

A Physics-Constrained Deep-Learning Framework based on Long-Term Remote-Sensing Data for Retrieving Vertical Distribution of PM_{2.5} Chemical Components

Hongyi Li¹, Ting Yang¹, Yele Sun¹, Zifa Wang^{1,2}

5 ¹ State Key Laboratory of Atmospheric Environment and Extreme Meteorology, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing 100029, China

² College of Earth and Planetary Sciences, University of Chinese Academy of Sciences, Beijing 100049, China

Correspondence to: Ting Yang (tingyang@mail.iap.ac.cn)

Abstract. The vertical distribution of PM_{2.5} chemical components is crucial for identifying the causes of atmospheric
10 pollution and its impact on climate change and extreme weather. By integrating long-term lidar measurements, deep-learning
algorithms and a physics-constrained optimization method, this paper presents a novel lidar-based retrieval framework to
obtain vertical mass concentration profiles of PM_{2.5} chemical components for the first time. Identifiable components include
sulfate (SO₄²⁻), nitrate (NO₃⁻), ammonium (NH₄⁺), organic matter (OM) and black carbon (BC), which extend beyond the
component types that traditional remote-sensing retrievals can identify. A 1-year retrieved surface mass concentrations of
15 these components closely aligned with the observations, with Pearson correlation coefficient values ranging from 0.87 to
0.97. The retrieval framework applied to varying non-training spatiotemporal scenarios showed moderate generalization
capability, although a tendency toward underestimation is observed. Tower and aircraft-based field campaigns indicate that
the retrieved and observed vertical profiles of these components exhibited consistent patterns in mass concentrations and
proportions. Subsequently, an explainable method was incorporated into the retrieval framework to quantify the multivariate
20 driving effects on vertical profile retrieval. Results showed that the extinction coefficient and representative indicators within
physiochemical processes contributed significantly to mass concentrations of these components. Finally, a dataset of vertical
mass concentration profiles of these components over six years in a Chinese megacity (Beijing) was generated by the
retrieval framework, revealing the dominant roles of OM and NO₃⁻ in PM_{2.5} throughout the entire boundary layer across all
seasons. As a result of the continued implementation of clean air policies in China, these components exhibited significant
25 decreases during 2021-2022 compared with 2017-2018. Our retrieval framework offers a novel approach for acquiring
vertical profiles of PM_{2.5} chemical components, thereby providing a new perspective on elucidating the vertical evolution of
atmospheric pollutants.

1 Introduction

PM_{2.5} is a complex mixture composed of varying chemical components (Tao et al., 2017), mainly including sulfate (SO₄²⁻), nitrate (NO₃⁻), ammonium (NH₄⁺), organic matter (OM) and black carbon (BC). The diverse physiochemical properties arising from various chemical components yield distinct effects on the environment (Tan et al., 2018), climate change (Menon et al., 2002; Zhu et al., 2024) and human health (Kim et al., 2022). Vertical detection technologies have revealed that chemical components are primarily distributed at varying heights within the atmospheric boundary layer and contribute to environmental pollution through internal physiochemical processes (Morgan et al., 2009; Yang et al., 2024; Sun et al., 2015). Additionally, the proportion and vertical distribution of chemical components can regulate radiation flux at both the top of the atmosphere and at the surface by directly affecting light absorption and scattering, as well as the microphysical properties of clouds, thereby influencing climate change and extreme weather (Zhao et al., 2024). Consequently, characterizing the vertical structures of chemical components is essential for identifying the causes of PM_{2.5} pollution and the response mechanisms related to climate change and extreme weather.

Field campaigns are widely conducted to obtain vertical profiles of PM_{2.5} chemical components by mounting observation instruments on meteorological towers, aircraft, tethered balloons and unmanned aerial vehicles. However, these platforms are constrained by sparse detection sites and heights, limited flight schedules, and high observation costs (Dubey et al., 2022), hindering the time-continuous acquisition of vertical profiles of PM_{2.5} chemical components within the whole boundary layer over a long-term period. Continuous remote-sensing lidar detection technologies with high temporal and vertical resolution serve as robust pathways for the constant identification of PM_{2.5} and its components across all altitudes (Matus et al., 2024; Toth et al., 2022; Wang et al., 2022). Additionally, both satellite-based lidar and ground-based lidar networks, such as China Lidar Joint Observation Network (LiDARNET, <https://lidar.pku.edu.cn/>, last access: 25 July 2025), Asian Dust Network (AD-NET) (Sugimoto et al., 2005), Micro Pulse Lidar Network (MPLNET) (Welton et al., 2001), and European Aerosol Research Lidar Network (EARLINET) (Ansmann et al., 2003), provide remote sensing capabilities with extensive spatial coverage.

Retrieval algorithms for the lidar have been progressively developed over the past 20 years. Earlier studies utilized lidar depolarization ratios to identify dust and non-dust aerosols (Sugimoto et al., 2003; Tesche et al., 2009). Subsequently, additional lidar parameter constraints, such as multi-wavelength backscatter coefficient and lidar ratio, were incorporated to identify dust aerosol, water-soluble aerosols, black carbon, and sea salt based on the assumption of external mixing (Nishizawa et al., 2011; Nishizawa et al., 2017). Hara et al. (2018) considered the hygroscopic growth of water-soluble aerosols and their internal mixing with BC to mitigate the overestimation of BC retrieval (Hara et al., 2018). By integrating the ground-based lidar and sun-photometer, Wang et al. (2022) significantly increased the identifiable aerosol component types, including ammonium nitrate-like, water-insoluble organic matter, water-soluble organic matter, black carbon and fine-

mode aerosol water content (Wang et al., 2022). However, the aerosol component type retrieved from existing lidar retrieval algorithms that utilize aerosol optical properties is not equivalent to the conventional chemical component type. Due to similar optical properties exhibited by PM_{2.5} chemical components, the identification of chemical component types seems to be beyond the scope of remote-sensing retrieval (Wang et al., 2022). Moreover, the multiple parameterization assumptions introduced by existing lidar retrieval algorithms increase the uncertainties in component retrieval.

Data-driven machine learning can interpret the nonlinear relationships between PM_{2.5} chemical components and various driving factors without the constraints imposed by the inherent properties of these components (Li et al., 2025a). Meng et al. (2018) utilized a random forest algorithm to predict national mass concentrations of SO₄²⁻, NO₃⁻, organic carbon (OC) and elemental carbon (EC), achieving R² values ranging from 0.71 to 0.86 on a daily scale (Meng et al., 2018). Based on this algorithm, Lv et al. (2021) further achieved the hourly predictions of the aforementioned chemical components and NH₄⁺ with R values of 0.71-0.81 (Lv et al., 2021). Subsequently, deep learning algorithms are employed to accurately characterize complex nonlinear relationships and effectively extract data features, thereby enhancing the predictive ability of hourly mass concentrations of PM_{2.5} chemical components (Lee et al., 2023; Liu et al., 2023; Li et al., 2025a). However, current studies primarily focus on predicting the ground-level mass concentrations of PM_{2.5} chemical components but cannot interpret the vertical distribution of these components. Furthermore, existing prediction models are susceptible to the quantity and quality of available training data due to the absence of physical constraints, limiting their spatiotemporal generalization capabilities.

In this study, we proposed a novel physics-constrained deep-learning framework that utilized lidar data to retrieve vertical profiles of five PM_{2.5} chemical components (SO₄²⁻, NO₃⁻, NH₄⁺, OM and BC) for the first time. Our retrieval framework effectively mitigates the limitations of remote-sensing retrieval algorithms in identifying chemical components, as well as the deficiencies and limited generalization capabilities of purely data-driven machine learning techniques in characterizing vertical profiles of these components. Detailed descriptions of the retrieval framework and the data utilized are provided in Sect. 2., while Sect. 3 discusses the validation of the retrieval framework, the assessment of feature importance, and applications of this framework. Section 4 presents the conclusion.

2 Data and methodology

2.1 Data

2.1.1 Lidar measurement

The $\sigma_{b_{sc,532}}$ data for deep learning module training and PM_{2.5} chemical component retrieving is obtained from a ground-based dual-wavelength polarization Mie lidar at the Institute of Atmospheric Physics (IAP), Chinese Academy of Sciences (CAS), Beijing (39.98°N, 116.38°E). This Mie lidar has consistently detected optical signals since 2017, offering a temporal

resolution of 15 minutes and a vertical resolution of 6 m. The lidar specification parameters and data preprocessing are detailed in Text S1 (and Table S1) and Text S2 of the supplement, respectively. The $\sigma_{bsc,532}$ data from February 8-15th, 2021 at 23 lidar sites in the North China Plain (NCP), provided by the China National Environmental Monitoring Center (CNEMC), were utilized to assess the spatial generalization ability. The multi-site data offers a temporal resolution of 5-20 minutes and a vertical resolution of 7.5 m. To generate an hourly resolution lidar dataset, minute-level data were resampled using a simple averaging method. Specifically, the arithmetic mean was calculated from all valid minute-level data points within each non-overlapping one-hour window aligned to the start of each hour (e.g., from 00:00 to 00:59).

2.1.2 Auxiliary data for Retrieval

100 The data of multiple meteorological parameters for deep learning module training and PM_{2.5} chemical component retrieving can be obtained from the 5th Generation European Centre for Medium-Range Weather Forecasts (ECMWF) ReAnalysis (ERA5, <https://cds.climate.copernicus.eu/datasets>, last access: 25 July 2025), which provides the hourly data on pressure levels (1000-1 hPa) from 1940 to present with a spatial resolution of $0.25^\circ \times 0.25^\circ$. The data of fine soil, coarse mass and fine sea salt for physics-constrained optimization can be obtained from 4th Generation ECMWF Atmospheric Composition Reanalysis (EAC4, <https://ads.atmosphere.copernicus.eu/datasets>, last access: 25 July 2025), which provides the 3-hourly data on pressure levels (1000-1 hPa) from 2003 to 2024 with a spatial resolution of $0.75^\circ \times 0.75^\circ$. The mass concentration ($\mu\text{g m}^{-3}$) of fine soil is approximately estimated by the mixing ratio (kg kg^{-1}) of dust aerosol with a diameter of 0.03-0.9 μm . The mass concentration ($\mu\text{g m}^{-3}$) of coarse mass is approximately estimated by the mixing ratio (kg kg^{-1}) of dust aerosol with a diameter of 0.9-20 μm . The mass concentration ($\mu\text{g m}^{-3}$) of fine sea salt is approximately estimated by the mixing ratio (kg kg^{-1}) of sea salt aerosol with a diameter of 0.03-5 μm . The pressure levels (hPa) of ERA5 and EAC4 are converted to geometric heights (m), and the 3-hour EAC4 data is converted to hourly data through a linear interpolation method. The grid cells of EAC4 and ERA5 that contain the lidar sites were extracted using the k-nearest neighbor search method based on longitude and latitude data (Friedman et al., 1977). The lidar data and the reanalysis data were interpolated onto a preset vertical grid with a height range of 50 m to 3 km using linear interpolation. The preset height information is presented in Text S2 of the supplement.

2.1.3 Surface observations

Ground-level mass concentrations of NH_4^+ , SO_4^{2-} , NO_3^- , OM and BC at the Beijing lidar site ($39.98^\circ\text{N}, 116.38^\circ\text{E}$) were collected for training the deep learning module and validating retrievals by a high-resolution time-of-flight aerosol mass spectrometer, with a temporal resolution of 1 hour, covering the periods from January 1, 2021, to March 31, 2022, and June 1 to August 31, 2022. Ground-level mass concentrations of the five PM_{2.5} chemical components at 23 non-training NCP sites were provided by CNEMC. Besides, ground-level PM_{2.5} mass concentrations, approximately equal to the sum of the mass concentrations of the five chemical components, are available on CNEMC data release website (<https://www.cnemc.cn/>, last access: 25 July 2025).

