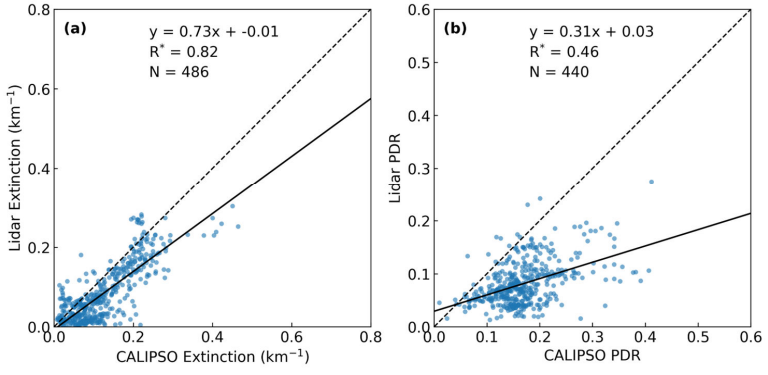


Figure 4: Scatter plots of the annual mean EXT (a) and PDR (b) retrieved by CALIPSO and ground–based lidar for 2018–2023. The black dashed line is the 1:1 line, and the solid line shows the linear least–squares fit. The sample size and fitting parameters are shown in the upper corner. An asterisk (*) next to R indicates significance at P < 0.01.

https://doi.org/10.5194/essd-2026-677
Preprint. Discussion started: 23 September 2026