2.1.4 Aircraft-based and tower-based measurements

125 The aircraft-based vertical profiles for retrieval independent verification were sampled in a flight experiment at an airport
site in Shijiazhuang (37.54°N, 114.35°E). The flight time schedules (LT, local time) are detailed in Table 1. The tower-based
vertical profiles at altitudes of 16 m, 102 m and 280 m were sampled at a 325-m meteorological tower located at the IAP,
CAS in Beijing (39.98°N, 116.38°E) for 10 days (27 and 30 December 2023; 2, 5, 9, 12, 15, 18, 24, and 27 January 2024). A
flow sampler with a flow rate of 42.8 L min⁻¹ and the 47-mm quartz filter membranes were utilized to collect PM_{2.5} chemical
130 component samples in the aircraft-based and tower-based sampling experiments. Furthermore, the 325 m tower-based
vertical profiles from December 30, 2018 to January 2, 2019 were also collected from Lei et al. (2021)'s study (Lei et al.,
2021).

Table 1: Flight time schedules (LT, local time), corresponding surface temperature and relative humidity.

Date	Flight time	Sampling height (m)	Surface temperature (°C)	Surface relative humidity (%)
September 26 2024	19:10-21:10	2100	19.2-22.9	87.5-95
October 10 2024	19:40-21:40	600	18.8-19.2	29-30
December 11 2024	15:00-16:00	1200	4.3-4.9	31-34
December 11 2024	16:00-17:00	1500	3.3-4.3	34-38

135

2.2 Methodology

2.2.1 Retrieval framework

This paper proposed a novel retrieval framework for retrieving the vertical concentration profiles of five PM_{2.5} chemical
components (NH₄⁺, SO₄²⁻, NO₃⁻, OM and BC) from the lidar aerosol extinction coefficient at 532 nm ($\sigma_{b_{sc},532}$). As shown in
140 Fig. 1, the retrieval framework mainly consists of a deep learning module and a physics-constrained optimization module.
The input datasets of the deep learning module include the surface observation data, meteorological data and ground-based
lidar data (Fig. 1a). Specifically, the aerosol extinction coefficient at 532 nm ($\sigma_{b_{sc},532}$) and multiple meteorological
parameters (u-component wind, v-component wind, temperature, relative humidity, specific humidity, vertical velocity and
geopotential) serve as input features, while the concentrations of the five PM_{2.5} chemical components (NH₄⁺, SO₄²⁻, NO₃⁻,
145 OM and BC) serve as target features. The deep learning module (Fig. 1b), mainly consisting of the Convolutional Neural
Network (CNN), Bidirectional Long Short-Term Memory (BiLSTM), attention mechanism and Bayesian optimization, is
utilized to establish the nonlinear relationship between input and target features. The input datasets of the physics-
constrained optimization module include the ground-based lidar data, aerosol auxiliary data and deep learning intermediate
output (Fig. 1a, c), which provide fundamental input for establishing a multi-object function based on the Interagency
150 Monitoring of Projected Visual Environment (IMPROVE) equation. The physics-constrained optimization module

incorporates the multi-object loss function with the Non-dominated Sorting Genetic Algorithm II (NSGA-II) to implement external physical constraints (Fig. 1c), thus enhancing the extrapolation capability of the deep learning module and generating high-quality vertical concentration profiles of the five PM_{2.5} chemical components. Detailed descriptions of the deep learning algorithms, hyperparameter tuning, and physics-constrained optimization used in this work will be presented in subsequent sections. The brief workflow of the retrieval framework is summarized as follows.

Step 1. The multi-source input datasets undergo matching across spatiotemporal and vertical dimensions. All input and output data are uniformly time-resolved to hourly intervals, while vertical data are uniformly vertically resolved into 10 layers ranging from 50 m to 3 km.

Step 2. The input data of the deep learning module are normalized by Z-score normalization to stabilize the training process, accelerate training convergence, and enhance model robustness (Al-Faiz et al., 2018; Cabello-Solorzano et al., 2023).

Step 3. Training deep learning module by using the normalized surface-level input data.

Step 4. Generating the normalized concentrations of the five PM_{2.5} chemical components at each vertical layer by feeding the normalized height-level input data into the deep learning module.

Step 5. Denormalizing the deep-learning output by using the inverse Z-score transformation, with the mean and standard deviation statistics derived from the original training set, thereby recovering the physical mass concentration unit ($\mu\text{g m}^{-3}$).

Step 6. Optimizing the denormalized deep learning output through implementing an external physics constraint to obtain the high-quality vertical concentration profiles of the five PM_{2.5} chemical components. Repeat steps 4-6 until the retrieval task is complete.

170

Figure 1: Remote-sensing retrieval framework for vertical distribution of five PM_{2.5} chemical components (NH₄⁺, SO₄²⁻, NO₃⁻, OM and BC). (U: U-component wind; V: V-component wind; T: Temperature; RH: Relative Humidity; q: Specific Humidity; w: Vertical Velocity; Z: Geopotential; $\sigma_{ext,532}$: Aerosol Extinction Coefficient at 532 nm; CNN: Convolutional Neural Network; ReLU: Rectified Linear Unit; FC: Fully Connected; BiLSTM: Bidirectional Long Short-Term Memory; IMPROVE: Interagency Monitoring of Projected Visual Environment; NSGA-II: Non-dominated Sorting Genetic Algorithm II).

175

2.2.2 Deep learning and hyperparameter tuning

The deep learning module is the core of the retrieval framework that generates the normalized vertical profiles of five PM_{2.5} chemical components by feeding the lidar-based aerosol extinction coefficient at 532 nm ($\sigma_{ext,532}$) and multiple meteorological parameters. The deep learning module is designed by numerous neural network layers (Fig. 1, red part), including the CNN layer, Average Pooling layer, Rectified Linear Unit (ReLU) layer, Fully Connected (FC) layer, Attention Mechanism layer, Sigmoid layer, Flatten layer, BiLSTM layer, Dropout layer and Regression Output layer. The CNN and BiLSTM layers, coupled with the Attention Mechanism (AM), are designed to effectively capture the multivariate and temporal characteristics in the training data, thereby establishing a robust nonlinear mapping between the input and output features. The hybrid CNN-BiLSTM-AM architecture consistently outperforms single-architecture models in predictive tasks, as evidenced by numerous studies. (Kavianpour et al., 2023; Ma et al., 2022; Shan et al., 2021; Zhang et al., 2023). Other

185

layers are responsible for data input, structural transformation, normalization, nonlinear process, pooling process, neuron removal and data output, enhancing the training performance and preventing overfitting. Here, we review the description of three key layers, and the description of other layers can be found in our previous work (Li et al., 2025a).

190 CNN is a variant of the multilayer perceptron that efficiently identifies the relevant features through local perception, sparse connections and sharing of weight and bias (Alzubaidi et al., 2021). The convolutional layer in CNN performs convolutional computation on input across spatial dimensionality using learnable kernels to extract local features and enhance training efficiency (O'shea and Nash, 2015). Then the convolutional output is typically enhanced nonlinearly by the ReLU layer (Eq. (1)) or down-sampled nonlinearly by the pooling layer in a CNN architecture.

$$195 \quad y_t = \max(0, f(\mathbf{w} \times x_t + b_t)), \quad (1)$$

Where y_t is the nonlinearly enhanced convolutional output at timestep t , $f(w \times x_t + b_t)$ is the original convolutional output at timestep t , x_t is the input data at timestep t , w is the weight matrix and b_t is the bias term.

The attention mechanism layer is incorporated with CNN to amplify the weight of key information and mitigate the interference of redundant information, leading to an enhancement in the quality of the CNN output (Wang and Zhang, 2025).
 200 The attention mechanism is inspired by the ability of human vision to selectively focus on key information (Guo et al., 2022). Our retrieval framework integrates a data-driven channel attention mechanism, which rescales the original feature channels from the convolutional layers through element-wise multiplication using learned attention weights, thereby enhancing the importance of key features and reduce the interference of irrelevant features. The attention weights are generated by the FC
 205 layers with a sigmoid activation function (Eq. (2)) and then performs Schur product operation with CNN multivariate output (Eq. (3)).

$$\mathbf{W} = \text{sigmoid} \left(\text{FC} \left(\text{Pooling}(y_{1,i}) \right) \right), \quad (2)$$

$$y_{2,i} = \mathbf{W} \cdot y_{1,i}, \quad (3)$$

Where $y_{1,i}$ is the CNN multivariate output. Pooling and FC layers are responsible for down-sampling and feature learning,
 210 respectively, thus predicting the importance of i^{th} feature. The sigmoid activation function is utilized to calculate the attention weight (\mathbf{W}). $y_{2,i}$ is the reweighted multivariate output.

BiLSTM is a variant of Recurrent Neural Networks (RNNs) that learns long-timestep information bidirectionally and avoids the gradient vanishing or explosion of traditional RNNs (Kavianpour et al., 2023). Previous studies have indicated that
 215 BiLSTM outperforms LSTM in regression tasks due to the insufficient utilization of future information in LSTM (Siami-Namini et al., 2019; Yang and Wang, 2022). Therefore, the BiLSTM layer is integrated into the deep-learning module to

fully capture the temporal characteristics of the CNN attention-weighted multivariate output. The BiLSTM layer is realized by the forward LSTM and backward LSTM (Eq. (4)). Both the forward and backward LSTM consist of cell states, forget gates, input gates, output gates, and activation functions, which are responsible for transmission, screening and processing of temporal information. The final LSTM output is obtained by output gates and cell states (Eq. (5)). A detailed description of BiLSTM components can be found in our previous work (Li et al., 2025a).

$$H_t = [\vec{h}_t; \overleftarrow{h}_{end-t+1}], \quad (4)$$

$$h_t = o_t \times \tanh(C_t), \quad (5)$$

Where H_t is the final output of BiLSTM at timestep t , which is obtained by concatenating the forward output \vec{h}_t and backward output value $\overleftarrow{h}_{end-t+1}$. h_t is the final output of LSTM at timestep t , o_t is the output of output gate at timestep t , \tanh is an activation function that regulates the values transmitted in neural networks by compressing the values to a range of from -1 to 1. C_t is the output of the cell state at timestep t .

Hyperparameter tuning is crucial for improving the performance of deep neural networks. Bayesian optimization can determine global optima with higher efficiency (Shahriari et al., 2016) and has been widely employed in hyperparameter optimization of varying machine learning models (Wu et al., 2019a). The primary process of Bayesian optimization involves establishing search spaces for hyperparameters and the corresponding objective function, followed by the determination of the optimal solution by minimizing the objective function (Eq. (6)). Bayesian optimization utilizes a probabilistic surrogate model to iteratively estimate the complex unknown objective function based on the current query point and then identifies the next most promising query point by an acquisition function (Shahriari et al., 2016). The probabilistic surrogate model and the acquisition function in this study are the Gaussian process regression model (Rasmussen, 2004) and the Expected-Improvement-Per-Second-Plus function (Gelbart et al., 2014), respectively. The number of optimization iteration is set to 30 and the final optimal settings of model hyperparameters are presented in Table S2 of the supplement.

$$x^* = \operatorname{argmin} f(x), x \in X \subseteq R^d, \quad (6)$$

Where x^* is the optimal scheme of multiple hyperparameters, x is the decision vector composed of d -dimensional hyperparameters, X is the search space that consists of all possible decision vectors, $f(x)$ is the unknown objective function.

2.2.3 Physics-constrained optimization scheme

The normalized vertical profiles of $PM_{2.5}$ chemical components generated by the deep learning module are denormalized by the statistical characteristics of the initial input data of the surface-level observations. To reduce the retrieval error induced by the inherent extrapolation limitations of deep learning modules, a physics-constrained optimization scheme is incorporated into the retrieval framework based on a revised Interagency Monitoring of Projected Visual Environment

(IMPROVE) Equation (Pitchford et al., 2007) and Non-dominated Sorting Genetic Algorithm II (NSGA-II) (Verma et al., 2021).

250 The revised IMPROVE Equation interprets the particle extinction coefficient (σ) through the concentrations (M) and the optical and microphysical characteristics of PM_{2.5} chemical components (Eq. (7)).

$$\sigma(M) = \theta_s^{SNA} f(RH)(M(SO_4^{2-}) + M(NO_3^-) + M(NH_4^+)) + \theta_s^{OC} M(OC) + \theta_s^{FS} M(Fine\ Soil) + \theta_s^{CM} M(Coarse\ Mass) + \theta_s^{FSS} f_{FSS}(RH)M(Fine\ Sea\ Salt) + \theta_a^{BC} M(BC) + Rayleigh\ Scattering, \quad (7)$$

Where $\sigma(M)$ is the estimated particle extinction coefficient (km⁻¹), θ_s is the scattering efficiency (m² mg⁻¹), θ_a is the mass absorption efficiency (m² mg⁻¹), respectively. $f(RH)$ and $f_{FSS}(RH)$ account for the increase in light scattering induced by hygroscopic growth of sulfate, nitrate and ammonium (SNA), as well as fine sea salt (FSS). θ_s^{FS} , θ_s^{CM} , $\theta_s^{FSS} f_{FSS}$ and θ_a^{BC} are set to 0.001 m² mg⁻¹, 0.0006 m² mg⁻¹, 0.0017 m² mg⁻¹ and 0.01 m² mg⁻¹, respectively. M are the mass concentrations (μg m⁻³) of the PM_{2.5} chemical components. *Rayleigh Scattering* is set to 0.01 km⁻¹. θ_s^{SNA} and θ_s^{OC} are determined by Eq. (8)-(9).

$$\theta_s^{SNA} = 0.003 \times (0.7 + 0.002 \times (M(SO_4^{2-}) + M(NO_3^-) + M(NH_4^+) + M(OC))), \quad (8)$$

$$260 \quad \theta_s^{OC} = 0.00363 \times (0.7 + 0.002 \times (M(SO_4^{2-}) + M(NO_3^-) + M(NH_4^+) + M(OC))), \quad (9)$$

To implement the physics-constrained optimization, we first introduce a scale factor ($\gamma_{i,h}$) for each chemical component at each vertical layer, which is used to correct the initial mass concentrations (Eq. (10)). Then we determine the optimal scale factors through minimizing a multi-objective function (Eq. (11)). The Pearson correlation coefficient (CORR) and root mean square error (RMSE) quantified by the lidar-observed and the IMPROVE-simulated extinction coefficient serve as two objective values in the multi-objective function. The NSGA-II algorithm is utilized to determine the optimal scale factors by solving the multi-objective function that simultaneously enhances the correlation and reduces the discrepancy between the IMPROVE-estimated and lidar-observed extinction coefficients.

$$M_{regulated}^{i,h} = \gamma_{i,h} \times M_{original}^{i,h}, \quad i = SO_4^{2-}, NO_3^-, NH_4^+, OM, \text{ and } BC, \quad (10)$$

$$270 \quad \gamma_{i,h} = \min(f_{RMSE}(\gamma), f_{CORR}(\gamma)), \quad (11)$$

Where $M_{regulated}^{i,h}$ (μg m⁻³) is the regulated mass concentration of the i^{th} chemical component at an altitude of h (m), $\gamma_{i,h}$ is the scale factor for the i^{th} chemical component at an altitude of h (m), and $M_{original}^{i,h}$ (μg m⁻³) is the original mass concentration of the i^{th} chemical component at an altitude of h (m). $f_{RMSE}(\gamma)$ is the RMSE-based objective function (Eq. (12)) and $f_{CORR}(\gamma)$ is the CORR-based objective function (Eq. (13)).

$$275 \quad f_{RMSE}(\gamma) = \sqrt{\frac{\sum_{k=1}^K (\sigma_k^{obs} - \sigma_k(\gamma \times M))^2}{K}}, \quad (12)$$

$$f_{CORR}(\gamma) = -\frac{\sum_{k=1}^K \left(\frac{\sigma_k(\gamma \times M) - \bar{\sigma}(\gamma \times M)}{std(\sigma(\gamma \times M))} \right) \left(\frac{\sigma_k^{obs} - \bar{\sigma}^{obs}}{std(\sigma^{obs})} \right)}{K-1}, \quad (13)$$

Where K is the total number of samples, σ_k^{ob} is the k^{th} observed extinction coefficient, $\sigma_k(\gamma \times M)$ is the k^{th} simulated extinction coefficient, $\overline{\sigma(\gamma \times M)}$ is the average of simulated extinction coefficient, $\overline{\sigma^{obs}}$ is the average of observed extinction coefficient, $std(\sigma(\gamma \times M))$ is the standard deviation of simulated extinction coefficient, and $std(\sigma^{obs})$ is the standard deviation of observed extinction coefficient.

NSGA is capable of simultaneously optimizing the multi-objective function by generating a Pareto front that consists of an ensemble of non-dominated solutions (Srinivas and Deb, 1994). The non-dominated solutions in a Pareto front meet the criterion that one objective cannot be further improved without compromising other objectives. However, the initial version of NSGA has several limitations. First, NSGA has a high computational complexity of $O(MN^3)$, where M is the number of objective functions, and N is the size of the population. Second, NSGA utilizes a sharing parameter to preserve the diversity of the population that dominates the choice of Pareto non-dominated solutions, resulting in the introduction of parameter uncertainty into the algorithm. Third, NSGA lacks an elitism mechanism, leading to the incorrect removal of advantageous solutions. NSGA-II is an improved NSGA with a lower computational complexity of $O(MN^2)$ and an elitism mechanism that retains the dominant members of the parent and offspring generations during iterative evolution (Deb et al., 2002). Moreover, NSGA-II replaces the sharing parameters in NSGA with the crowding distance operator, mitigating the uncertainty of sharing parameters and the high computational complexity of sharing functions.

NSGA-II implements multi-objective optimization by two primary procedures, namely non-dominated sorting and crowding distance calculation. The non-dominated sorting progressively identifies the Pareto front at each rank from a population of size N . The Pareto front at the second rank is derived from a population that excludes the Pareto front at the first rank. The crowding distance is utilized to quantify the priority of all optimal solutions within a Pareto front, defined as the normalized distance of two nearest optimal solutions on either side (Eq. (14)).

$$d_i = \sum_{i=1}^K \frac{f_m^{i+1} - f_m^{i-1}}{f_m^{max} - f_m^{min}}, \quad (14)$$

Where d_i is the crowding distance of the i^{th} intermediate Pareto optimal solution, K is the number of Pareto optimal solutions in a Pareto front, f_m^{i+1} is the m^{th} objective value induced by the $(i + 1)^{\text{th}}$ Pareto optimal solution, f_m^{i-1} is the m^{th} objective value induced by the $(i - 1)^{\text{th}}$ Pareto optimal solution, f_m^{max} and f_m^{min} are the m^{th} maximum and minimum objective values, respectively.

The workflow of NSGA-II is summarized as follows (Fig. 2).

- a) Randomly generating an initial population (A_1) of size N . Performing selection, crossover and mutation operations on A_1 to generate an offspring population (B_1) of size N . The parent population (A_1) and the offspring population (B_1) are combined to form a new population (C_1) of size $2N$.
- b) Performing a rapid non-dominated sorting on C_1 to generate the Pareto fronts ($P_i, i = 1, 2, \dots, n$) at different ranks.

- 310 c) Filling the next population (A_2) of size N with P_i based on the rank order.
d) When A_2 is filled to the point of insufficient capacity to contain the entire P_i , the optimal solutions in P_i are inserted into A_2 in a priority order identified by the non-dominated sorting and crowding distance until the size of A_2 reaches N .
e) Performing selection, crossover and mutation operations on A_2 to generate an offspring population (B_2) of size N . The parent population (A_2) and offspring population (B_2) combine to form a new population (C_2) of size $2N$.
315 f) Iterating steps (b) to (e) until the convergence criteria are satisfied.

Figure 2: Brief workflow of NSGA-II (A: the parent population; B: the offspring population; C: the new population; P: the Pareto front).

2.2.4 Framework training and evaluation

- 320 An hourly multivariate dataset with extensive temporal coverage was employed to train and evaluate the deep learning module. To maintain temporal independence, the training (and validation) set was constructed from a 1-year (2021) time-series dataset obtained from a Beijing site (Fig. S1), while the testing set contains an independent 6-month (Jan 1-Mar 31 and Jun 1 to Aug 31, 2022) time-series dataset obtained from the same site. A 10-fold time-series cross-validation (CV) scheme was designed for the training (and validation) set to preserve its temporal order and prevent future information leakage,
325 which is detailed in Text S3 and Fig. S2 of the supplement. The iteration number of Bayesian optimization is set to 20.

- To fully evaluate the performance of the retrieval framework in predicting vertical profiles of NH_4^+ , SO_4^{2-} , NO_3^- , OM and BC, we conduct three retrieval experiments: (1) We compare the retrieved mass concentrations with the observed values at the surface level during a training year (2021) and three non-training years (2017, 2018 and 2024) to validate the temporal
330 generalization in all seasons and under diverse meteorological conditions. (2) We assess the spatial generalization ability by applying the retrieval framework to 23 non-training lidar sites in the NCP from February 8-15th, 2021 and comparing the retrieved mass concentrations with observations at the surface level. The spatial distribution of the 23 non-training lidar sites is presented in Fig. S1. (3) We validate the retrieved vertical profiles by aircraft-based and tower-based vertical observations during several non-training episodes. Subsequently, SHapley Additive exPlanations (SHAP), a local explainable technology
335 (Lundberg et al., 2020), has been widely employed in prediction interpretation for varying machine learning models (Li et al., 2025b; Hou et al., 2022), is integrated into the deep learning module to quantify the impact of multivariate input features on the retrieval of $\text{PM}_{2.5}$ chemical components. Finally, we applied this retrieval framework to generate a long-term vertical profile dataset for five $\text{PM}_{2.5}$ chemical components in a megacity over six years of 2017-2018 and 2021-2024.