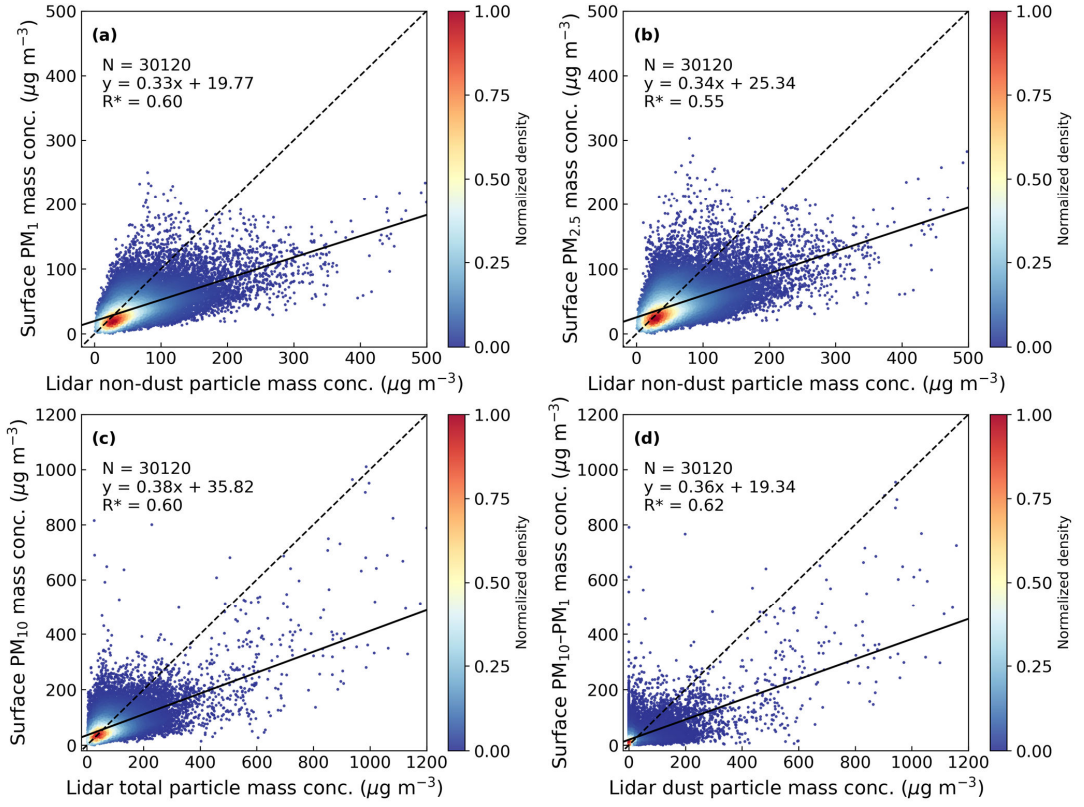


Figure 5: Comparison between lidar–derived particle mass concentration and surface in situ PM measurements using density–colored scatter plots. (a) Scatter plot of lidar–derived non–dust particle mass concentration versus surface PM_1 mass concentration. (b) Lidar–derived non–dust particle mass concentration versus surface $PM_{2.5}$ mass concentration. (c) Lidar–derived total particle mass concentration versus surface PM_{10} mass concentration. (d) Lidar–derived dust particle mass concentration versus surface PM_{10}-PM_1 (serving as a dust proxy) mass concentration. The sample size and fitting parameters are presented in the upper–left corner of each panel. An asterisk (*) next to R indicates significance at $P < 0.01$.

https://doi.org/10.5194/essd-2026-677
Preprint. Discussion started: 23 September 2026

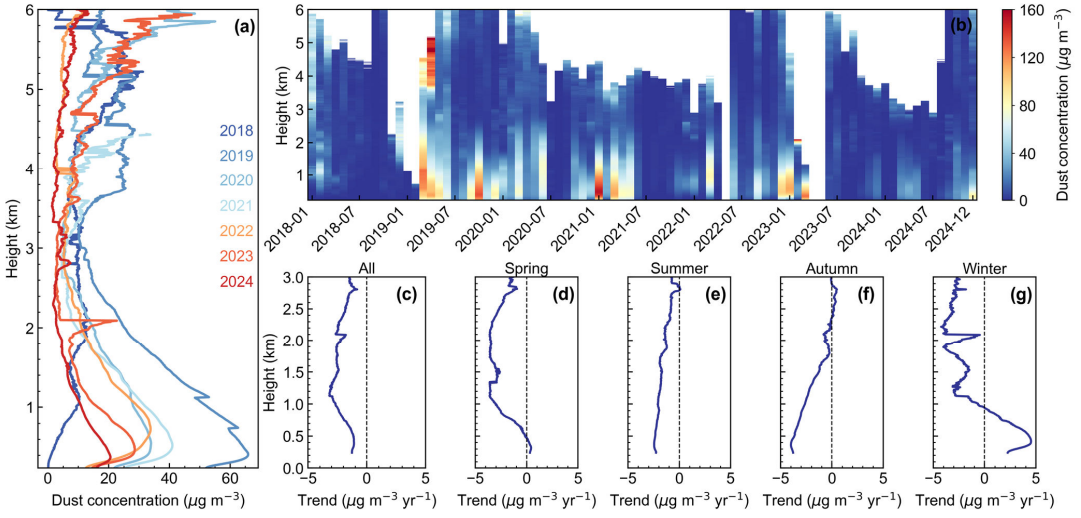


Figure 6: Annual mean profiles and seasonal changes analysis of dust mass concentration from 2018 to 2024. (a) Annual mean dust mass concentration profiles across 0–6 km. (b) Time series of monthly mean dust mass concentration profiles from 2018 to 2024. Interannual changes of dust mass concentration profiles for (c) the entire period, (d) spring, (e) summer, (f) autumn, and (g) winter.