3 Results and discussion

340 3.1 Validation

3.1.1 Evaluation of the deep learning module performance

The 10-fold CV sets and a testing set with temporal independence are utilized to evaluate the predictive performance of the deep learning module, which is quantified by the discrepancies between simulations and observations at ground level for NH_4^+ , SO_4^{2-} , NO_3^- , OM and BC. Overall, the scatter distribution and fitted regression line closely align with the 1:1 line in
345 both the 10-fold CV (Fig. 3a1-a5) and temporally independent testing phases (Fig. 3b1-b5). The error distributions are concentrated around 0, with mean errors between $-1.78 \pm 8.15 \mu\text{g m}^{-3}$ and $-0.13 \pm 0.94 \mu\text{g m}^{-3}$ during the 10-fold CV phase (Fig. S3a1-a5) and between $-1.36 \pm 7.40 \mu\text{g m}^{-3}$ and $-0.07 \pm 1.00 \mu\text{g m}^{-3}$ during the temporally independent testing phase (Fig. S3b1-b5), demonstrating strong consistency between observations and simulations. Notably, the error distributions for the validation and independent testing sets are closely aligned, indicating that the deep learning module is robust and generalizes
350 well to unseen data. Specifically for the 10-fold CV process (Fig. 3a1-a5), the CORR values for the five $\text{PM}_{2.5}$ chemical components range from 0.76 to 0.86, indicating that the deep learning module accurately interprets the relationship between multivariate input features and the five $\text{PM}_{2.5}$ chemical components. The RMSE values range from 0.95 to $8.35 \mu\text{g m}^{-3}$, indicating a low discrepancy between simulations and observations. Compared to the 10-fold CV process, the temporally independent testing yields slightly lower CORR values (0.69-0.79) and higher RMSE values (1.00 - $8.87 \mu\text{g m}^{-3}$), showing a
355 slight underestimation for the five $\text{PM}_{2.5}$ chemical components (Fig. 3b1-b5). It is expected that the statistical results from the temporally independent testing are less robust than those from the 10-fold CV, since the temporally independent testing set aggregates a broader spectrum of temporal patterns compared to the validation set at each fold. Our statistical results from the 10-fold CV exhibit similarities or even improvements compared to those reported in other studies that predicting $\text{PM}_{2.5}$ chemical component concentrations based on machine learning models (Lv et al., 2021; Lin et al., 2022; Araki et al.,
360 2022; Liu et al., 2023), indicating that the deep learning module demonstrates strong prediction capabilities.

Figure 3: Scatterplots of the simulations ($\mu\text{g m}^{-3}$) versus the observations ($\mu\text{g m}^{-3}$) with probability density (%) for NH_4^+ , NO_3^- , SO_4^{2-} , OM and BC during the 10-fold cross-validation process (a1-a5) and temporally independent testing process (b1-b5). The dotted grey lines represent the 2:1, 1:1, and 1:2 lines, and the solid red line represents the fitted regression line. CORR represents the correlation coefficient, and RMSE represents root mean square error.
365

3.1.2 Comparison with ground-level observations

The retrieval framework was applied to retrieve the vertical profiles of NH_4^+ , NO_3^- , SO_4^{2-} , OM and BC in a Beijing lidar site (39.98°N , 116.38°E) over a training year (2021) and three non-training years (2017, 2018 and 2024). As illustrated in Fig. 4a1-a5, the weekly-smoothed variations in the retrieved surface concentrations of the five $\text{PM}_{2.5}$ chemical components
370 demonstrate strong consistency with the observed surface concentrations for the training year, indicating that the retrieval framework adequately captures the temporal characteristics of these chemical components. The CORR values between the

retrieved and observed concentrations range from 0.87 to 0.97, surpassing those of the deep learning module (Fig. 4a1-a5 and Fig. 3b1-b5). In addition, the RMSE values for all five chemical components ($0.57\text{--}4.98 \mu\text{g m}^{-3}$) are consistently lower than those from the deep-learning module (Fig. 4a1-a5 and Fig. 3b1-b5). These results demonstrate that the physics-
375 constrained optimization effectively enhances the retrieval accuracy of chemical component concentrations.

For the non-training years, the retrieved surface concentrations of a sum of five $\text{PM}_{2.5}$ chemical components are compared to the observed surface $\text{PM}_{2.5}$ concentrations, owing to the absence of long-term observations for individual chemical components. As shown in Fig. 4b-d, the weekly-smoothed variations in the retrieved surface $\text{PM}_{2.5}$ concentrations closely
380 align with the observed values in 2017, 2018 and 2024. The high values of surface $\text{PM}_{2.5}$ concentration observed in March-April and November of 2018 and 2024 are effectively captured by the retrieval framework. These results indicate that the retrieval framework roughly interprets the changes in concentrations of various chemical components across different periods, exhibiting fundamental temporal generalization capabilities. However, the retrieved concentrations show some overestimation cases during autumn in 2018 and spring in 2024, potentially associated with the uncertainties induced by the
385 training data. The training data may lack a sufficiently diverse spectrum of meteorological conditions and pollution patterns, which limits the temporal generalizability of the retrieval framework across all complex and dynamic atmospheric scenarios. Future efforts should enhance retrieval accuracy by augmenting the training data with observations spanning a wider range of temporal conditions.

390 **Figure 4: Weekly-smoothed variations in the retrieved and observed concentrations ($\mu\text{g m}^{-3}$) of NH_4^+ (a1), NO_3^- (a2), SO_4^{2-} (a3), OM (a4) and BC (a5) in 2021. (b) same as (a1-a5) but for $\text{PM}_{2.5}$ in 2017. (c) same as (a1-a5) but for $\text{PM}_{2.5}$ in 2018. (d) same as (a1-a5) but for $\text{PM}_{2.5}$ in 2024. CORR represents the correlation coefficient, RMSE represents root mean square error.**

The retrieval framework was also applied to retrieve the vertical profiles of the five $\text{PM}_{2.5}$ chemical components at 23 non-
395 training NCP lidar sites over a short-term period of February 8-15th, 2021, aiming to validate its spatial generalization capabilities. Compared with the observed surface concentrations at 23 non-training sites, the retrieved surface concentrations exhibit a more clustered data distribution and exhibit a tendency toward underestimation across all components (Fig. 5a). The site-averaged CORR values for the five chemical components range from 0.21 to 0.46, with RMSE values spanning $2.7 \mu\text{g m}^{-3}$ to $20.37 \mu\text{g m}^{-3}$ (Fig. S4). From a spatial perspective (Fig. 5b1-b5), non-training NCP sites located closer to the
400 Beijing lidar site exhibit higher CORR values, with the highest reaching 0.71 (NH_4^+), 0.56 (NO_3^-), 0.81 (SO_4^{2-}), 0.48 (OM) and 0.41 (BC). Conversely, the RMSE values are not affected by the distance from the Beijing lidar site (Fig. 5c1-c5), with the lowest reaching $2.91 \mu\text{g m}^{-3}$ (NH_4^+), $6.15 \mu\text{g m}^{-3}$ (NO_3^-), $3.05 \mu\text{g m}^{-3}$ (SO_4^{2-}), $6.59 \mu\text{g m}^{-3}$ (OM) and $0.78 \mu\text{g m}^{-3}$ (BC). However, several sites exhibit poor retrieval performance, with CORR values ranging from ~ 0.20 to ~ 0.30 (Fig. S5), which is primarily attributed to limitations in the spatial representativeness of the training data. The deep-learning module was
405 trained exclusively on a long-term dataset from a single site in Beijing, which is insufficient to capture the spatial variability

in emission intensity, as well as local meteorological and geographical conditions across the broader NCP. As a result, the spatial extrapolation capability of the deep-learning module is constrained. Although the retrieval framework can retrieve $PM_{2.5}$ chemical component concentrations at spatially distributed lidar sites, future work should incorporate long-term datasets from varying locations to enhance spatial generalization and extrapolation performance.

410

Figure 5: Data distribution properties of retrieved and observed surface mass concentration ($\mu\text{g m}^{-3}$) of NH_4^+ , NO_3^- , SO_4^{2-} , OM and BC at 23 non-training NCP lidar sites over a period of February 8-15th, 2021, presented by a combination of boxplots and kernel density (a). Spatial distribution of Pearson correlation coefficient (CORR) between retrieved and observed surface mass concentration of NH_4^+ (b1), NO_3^- (b2), SO_4^{2-} (b3), OM (b4) and BC (b5). (c1-c5) Same as (b1-b5) but for root mean square error (RMSE, $\mu\text{g m}^{-3}$). The geographic basemap is hosted by Esri (<https://www.esri.com/en-us/home>).

415

3.1.3 Verification of retrieved vertical profiles

In addition to the spatiotemporal verification of surface-level mass concentrations, tower-based and aircraft-based observational experiments were conducted to validate the retrieved vertical profiles of five $PM_{2.5}$ chemical components during non-training periods. From the surface to ~ 200 m altitude, the retrieved and observed vertical profiles exhibit similar vertical patterns during a period from December 30, 2018 to January 2, 2019 in Beijing, with higher concentrations occurring at altitudes of 50-80 m for NH_4^+ , NO_3^- , SO_4^{2-} and OM (Fig. 6a1, a2). Specifically, as presented in Table S3, the CORR values are no less than 0.66 for all four $PM_{2.5}$ chemical components. However, the RMSE value for OM ($23.04 \mu\text{g m}^{-3}$) is notably higher than that for the other components (4.08 - $10.48 \mu\text{g m}^{-3}$), indicating limitations in the retrieval framework when representing the vertical profile of OM during winter pollution episodes. This discrepancy may be associated with retrieval uncertainties arising from input data quality and imposed physical constraints. Additionally, the retrieved and observed proportions of NH_4^+ , NO_3^- , SO_4^{2-} , OM and BC demonstrate significant consistency (Fig. 6b1, b2). Among these chemical components, NO_3^- and OM contribute the largest proportions, followed by NH_4^+ and SO_4^{2-} , while BC contributes the smallest fraction. This proportional characteristic is evident in both the retrieved and observed proportions at altitudes of 600 m and 1200 m (Fig. 6c1, c2). Due to the lack of NH_4^+ measurements at 1500 m and the absence of both NH_4^+ and SO_4^{2-} measurements at 2100 m, the proportions at these altitudes are statistically inferred from the remaining chemical components. The results indicate overall consistency between retrieved and observed proportions at altitudes of 1500 m and 2100 m, although the proportion of NO_3^- is slightly overestimated at 2100 m and underestimated at 1500 m. Overall, the tower-based and aircraft-based verifications indicate that the retrieval framework achieves high accuracy in retrieving the vertical profiles of the five $PM_{2.5}$ chemical components during non-training period, demonstrating its robust generalization capability and reliability when applied to independent datasets.

425
430
435

Figure 6: Vertical profiles ($\mu\text{g m}^{-3}$) of NH_4^+ , NO_3^- , SO_4^{2-} , and OM from retrieval (a1) and tower-based observation (a2) during a period from December 30, 2018 to January 2, 2019 in Beijing. The line represents the daily average of the hourly vertical profiles, and the shaded area represents the standard deviation. Averaged proportions of NH_4^+ , NO_3^- , SO_4^{2-} , OM, and BC from retrieval (b1) and tower-based observation (b2) for 10 days (December 27 and 30, 2023; January 2, 5, 9, 12, 15, 18, 24, and 27, 2024). (c1 and c2) Same as (b1 and b2) but for aircraft-based verification for 3 days (September 26, October 10, December 11, 2024).

440

3.2 Assessment of feature importance

The predictive performance of the deep learning module is intricately connected to the input features (Blum and Langley, 1997). Although the module incorporates the CNN and attention mechanism layer to mitigate issues related to feature dimension, the impact of input features on the module predictions remains ambiguous, which impedes module interpretability and restricts the capacity to enhance the module performance through effective feature selection. The SHAP method is employed to quantify the relative contributions of 8 input features to the predictions of the five PM_{2.5} chemical components at various heights and to identify the impact of the input features on the decision-making processes of the deep learning module. The coexistence of a high feature value with a positive SHAP value in a specific feature implies an amplification of concentration prediction at elevated levels.

Figure 7a1-a5 depicts that the aerosol extinction coefficient at 532 nm (EXT), relative humidity (RH) and v-component wind (VW) are the significant input features for predicting the five PM_{2.5} chemical components with an averaged relative contribution of 14.43 %, 15.84 % and 16.77 %. These features largely affect the vertical structure, chemical and physical processes, respectively. Specifically, EXT characterizes the vertical distribution of a total of the five PM_{2.5} chemical components and plays a crucial indicative role in vertical profile predictions (Tao et al., 2016). RH is a key driving factor in aerosol hygroscopic growth, aqueous-phase chemical reactions, and heterogeneous reactions, significantly contributing to the mass concentrations of varying chemical components as reported in numerous studies (Fang et al., 2019; Wang et al., 2020; Gao et al., 2020; Liang et al., 2019). VW primarily affects latitudinal transboundary transport, which is a dynamic forcing in the southwest-northeast transport channel of the Beijing-Tianjin-Hebei (BTH) region (Yang et al., 2024). Notably, the relative contribution of EXT decreases with height from the surface (50 m) to the free atmosphere (1900 m), while the relative contribution of VW exhibits an opposite trend. The aerosol content in the upper planetary boundary layer is relatively low, and the weakened lidar aerosol signal is susceptible to interference from noise signals, restricting the indicative effect of EXT on chemical component concentrations. Conversely, pollution transport in the upper planetary boundary layer is less affected by interference from complex underlying surfaces than near-surface transport (Wu et al., 2019b), amplifying the driving effect of high-altitude VW on chemical component concentrations. Specific humidity (SH) and geopotential (GEOP) also provided important contributions (13.04% and 12.85%, respectively). SH is related to the vertical diffusion and wet scavenging of pollutants (Chatfield et al., 2020) and GEOP identifies the synoptic meteorological patterns that affect both horizontal process (Jia et al., 2022; Wang et al., 2021) and vertical distribution of pollutants within the boundary layer (Miao et al., 2022; Xu et al., 2019).

Figure 7: Relative contribution of 8 input features on predictive NH₄⁺ (a1), NO₃⁻ (a2), SO₄²⁻ (a3), OM (a4) and BC (a5) at altitudes of 50 m, 766 m and 1900 m. SHAP values with feature values of 8 input features for predictive NH₄⁺ (b1), NO₃⁻ (b2), SO₄²⁻ (b3), OM (b4) and BC (b5) at an altitude of 50 m. (c1-c5) Same as (b1-b5) but for an altitude of 766 m. (d1-d5) Same as (b1-b5) but for an altitude of 1900 m. F1: extinction coefficient at 532 nm, EXT; F2: Geopotential, GEOP; F3: Relative humidity, RH; F4: Specific humidity, SH; F5: Temperature, TEMP; F6: U-component wind, UW; F7: V-component wind; F8: Vertical velocity, VV.

Figure 7b1-d5 further determines the impact of the input features on the decision-making processes of the deep learning module. From Fig. 7b1-b5, the elevated levels of EXT, GEOP, and VW significantly enhance the concentration predictions of the five PM_{2.5} chemical components in the near-surface layer (50 m), while high-level RH exert either positive or negative effects on predictions. High RH not only facilitates aqueous-phase and heterogeneous chemical reactions, positively contributing to predictions, but also promotes aerosol coalescence, leading to dry and wet deposition that negatively contributes to predictions (Chen et al., 2020). The results in the middle of the boundary layer (766 m) are consistent with those observed in the near-surface layer (Fig. 7c1-c5). Particularly, the positive driving effect of lower VV values on predictions is more significant, with downward wind contributing positively to predictions, which is attributed to the fact that sinking airflows inhibit the dispersion of chemical components, thereby exacerbating aggregation and increasing concentration (Yang et al., 2022). The results in the free atmosphere (1900 m) align with those in the middle of the boundary layer (Fig. 7d1-d5). Notably, the influence of UW on predictions is more apparent, as the westerly wind positively contributes to the predictions, which is primarily due to the elevated emission sources located in the southwestern BTH region (Yang et al., 2024). Strong prevailing southwesterly winds at high altitudes enhance the regional transport of atmospheric pollutants, leading to an increase in concentration.

3.3 Application of the retrieval framework

The retrieval framework was applied to generate a long-term dataset of vertical profiles for NH₄⁺, NO₃⁻, SO₄²⁻, OM and BC over six years (2017-2018, 2021-2024) at a Beijing lidar site. Figure 8 shows the averaged vertical profiles for the five PM_{2.5} chemical components in spring (MAM) (Fig. 8a1), summer (JJA) (Fig. 8a2), autumn (SON) (Fig. 8a3) and winter (DJF) (Fig. 8a4) during the six years. OM mass concentrations are consistently the highest across all four seasons, followed by NO₃⁻, while the mass concentrations of NH₄⁺, SO₄²⁻ and BC remain relatively low. The high proportions of OM and NO₃⁻ in Chinese PM_{2.5} pollution were frequently reported in recent studies (Zhang et al., 2024; Liu et al., 2022). Since the implementation of the Air Pollution Prevention and Control Action Plan during 2013-2017 and the Three-year Action Plan to Win the Blue-Sky Defense War during 2018-2020 in China, effective reductions in sulfur dioxide (SO₂) have gradually shifted the dominated chemical component of PM_{2.5} pollution from SO₄²⁻ to OM and NO₃⁻ (Niu et al., 2022). Furthermore, the decreased SO₄²⁻ mass concentrations have amplified the competitive effect of NO₃⁻ on capturing NH₃ and NH₄⁺ in the thermodynamic equilibrium process, increasing NO₃⁻ mass concentrations (Geng et al., 2024). In comparison to the mass concentrations of the five PM_{2.5} chemical components in MAM, SON and DJF, summertime mass concentrations are notably lower, which are attributed to reduced heating activities and enhanced wet deposition during summer periods (Liu et al., 2015; Ji et al., 2019). Moreover, the summertime vertical distributions of the five chemical components are relatively uniform, which may be attributed to the enhanced atmospheric vertical mixing effects induced by the unstable boundary layer (Roostaei et al., 2024).

510 **Figure 8: Vertical distribution of mass concentrations ($\mu\text{g m}^{-3}$) for NH_4^+ , NO_3^- , SO_4^{2-} , OM and BC in spring (MAM, a1), summer (JJA, a2), autumn (SON, a3) and winter (DJF, a4) over six years (2017-2018, 2021-2024). Averaged vertical profiles of mass concentrations ($\mu\text{g m}^{-3}$) for NH_4^+ , NO_3^- , SO_4^{2-} , OM and BC from 2017 to 2018 (b1), from 2021 to 2022 (b2), and from 2023 to 2024 (b3). Annual change rates ($\mu\text{g m}^{-3} \text{a}^{-1}$) of mass concentrations for NH_4^+ , NO_3^- , SO_4^{2-} , OM and BC at various altitudes from 2021 to 2024 (c).**

515

Figure 8 also shows the averaged vertical profiles for 2017-2018 (Fig. 8b1), 2021-2022 (Fig. 8b2) and 2023-2024 (Fig. 8b3), as well as the annual change rate during 2021-2024 in Beijing (Fig. 8c). During 2017-2018, the implementation of clean air policies resulted in mass concentrations of NH_4^+ , SO_4^{2-} and BC remaining below $8 \mu\text{g m}^{-3}$ (Fig. 8b1). However, the mass concentrations of NO_3^- exceeded $13 \mu\text{g m}^{-3}$ at altitudes below ~ 125 m, and those of OM exceeded $15 \mu\text{g m}^{-3}$ below ~ 100 m, likely due to the nonlinear response to emission reduction (Li et al., 2021). Compared to 2017-2018, the mass concentrations of NH_4^+ , NO_3^- , SO_4^{2-} , OM and BC decreased significantly during 2021-2022 with average reductions of 8.36 %, 10.65 %, 6.58 %, 8.58 % and 5.85 %, respectively, from 50 m to 3000 m (Fig. 8b2). These decreases are attributed to the continued implementation of clean air policies and reduced emissions associated with the COVID-19 pandemic control in China (Kang et al., 2020). During 2023-2024, the mass concentrations of NH_4^+ , NO_3^- , SO_4^{2-} , OM and BC increased relative to the 2021-2022 levels, with average increases of 5.49 %, 6.43 %, 4.65 %, 5.75 % and 4.40 %, respectively, from 50 m to 3000 m (Fig. 8b3). This rebound over the NCP has been reported previously and is likely related to the offsetting effect of enhanced human activities following the relaxation of the COVID-19 pandemic lockdowns on the implementation of clean air policies (Song et al., 2025). From 2021 to 2024, the change rates of NH_4^+ , NO_3^- and SO_4^{2-} at approximately 50 m, 310 m and 770 m exhibit decreasing trends, with decrease rates ranging from 0.06 to $0.19 \mu\text{g m}^{-3} \text{a}^{-1}$. In contrast, the change rates of the five chemical components at approximately 80 m, 120 m, 1210 m and 1900 m show increasing trends, with the highest increase rate of $0.83 \mu\text{g m}^{-3} \text{a}^{-1}$ occurring at ~ 120 m for NO_3^- (Fig. 8c). In addition, OM exhibited a significant increase rate of $0.69 \mu\text{g m}^{-3} \text{a}^{-1}$ at ~ 120 m, which may be related to the low sensitivity of high-altitude organic aerosols to emission controls (Zhao et al., 2017). Future clean air policies should prioritize strengthening control measures for OM and NO_3^- within the lower and middle parts of the atmospheric boundary layer.

535 **3.4 Limitations and uncertainties**

The deep learning module in our retrieval framework can establish a powerful mapping between optical and meteorological features and $\text{PM}_{2.5}$ chemical species, and physics-based explicit constraints can enhance the reliability and expandability of the mapping relationships. However, several limitations and sources of uncertainty remain and should be acknowledged when interpreting the results and extending the framework to broader applications.

540

First, the spatial scope of the training data is predominantly restricted to the NCP region. Expanding the retrieval framework with data from more diverse geographical locations is necessary to improve its global transferability. Second, the current retrieval framework primarily relies on extinction coefficients at a wavelength of 532 nm, exhibiting dependence on specific

lidar instruments. Future retrieval framework should focus on integrating diverse optical features from additional
545 wavelengths to enhancing adaptability and transferability. Third, the auxiliary input data used in both the deep learning
module and the physics-constrained optimization are obtained from global reanalysis products, which may not fully capture
local atmospheric conditions at specific observational sites, thereby introducing representativeness errors into the retrievals.
Acquiring the vertical observational data for these auxiliary features can effectively mitigate the uncertainty induced by the
input data. Fourth, the IMPROVE equation applied as an external physical constraint may introduce additional uncertainty
550 into the retrievals due to its systematic estimation biases (Lowenthal and Kumar, 2016). Moreover, since the IMPROVE
equation was applied as an external physical constraint to optimize the retrievals of PM_{2.5} chemical components, the machine
learning model itself was not intrinsically constrained by physical principles during its training. Future work could
incorporate an internal physical constraint into the machine learning model to improve its physical interpretability by
formulating a hybrid loss function for training that combines the traditional data-fitting term with a physical term. Finally,
555 long-term acquisition of independent vertical profiling data from both tower-based and aircraft-based campaigns is essential
for a comprehensive assessment of the robustness of the vertical retrievals with respect to varying sites, aerosol types, and
seasons.