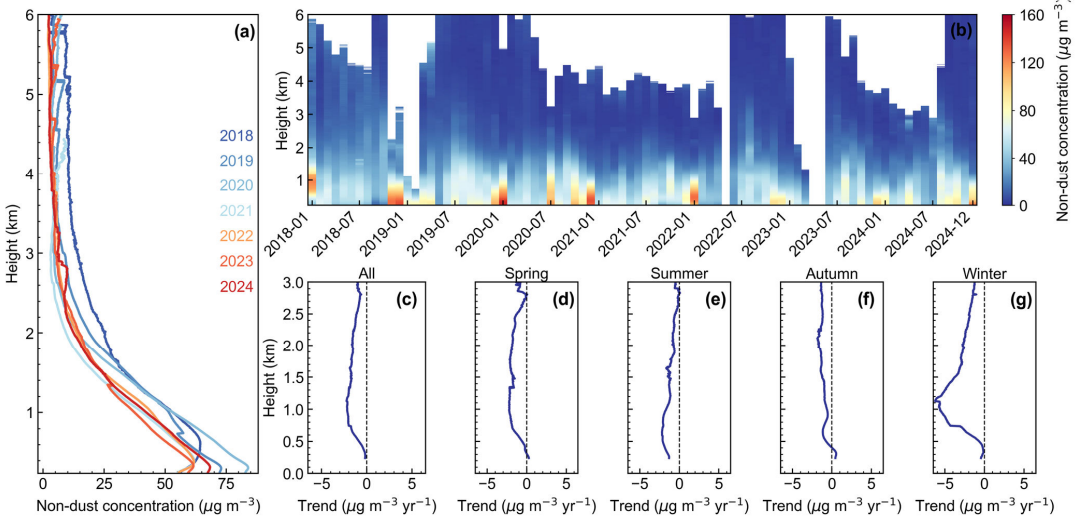


Figure 7: Same as Fig. 6, but for non–dust aerosol mass concentration.

https://doi.org/10.5194/essd-2026-677
Preprint. Discussion started: 23 September 2026

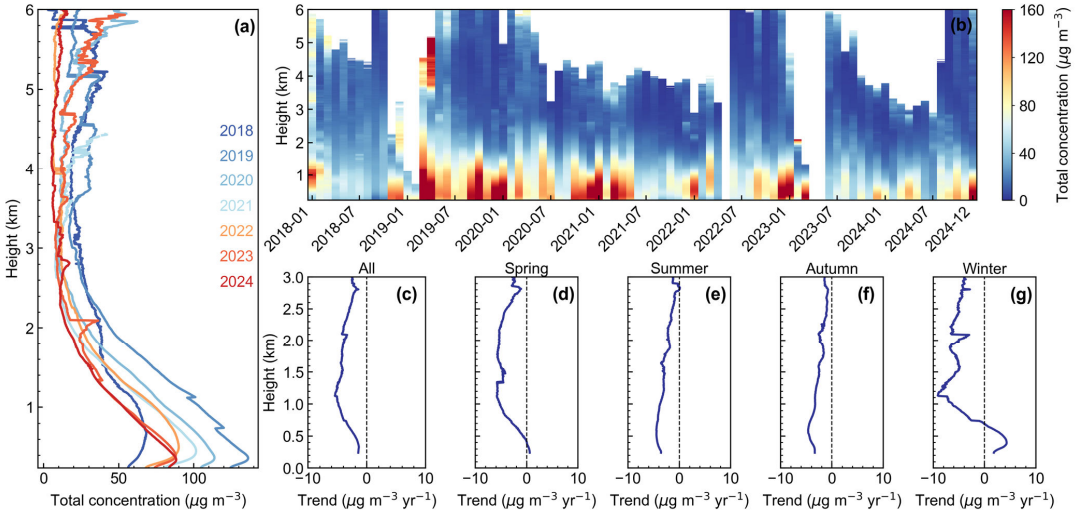


Figure 8: Same as Fig. 6, but for total aerosol mass concentration.

https://doi.org/10.5194/essd-2026-677
Preprint. Discussion started: 23 September 2026

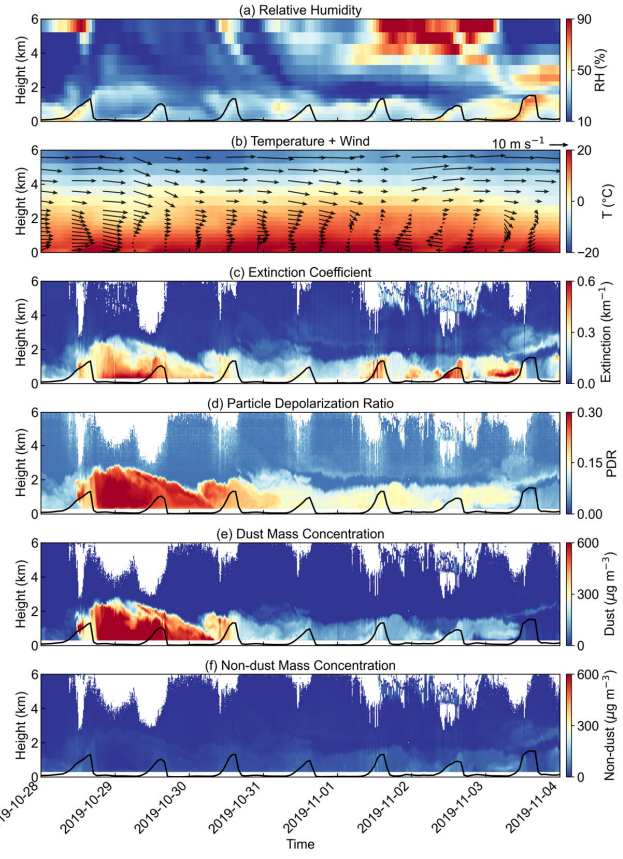


Figure 9: Time–height plots of meteorological conditions and aerosol properties for representative dust events. (a) Relative humidity. (b) Temperature overlaid with ERA5 wind vectors. The wind direction is indicated by the arrow orientation (downward denotes northerly flow), and the wind speed is indicated by the arrow's length. Lidar–retrieved (c) EXT, (d) PDR, (e) dust mass concentration, and (f) non–dust mass concentration. The black line in panels (a) and (c–f) denotes the PBL height.

https://doi.org/10.5194/essd-2026-677
Preprint. Discussion started: 23 September 2026

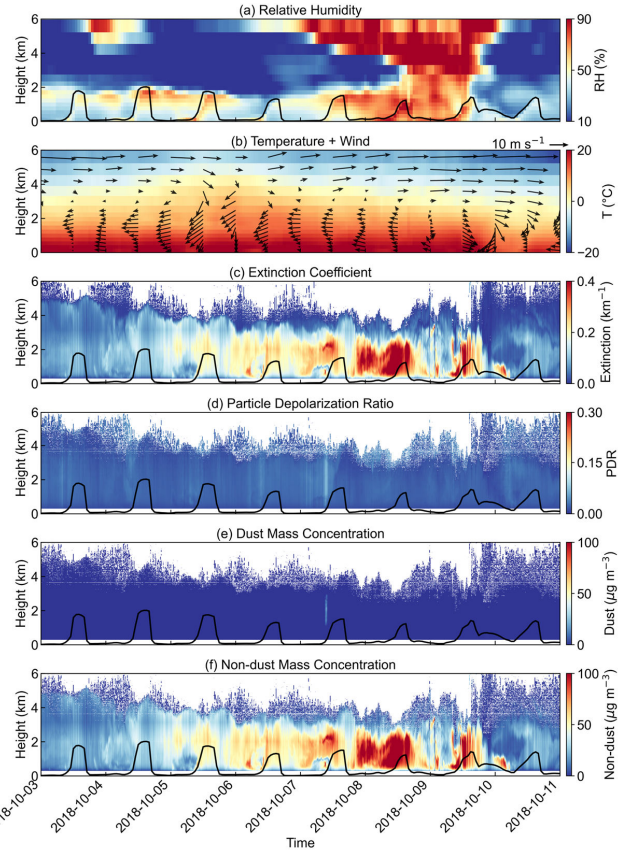


Figure 10: Same as Fig. 9, but for a representative non–dust pollution event.