4 Conclusions

This study proposes a novel lidar-based retrieval framework for obtaining the vertical profiles of five PM_{2.5} chemical
560 components (NH₄⁺, SO₄²⁻, NO₃⁻, OM and BC) for the first time. A long-term multivariate dataset was utilized to train a
complex deep-learning module in the retrieval framework, thus interpreting the nonlinear relationship among lidar
parameters, meteorological parameters and PM_{2.5} chemical components. A physics-constrained optimization module was
integrated into the retrieval framework, enhancing the generalization capabilities of predicting vertical profiles across diverse
spatiotemporal scenarios.

565 In situ surface observations of hourly mass concentrations of PM_{2.5} and its five chemical components over a training year
and three non-training years were used to validate the accuracy of the retrieval framework in interpreting temporal variations.
The results showed that the Pearson correlation coefficient values between the retrieved and observed concentrations ranged
from 0.87 to 0.97 during the training year, and the variations in the retrieved surface PM_{2.5} mass concentrations closely
570 aligned with the observations during the non-training year, indicating the robust capabilities of temporal prediction and
generalization in the retrieval framework. The retrieval framework was then applied to obtain the mass concentrations of five
PM_{2.5} chemical components at 23 non-training sites. The retrieved results exhibited patterns that are moderately consistent
with the corresponding observations. However, limitations remained in accurately capturing short-term temporal variations,
with a general tendency toward underestimation. Tower-based and aircraft-based field campaigns at altitudes ranging from
575 surface to 2100 m were conducted to validate the accuracy of the retrieved vertical profiles of NH₄⁺, SO₄²⁻, NO₃⁻, OM and

BC. The tower-based and aircraft-based verifications indicate that the retrieved and observed vertical profiles of these components exhibited consistent patterns in mass concentrations and proportions, demonstrating the robust capabilities of the retrieval framework in obtaining high-precision vertical profiles from non-training datasets.

580 Subsequently, SHapley Additive exPlanations (SHAP), an explainable technology, is integrated into the deep learning module to quantify the impact of multivariate input features on the retrieval of PM_{2.5} chemical components. The results showed that the aerosol extinction coefficient at 532 nm, relative humidity and v-component wind are the dominant input features for predicting the five PM_{2.5} chemical components with an averaged relative contribution of 14.43 %, 15.84 % and 16.77 %. The driving effect of the input features on the decision-making processes of the deep learning module was also
585 determined by SHAP values.

Finally, we applied this framework to generate a long-term dataset of vertical profiles for NH₄⁺, SO₄²⁻, NO₃⁻, OM and BC over six years (2017-2018, 2021-2024). From this dataset, we found that OM mass concentrations are consistently the highest across all four seasons, followed by NO₃⁻, while the mass concentrations of NH₄⁺, SO₄²⁻ and BC remain relatively
590 low. From 2021 to 2024, the change rates of NH₄⁺, NO₃⁻ and SO₄²⁻ at approximately 50 m, 310 m and 770 m exhibit decreasing trends, with decrease rates ranging from 0.06 to 0.19 μg m⁻³ a⁻¹. However, OM and NO₃⁻ exhibited significant increase rates of 0.69 μg m⁻³ a⁻¹ and 0.83 μg m⁻³ a⁻¹, respectively, at an altitude of ~120 m. Future clean air policies should prioritize strengthening control measures for OM and NO₃⁻ within the lower and middle parts of the atmospheric boundary layer. Our new retrieval framework offers a novel approach to acquiring vertical profiles of PM_{2.5} chemical components.
595 Future efforts should aim to mitigate the overestimation of carbonaceous aerosols by regulating the parameters involved in the physics-constrained optimization process.

Code and data availability

The source codes and related data in this manuscript are freely available upon request through the corresponding author (tingyang@mail.iap.ac.cn).

600 Author contributions

HL developed the retrieval framework, carried out the analysis and verification, as well as wrote this paper. TY provided scientific guidance and wrote this paper. TY and YS provided various measurement data. ZW did overall supervision. All authors reviewed and revised this paper.

Competing interests

605 The authors declare that they have no conflict of interest.

Acknowledgements

Ting Yang would like to express gratitude towards the Program of the Youth Innovation Promotion Association (CAS). We thank for the technical support of the National large Scientific and Technological Infrastructure “Earth System Numerical Simulation Facility” (<https://cstr.cn/31134.02>. EL), and the data support of the China National Environmental Monitoring
610 Center.

Financial support

This work was supported by National Natural Science Foundation of China (NSFC) Excellent Young Scientists Fund (grant no. 42422506), National Key Research and Development Program of China (grant no. 2023YFC3705801), and the National Natural Science Foundation of China (grant no. 42275122).

615 References

- Al-Faiz, M. Z., Ibrahim, A. A., and Hadi, S. M.: The effect of Z-Score standardization (normalization) on binary input due the speed of learning in back-propagation neural network, *Iraqi j. inf. commun. technol.*, 1, 42-48, <https://doi.org/10.31987/ijict.1.3.41>, 2018.
- Alzubaidi, L., Zhang, J., Humaidi, A. J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., Santamaría, J., Fadhel, M. A., Al-Amidie,
620 M., and Farhan, L.: Review of deep learning: concepts, CNN architectures, challenges, applications, future directions, *J. Big Data*, 8, 53, <https://doi.org/10.1186/s40537-021-00444-8>, 2021.
- Ansmann, A., Bösenberg, J., Chaikovsky, A., Comerón, A., Eckhardt, S., Eixmann, R., Freudenthaler, V., Ginoux, P., Komguem, L., Linné, H., Márquez, M. Á. L., Matthias, V., Mattis, I., Mitev, V., Müller, D., Music, S., Nickovic, S., Pelon, J., Sauvage, L., Sobolewsky, P., Srivastava, M. K., Stohl, A., Torres, O., Vaughan, G., Wandinger, U., and Wiegner, M.: Long-
625 range transport of Saharan dust to northern Europe: The 11–16 October 2001 outbreak observed with EARLINET, *J. Geophys. Res.: Atmos.*, 108, 4783, <https://doi.org/10.1029/2003JD003757>, 2003.
- Araki, S., Shimadera, H., and Shima, M.: Continuous estimations of daily PM_{2.5} chemical components from temporally sparse monitoring data using a machine learning approach, *Atmos. Pollut. Res.*, 13, 101580, <https://doi.org/10.1016/j.apr.2022.101580>, 2022.
- 630 Blum, A. L. and Langley, P.: Selection of relevant features and examples in machine learning, *Artif. Intel.*, 97, 245-271, [https://doi.org/10.1016/S0004-3702\(97\)00063-5](https://doi.org/10.1016/S0004-3702(97)00063-5), 1997.

- Cabello-Solorzano, K., Ortigosa de Araujo, I., Peña, M., Correia, L., and J. Tallón-Ballesteros, A.: The Impact of Data Normalization on the Accuracy of Machine Learning Algorithms: A Comparative Analysis, 18th International Conference on Soft Computing Models in Industrial and Environmental Applications (SOCO 2023), Cham, 344-353, https://doi.org/10.1007/978-3-031-42536-3_33, 2023.
- 635
- Chatfield, R. B., Sorek-Hamer, M., Esswein, R. F., and Lyapustin, A.: Satellite mapping of PM_{2.5} episodes in the wintertime San Joaquin Valley: a “static” model using column water vapor, *Atmos. Chem. Phys.*, 20, 4379-4397, <https://doi.org/10.5194/acp-20-4379-2020>, 2020.
- Chen, Z., Chen, D., Zhao, C., Kwan, M.-p., Cai, J., Zhuang, Y., Zhao, B., Wang, X., Chen, B., Yang, J., Li, R., He, B., Gao, B., Wang, K., and Xu, B.: Influence of meteorological conditions on PM_{2.5} concentrations across China: A review of methodology and mechanism, *Environ. Int.*, 139, <https://doi.org/10.1016/j.envint.2020.105558>, 2020.
- 640
- Deb, K., Pratap, A., Agarwal, S., and Meyarivan, T.: A fast and elitist multiobjective genetic algorithm: NSGA-II, *IEEE Trans. Evol. Comput.*, 6, 182-197, <https://doi.org/10.1109/4235.996017>, 2002.
- Dubey, R., Patra, A. K., and Nazneen: Vertical profile of particulate matter: A review of techniques and methods, *Air Qual. Atmos. Hlth.*, 15, 979-1010, <https://doi.org/10.1007/s11869-022-01192-1>, 2022.
- 645
- Fang, Y., Ye, C., Wang, J., Wu, Y., Hu, M., Lin, W., Xu, F., and Zhu, T.: Relative humidity and O₃ concentration as two prerequisites for sulfate formation, *Atmos. Chem. Phys.*, 19, 12295-12307, <https://doi.org/10.5194/acp-19-12295-2019>, 2019.
- Friedman, J. H., Bentley, J. L., and Finkel, R. A.: An algorithm for finding best matches in logarithmic expected time, *ACM T. Math. Software (TOMS)*, 3, 209-226, 1977.
- 650
- Gao, J., Wei, Y., Shi, G., Yu, H., Zhang, Z., Song, S., Wang, W., Liang, D., and Feng, Y.: Roles of RH, aerosol pH and sources in concentrations of secondary inorganic aerosols, during different pollution periods, *Atmos. Environ.*, 241, 117770, <https://doi.org/10.1016/j.atmosenv.2020.117770>, 2020.
- Gelbart, M. A., Snoek, J., and Adams, R. P.: Bayesian optimization with unknown constraints, arXiv preprint arXiv, 1403, 5607, <https://doi.org/10.48550/arXiv.1403.5607>, 2014.
- 655
- Geng, G., Liu, Y., Liu, Y., Liu, S., Cheng, J., Yan, L., Wu, N., Hu, H., Tong, D., Zheng, B., Yin, Z., He, K., and Zhang, Q.: Efficacy of China’s clean air actions to tackle PM_{2.5} pollution between 2013 and 2020, *Nat. Geosci.*, 17, 987-994, <https://doi.org/10.1038/s41561-024-01540-z>, 2024.
- Guo, M. H., Xu, T. X., Liu, J. J., Liu, Z. N., Jiang, P. T., Mu, T. J., Zhang, S. H., Martin, R. R., Cheng, M. M., and Hu, S. M.: Attention mechanisms in computer vision: A survey, *Comput. Vis. Media*, 8, 331-368, <https://doi.org/10.1007/s41095-022-0271-y>, 2022.
- 660
- Hara, Y., Nishizawa, T., Sugimoto, N., Osada, K., Yumimoto, K., Uno, I., Kudo, R., and Ishimoto, H.: Retrieval of Aerosol Components Using Multi-Wavelength Mie-Raman Lidar and Comparison with Ground Aerosol Sampling, *Remote Sens.*, 10, <https://doi.org/10.3390/rs10060937>, 2018.
- 665
- Hou, L., Dai, Q., Song, C., Liu, B., Guo, F., Dai, T., Li, L., Liu, B., Bi, X., Zhang, Y., and Feng, Y.: Revealing Drivers of

- Haze Pollution by Explainable Machine Learning, *Environ. Sci. Technol. Lett.*, 9, 112-119, <https://doi.org/10.1021/acs.estlett.1c00865>, 2022.
- Ji, W., Wang, Y., and Zhuang, D.: Spatial distribution differences in PM_{2.5} concentration between heating and non-heating seasons in Beijing, China, *Environ. Pollut.*, 248, 574-583, <https://doi.org/10.1016/j.envpol.2019.01.002>, 2019.
- 670 Jia, Z., Doherty, R. M., Ordóñez, C., Li, C., Wild, O., Jain, S., and Tang, X.: The impact of large-scale circulation on daily fine particulate matter (PM_{2.5}) over major populated regions of China in winter, *Atmos. Chem. Phys.*, 22, 6471-6487, <https://doi.org/10.5194/acp-22-6471-2022>, 2022.
- Kang, Y.-H., You, S., Bae, M., Kim, E., Son, K., Bae, C., Kim, Y., Kim, B.-U., Kim, H. C., and Kim, S.: The impacts of COVID-19, meteorology, and emission control policies on PM_{2.5} drops in Northeast Asia, *Sci. Rep.*, 10, 22112, <https://doi.org/10.1038/s41598-020-79088-2>, 2020.
- 675 Kavianpour, P., Kavianpour, M., Jahani, E., and Ramezani, A.: A CNN-BiLSTM model with attention mechanism for earthquake prediction, *J. Supercomput.*, 79, 19194-19226, <https://doi.org/10.1007/s11227-023-05369-y>, 2023.
- Kim, S., Yang, J., Park, J., Song, I., Kim, D.-G., Jeon, K., Kim, H., and Yi, S.-M.: Health effects of PM_{2.5} constituents and source contributions in major metropolitan cities, South Korea, *Environ. Sci. Pollut. Res.*, 29, 82873-82887, <https://doi.org/10.1007/s11356-022-21592-1>, 2022.
- 680 Lee, Y. S., Choi, E., Park, M., Jo, H., Park, M., Nam, E., Kim, D. G., Yi, S.-M., and Kim, J. Y.: Feature extraction and prediction of fine particulate matter (PM_{2.5}) chemical constituents using four machine learning models, *Expert Syst. Appl.*, 221, 119696, <https://doi.org/10.1016/j.eswa.2023.119696>, 2023.
- Lei, L., Sun, Y., Ouyang, B., Qiu, Y., Xie, C., Tang, G., Zhou, W., He, Y., Wang, Q., Cheng, X., Fu, P., and Wang, Z.: Vertical Distributions of Primary and Secondary Aerosols in Urban Boundary Layer: Insights into Sources, Chemistry, and Interaction with Meteorology, *Environ. Sci. Technol.*, 55, 4542-4552, <https://doi.org/10.1021/acs.est.1c00479>, 2021.
- 685 Li, H., Yang, T., Du, Y., Tan, Y., and Wang, Z.: Interpreting hourly mass concentrations of PM_{2.5} chemical components with an optimal deep-learning model, *J. Environ. Sci.*, 151, 125-139, <https://doi.org/10.1016/j.jes.2024.03.037>, 2025a.
- Li, H., Yang, T., Song, Y., Tian, P., He, J., Tan, Y., Tian, Y., Sun, Y., and Wang, Z.: Unveiling the intricate dynamics of PM_{2.5} sulfate aerosols in the urban boundary layer: A pioneering two-year vertical profiling and machine learning-enhanced analysis in global Mega-City, *Urban Clim.*, 61, 102424, <https://doi.org/10.1016/j.uclim.2025.102424>, 2025b.
- 690 Li, M., Zhang, Z., Yao, Q., Wang, T., Xie, M., Li, S., Zhuang, B., and Han, Y.: Nonlinear responses of particulate nitrate to NO_x emission controls in the megalopolises of China, *Atmos. Chem. Phys.*, 21, 15135-15152, <https://doi.org/10.5194/acp-21-15135-2021>, 2021.
- 695 Liang, L., Engling, G., Cheng, Y., Zhang, X., Sun, J., Xu, W., Liu, C., Zhang, G., Xu, H., Liu, X., and Ma, Q.: Influence of High Relative Humidity on Secondary Organic Carbon: Observations at a Background Site in East China, *J. Meteor. Res.*, 33, 905-913, <https://doi.org/10.1007/s13351-019-8202-2>, 2019.
- Lin, G. Y., Chen, H. W., Chen, B. J., and Chen, S. C.: A machine learning model for predicting PM_{2.5} and nitrate concentrations based on long-term water-soluble inorganic salts datasets at a road site station, *Chemosphere*, 289,

- 700 <https://doi.org/10.1016/j.chemosphere.2021.133123>, 2022.
- Liu, K., Zhang, Y., He, H., Xiao, H., Wang, S., Zhang, Y., Li, H., and Qian, X.: Time series prediction of the chemical components of PM_{2.5} based on a deep learning model, *Chemosphere*, 342, 140153, <https://doi.org/10.1016/j.chemosphere.2023.140153>, 2023.
- Liu, Q., Liu, D., Gao, Q., Tian, P., Wang, F., Zhao, D., Bi, K., Wu, Y., Ding, S., Hu, K., Zhang, J., Ding, D., and Zhao, C.:
705 Vertical characteristics of aerosol hygroscopicity and impacts on optical properties over the North China Plain during winter, *Atmos. Chem. Phys.*, 20, 3931-3944, <https://doi.org/10.5194/acp-20-3931-2020>, 2020.
- Liu, S., Geng, G., Xiao, Q., Zheng, Y., Liu, X., Cheng, J., and Zhang, Q.: Tracking Daily Concentrations of PM_{2.5} Chemical Composition in China since 2000, *Environ. Sci. Technol.*, 56, 16517-16527, <https://doi.org/10.1021/acs.est.2c06510>, 2022.
- Liu, Z., Hu, B., Wang, L., Wu, F., Gao, W., and Wang, Y.: Seasonal and diurnal variation in particulate matter (PM₁₀ and
710 PM_{2.5}) at an urban site of Beijing: analyses from a 9-year study, *Environ. Sci. Pollut. Res.*, 22, 627-642, <https://doi.org/10.1007/s11356-014-3347-0>, 2015.
- Lowenthal, D. H. and Kumar, N.: Evaluation of the IMPROVE Equation for estimating aerosol light extinction. *J. Air Waste Manage.*, 66, 726-737. <https://doi.org/10.1080/10962247.2016.1178187>, 2016.
- Lundberg, S. M., Erion, G., Chen, H., DeGrave, A., Prutkin, J. M., Nair, B., Katz, R., Himmelfarb, J., Bansal, N., and Lee,
715 S.-I.: From local explanations to global understanding with explainable AI for trees, *Nat. Mach. Intell.*, 2, 56-67, <https://doi.org/10.1038/s42256-019-0138-9>, 2020.
- Lv, L., Wei, P., Li, J., and Hu, J.: Application of machine learning algorithms to improve numerical simulation prediction of PM_{2.5} and chemical components, *Atmos. Pollut. Res.*, 12, 101211, <https://doi.org/10.1016/j.apr.2021.101211>, 2021.
- Ma, T., Xiang, G., Shi, Y., Liu, Y.: Horizontal in situ stresses prediction using a CNN-BiLSTM-attention hybrid neural
720 network, *Geomech. Geophys. Geo-energ. Geo-resour.* 8, 152, <https://doi.org/10.1007/s40948-022-00467-2>, 2022.
- Matus, A. V., Nowottnick, E. P., Yorks, J. E., and da Silva, A. M.: Enhancing surface PM_{2.5} air quality estimates in GEOS using CATS lidar data, *Earth and Space Sci.*, 12, e2024EA004078, <https://doi.org/10.1029/2024EA004078>, 2025.
- Meng, X., Hand, J. L., Schichtel, B. A., and Liu, Y.: Space-time trends of PM_{2.5} constituents in the conterminous United States estimated by a machine learning approach, 2005–2015, *Environ. Int.*, 121, 1137-1147,
725 <https://doi.org/10.1016/j.envint.2018.10.029>, 2018.
- Menon, S., Hansen, J., Nazarenko, L., and Luo, Y. F.: Climate effects of black carbon aerosols in China and India, *Science*, 297, 2250-2253, <https://doi.org/10.1126/science.1075159>, 2002.
- Miao, Y., Zhang, X., Che, H., and Liu, S.: Influence of Multi-Scale Meteorological Processes on PM_{2.5} Pollution in Wuhan, Central China, *Front. Environ. Sci.*, 10, <https://doi.org/10.3389/fenvs.2022.918076>, 2022.
- 730 Morgan, W. T., Allan, J. D., Bower, K. N., Capes, G., Crosier, J., Williams, P. I., and Coe, H.: Vertical distribution of sub-micron aerosol chemical composition from North-Western Europe and the North-East Atlantic, *Atmos. Chem. Phys.*, 9, 5389-5401, <https://doi.org/10.5194/acp-9-5389-2009>, 2009.
- Nishizawa, T., Sugimoto, N., Matsui, I., Shimizu, A., and Okamoto, H.: Algorithms to retrieve optical properties of three

- component aerosols from two-wavelength backscatter and one-wavelength polarization lidar measurements considering nonsphericity of dust, *J. Quant. Spectrosc. Radiat. Transfer*, 112, 254-267, <https://doi.org/10.1016/j.jqsrt.2010.06.002>, 2011.
- 735 Nishizawa, T., Sugimoto, N., Matsui, I., Shimizu, A., Hara, Y., Itsushi, U., Yasunaga, K., Kudo, R., and Kim, S.-W.: Ground-based network observation using Mie-Raman lidars and multi-wavelength Raman lidars and algorithm to retrieve distributions of aerosol components, *J. Quant. Spectrosc. Radiat. Transfer*, 188, 79-93, <https://doi.org/10.1016/j.jqsrt.2016.06.031>, 2017.
- 740 Niu, Y., Li, X., Qi, B., and Du, R.: Variation in the concentrations of atmospheric PM_{2.5} and its main chemical components in an eastern China city (Hangzhou) since the release of the Air Pollution Prevention and Control Action Plan in 2013, *Air Qual. Atmos. Hlth.*, 15, 321-337, <https://doi.org/10.1007/s11869-021-01107-6>, 2022.
- O'Shea, K. and Nash, R. J. a. e.-p.: An Introduction to Convolutional Neural Networks, <https://doi.org/10.48550/arXiv.1511.08458>, 2015.
- 745 Pitchford, M., Malm, W., Schichtel, B., Kumar, N., Lowenthal, D., and Hand, J.: Revised Algorithm for Estimating Light Extinction from IMPROVE Particle Speciation Data, *J. Air Waste Manage.*, 57, 1326-1336, <https://doi.org/10.3155/1047-3289.57.11.1326>, 2007.
- Rasmussen, C. E.: Gaussian Processes in Machine Learning, in: *Advanced Lectures on Machine Learning: ML Summer Schools 2003*, Canberra, Australia, February 2 - 14, 2003, Tübingen, Germany, August 4 - 16, 2003, Revised Lectures, edited by: Bousquet, O., von Luxburg, U., and Rätsch, G., Springer Berlin Heidelberg, Berlin, Heidelberg, 63-71, https://doi.org/10.1007/978-3-540-28650-9_4, 2004.
- Roostaei, V., Gharibzadeh, F., Shamsipour, M., Faridi, S., and Hassanvand, M. S.: Vertical distribution of ambient air pollutants (PM_{2.5}, PM₁₀, NO_x, and NO₂); A systematic review, *Heliyon*, 10, e39726, <https://doi.org/10.1016/j.heliyon.2024.e39726>, 2024.
- 755 Shahriari, B., Swersky, K., Wang, Z., Adams, R. P., and Freitas, N. d.: Taking the Human Out of the Loop: A Review of Bayesian Optimization, *Proceedings of the IEEE*, 104, 148-175, <https://doi.org/10.1109/JPROC.2015.2494218>, 2016.
- Shan, L., Liu, Y., Tang, M., Yang, M., Bai, X.: CNN-BiLSTM hybrid neural networks with attention mechanism for well log prediction, *J. Petrol. Sci. Eng.*, 205, 108838, <https://doi.org/10.1016/j.petrol.2021.108838>, 2021.
- Siarni-Namini, S., Tavakoli, N., and Namin, A. S.: The Performance of LSTM and BiLSTM in Forecasting Time Series, 2019 IEEE Int. Conf. on Big Data (Big Data), Los Angeles, CA, USA, 3285-3292, <https://doi.org/10.1109/BigData47090.2019.9005997>, 2019.
- 760 Song, Q., Huang, L., Zhang, Y., Li, Z., Wang, S., Zhao, B., Yin, D., Ma, M., Li, S., Liu, B., Zhu, L., Chang, X., Gao, D., Jiang, Y., Dong, Z., Shi, H., and Hao, J.: Driving Factors of PM_{2.5} Pollution Rebound in North China Plain in Early 2023, *Environ. Sci. Technol. Lett.*, 12, 305-312, <https://doi.org/10.1021/acs.estlett.4c01153>, 2025.
- 765 Srinivas, N. and Deb, K.: Multiobjective Optimization Using Nondominated Sorting in Genetic Algorithms, *Evol. Comput.*, 2, 221-248, <https://doi.org/10.1162/evco.1994.2.3.221>, 1994.
- 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, *Geophys. Res. Lett.*, 30, <https://doi.org/10.1029/2002gl016349>, 2003.
- 770 Sugimoto, N., Shimizu, A., Matsui, I., Uno, I., Arao, K., Dong, X., Zhao, S., Zhou, J., and Lee, C.-H.: Study of Asian Dust Phenomena in 2001–2003 Using A Network of Continuously Operated Polarization Lidars, *Water, Air, & Soil Pollution: Focus*, 5, 145-157, <https://doi.org/10.1007/s11267-005-0732-1>, 2005.
- Sun, Y., Du, W., Wang, Q., Zhang, Q., Chen, C., Chen, Y., Chen, Z., Fu, P., Wang, Z., Gao, Z., and Worsnop, D. R.: Real-Time Characterization of Aerosol Particle Composition above the Urban Canopy in Beijing: Insights into the Interactions
775 between the Atmospheric Boundary Layer and Aerosol Chemistry, *Environ. Sci. Technol.*, 49, 11340-11347, <https://doi.org/10.1021/acs.est.5b02373>, 2015.
- Tan, T., Hu, M., Li, M., Guo, Q., Wu, Y., Fang, X., Gu, F., Wang, Y., and Wu, Z.: New insight into PM_{2.5} pollution patterns in Beijing based on one-year measurement of chemical compositions, *Sci. Total Environ.*, 621, 734-743, <https://doi.org/10.1016/j.scitotenv.2017.11.208>, 2018.
- 780 Tao, J., Zhang, L., Cao, J., and Zhang, R.: A review of current knowledge concerning PM_{2.5} chemical composition, aerosol optical properties and their relationships across China, *Atmos. Chem. Phys.*, 17, 9485-9518, <https://doi.org/10.5194/acp-17-9485-2017>, 2017.
- Tao, Z., Wang, Z., Yang, S., Shan, H., Ma, X., Zhang, H., Zhao, S., Liu, D., Xie, C., and Wang, Y.: Profiling the PM_{2.5} mass concentration vertical distribution in the boundary layer, *Atmos. Meas. Tech.*, 9, 1369-1376, <https://doi.org/10.5194/amt-9-1369-2016>, 2016.
- 785 Tesche, M., Ansmann, A., Müller, D., Althausen, D., Mattis, I., Heese, B., Freudenthaler, V., Wiegner, M., Esselborn, M., Pisani, G., and Knippertz, P.: Vertical profiling of Saharan dust with Raman lidars and airborne HSRL in southern Morocco during SAMUM, *Tellus B: Chem. Phys. Meteor.*, 61, 144-164, <https://doi.org/10.1111/j.1600-0889.2008.00390.x>, 2009.
- Toth, T. D., Zhang, J., Vaughan, M. A., Reid, J. S., and Campbell, J. R.: Retrieving particulate matter concentrations over the
790 contiguous United States using CALIOP observations, *Atmos. Environ.*, 274, 118979, <https://doi.org/10.1016/j.atmosenv.2022.118979>, 2022.
- Verma, S., Pant, M., and Snasel, V.: A Comprehensive Review on NSGA-II for Multi-Objective Combinatorial Optimization Problems, *IEEE Access*, 9, 57757-57791, <https://doi.org/10.1109/ACCESS.2021.3070634>, 2021.
- Wang, F., Yang, T., Wang, Z., Wang, H., Chen, X., Sun, Y., Li, J., Tang, G., and Chai, W.: Algorithm for vertical distribution
795 of boundary layer aerosol components in remote-sensing data, *Atmos. Meas. Tech.*, 15, 6127-6144, <https://doi.org/10.5194/amt-15-6127-2022>, 2022.
- Wang, J., Li, J., Ye, J., Zhao, J., Wu, Y., Hu, J., Liu, D., Nie, D., Shen, F., Huang, X., Huang, D. D., Ji, D., Sun, X., Xu, W., Guo, J., Song, S., Qin, Y., Liu, P., Turner, J. R., Lee, H. C., Hwang, S., Liao, H., Martin, S. T., Zhang, Q., Chen, M., Sun, Y., Ge, X., and Jacob, D. J.: Fast sulfate formation from oxidation of SO₂ by NO₂ and HONO observed in Beijing haze, *Nat. Commun.*, 11, 2844, <https://doi.org/10.1038/s41467-020-16683-x>, 2020.
- 800 Wang, S. and Zhang, Y.: An attention-based CNN model integrating observational and simulation data for high-resolution

- spatial estimation of urban air quality, *Atmos. Environ.*, 340, 120921, <https://doi.org/10.1016/j.atmosenv.2024.120921>, 2025.
- Wang, X., Zhang, R., Tan, Y., and Yu, W.: Dominant synoptic patterns associated with the decay process of PM_{2.5} pollution episodes around Beijing, *Atmos. Chem. Phys.*, 21, 2491–2508, <https://doi.org/10.5194/acp-21-2491-2021>, 2021.
- 805 Welton, E., Campbell, J., Spinhirne, J., and Scott, V. S.: Global monitoring of clouds and aerosols using a network of micropulse lidar systems, Second International Asia-Pacific Symposium on Remote Sensing of the Atmosphere, Environment, and Space, SPIE, <https://doi.org/10.1117/12.417040>, 2001.
- Wu, J., Chen, X.-Y., Zhang, H., Xiong, L.-D., Lei, H., and Deng, S.-H.: Hyperparameter Optimization for Machine Learning Models Based on Bayesian Optimization, *J. Electron. Sci. and Technol.*, 17, 26-40, [https://doi.org/10.11989/JEST.1674-](https://doi.org/10.11989/JEST.1674-862X.80904120)
- 810 [862X.80904120](https://doi.org/10.11989/JEST.1674-862X.80904120), 2019a.
- Wu, L. B., Ren, H., Wang, P., Chen, J., Fang, Y. T., Hu, W., Ren, L. J., Deng, J. J., Song, Y., Li, J., Sun, Y. L., Wang, Z. F., Liu, C. Q., Ying, Q., and Fu, P. Q.: Aerosol Ammonium in the Urban Boundary Layer in Beijing: Insights from Nitrogen Isotope Ratios and Simulations in Summer 2015, *Environ. Sci. Technol. Lett.*, 6, 389-395, <https://doi.org/10.1021/acs.estlett.9b00328>, 2019b.
- 815 Xu, Y., Zhu, B., Shi, S., and Huang Y.: Two Inversion Layers and Their Impacts on PM_{2.5} Concentration over the Yangtze River Delta, China, *J. Appl. Meteor. Climatol.*, 58, 2349-2362, <https://doi.org/10.1175/JAMC-D-19-0008.1>, 2019.
- Yang, M. and Wang, J.: Adaptability of Financial Time Series Prediction Based on BiLSTM, *Procedia Comput. Sci.*, 199, 18-25, <https://doi.org/10.1016/j.procs.2022.01.003>, 2022.
- Yang, T., Li, H., Xu, W., Song, Y., Xu, L., Wang, H., Wang, F., Sun, Y., Wang, Z., and Fu, P.: Strong Impacts of Regional
- 820 Atmospheric Transport on the Vertical Distribution of Aerosol Ammonium over Beijing, *Environ. Sci. Technol. Lett.*, 11, 29-34, <https://doi.org/10.1021/acs.estlett.3c00791>, 2024.
- Yang, T., Wang, H., Li, H., Guo, X., Wang, D., Chen, X., Wang, F., Xin, J., Sun, Y., and Wang, Z.: Quantitative attribution of wintertime haze in coastal east China to local emission and regional intrusion under a stagnant internal boundary layer, *Atmos. Environ.*, 276, <https://doi.org/10.1016/j.atmosenv.2022.119006>, 2022.
- 825 Zhang, J., Su, Y., Chen, C., Guo, W., Tan, Q., Feng, M., Song, D., Jiang, T., Chen, Q., Li, Y., Li, W., Wang, Y., Huang, X., Han, L., Wu, W., and Wang, G.: Chemical composition, sources and formation mechanism of urban PM_{2.5} in Southwest China: a case study at the beginning of 2023, *Atmos. Chem. Phys.*, 24, 2803-2820, [https://doi.org/10.5194/acp-24-2803-](https://doi.org/10.5194/acp-24-2803-2024)
- [2024](https://doi.org/10.5194/acp-24-2803-2024), 2024.
- Zhang, J., Ye, L., Lai, Y.: Stock price prediction using CNN-BiLSTM-Attention model, *Mathematics*, 11, 1985, <https://doi.org/10.3390/math11091985>, 2023.
- 830 Zhao, C., Sun, Y., Yang, J., Li, J., Zhou, Y., Yang, Y., Fan, H., and Zhao, X.: Observational evidence and mechanisms of aerosol effects on precipitation, *Sci. Bull.*, 69, 1569-1580, <https://doi.org/10.1016/j.scib.2024.03.014>, 2024.
- Zhao, J., Du, W., Zhang, Y., Wang, Q., Chen, C., Xu, W., Han, T., Wang, Y., Fu, P., Wang, Z., Li, Z., and Sun, Y.: Insights into aerosol chemistry during the 2015 China Victory Day parade: results from simultaneous measurements at ground level
- 835 and 260 m in Beijing, *Atmos. Chem. Phys.*, 17, 3215-3232, <https://doi.org/10.5194/acp-17-3215-2017>, 2017.

Zhou, W., Sun, Y., Xu, W., Zhao, X., Wang, Q., Tang, G., Zhou, L., Chen, C., Du, W., Zhao, J., Xie, C., Fu, P., and Wang, Z.: Vertical Characterization of Aerosol Particle Composition in Beijing, China: Insights From 3-Month Measurements With Two Aerosol Mass Spectrometers, *J. Geophys. Res.: Atmos.*, 123, 13016-13029, <https://doi.org/10.1029/2018JD029337>, 2018.

840 Zhu, H., Yang, S., Zhao, H., Wang, Y., and Li, R.: Complex interplay of sulfate aerosols and meteorology conditions on precipitation and latent heat vertical structure, *npj Clim. Atmos. Sci.*, 7, 191, <https://doi.org/10.1038/s41612-024-00743-w>, 2024.