https://doi.org/10.5194/egusphere-2026-397
Preprint. Discussion started: 17 February 2026

Correcting Aerosol Extinction Coefficient Vertical Structure Biases in GEOS-Chem via a Physics-Informed Transformer with Physical Mechanism Diagnosis

Jiajun Xiong[1], Yi Wang[1], Jun Wang[2,3], Yanyu Wang[4], Meng Zhou[5,6], Minghui Tao[1], Wenhui Dong[1], Jhoon Kim[7], Lunche Wang[1]

[1]Hubei Key Laboratory of Regional Ecology and Environmental Change, School of Geography and Information Engineering, China University of Geosciences, Wuhan, 430074, China

[2]Department of Chemical and Biochemical Engineering, The University of Iowa, Iowa City, IA 52242, USA

[3]Center for Global and Regional Environmental Research, The University of Iowa, Iowa City, IA 52242, USA

[4]State Environmental Protection Key Laboratory of Formation and Prevention of Urban Air Pollution Complex, Shanghai Academy of Environmental Sciences, Shanghai 200233, China

[5]Goddard Earth Sciences Technology and Research, University of Maryland, Baltimore County, MA 21250, USA

[6]NASA/Goddard Space Flight Center, Global Modeling and Assimilation Office, Greenbelt, MD 20771, USA

[7]Department of Atmospheric Sciences, Yonsei University, Seoul, 03722, South Korea

Correspondence to: Yi Wang (wangyi34@cug.edu.cn)

https://doi.org/10.5194/egusphere-2026-397
Preprint. Discussion started: 17 February 2026

Abstract

We propose a physics-informed Transformer framework to correct biases in the Aerosol Extinction Coefficient (AEC, km^{-1}) profiles simulated by GEOS-Chem. Unlike standard Transformer, our framework features a dual-stream architecture with explicit physical constraints. It employs Gated Feature Fusion to integrate vertical structures (combining GEOS-Chem priors with MERRA-2 profiles) by dynamically identifying height-dependent drivers, and leverages Cross-Attention to incorporate MERRA-2 surface environmental constraints for modulating AEC vertical reconstruction with synoptic contexts. This approach effectively predicts systematic biases relative to Cloud-Aerosol Lidar with Orthogonal Polarization satellite observations and resolves AEC profiles, surpassing methods retrieving only aerosol layer heights. "Leave-One-Year-Out" validation over East Asia during 2017–2019 demonstrates significant AEC fidelity improvements, increasing R from 0.49–0.53 in the GEOS-Chem simulations to 0.66–0.73 and reducing RMSE by approximately 25%. The model effectively mitigates over-diffusion, significantly reducing AEC simulation biases in the critical near-surface layer while restoring smoothed biomass burning and dust plumes. Additionally, it exhibits robust cross-continental transferability, reproducing bias patterns over North American domain (R=0.70) without retraining, confirming the internalization of universal physicochemical relationships linking atmospheric states to simulation biases. Furthermore, interpretability analysis establishes a feedback loop from data-driven correction to physical model improvement. The model identifies temperature and sensible heat flux as primary drivers to constrain boundary layer mixing, and uses environmental proxies (e.g., vegetation indices) to diagnose deficiencies in dust uplift and secondary aerosol formation. These insights provide a physical basis for refining parameterization schemes in chemical transport models.

https://doi.org/10.5194/egusphere-2026-397
Preprint. Discussion started: 17 February 2026

Keywords: Aerosol extinction coefficient profiles, Deep learning, Chemical transport models, CALIOP, Model interpretability

https://doi.org/10.5194/egusphere-2026-397
Preprint. Discussion started: 17 February 2026

1. Introduction

Atmospheric aerosols are key modulators of the Earth's climate radiative forcing (Hayasaka et al., 2007; Stier et al., 2007; Koren et al., 2004; Liu et al., 2011; Wang et al., 2023; Kahn et al., 2005; Dong et al., 2023) and environmental health (Chen et al., 2022b; Song et al., 2025). As Aerosol Optical Depth (AOD) serves as a fundamental column-integrated parameter (Kaufman et al., 2002), it fails to reveal critical vertical distribution information (Winker et al., 2010; Lu et al., 2025b; Zhu et al., 2024). In reality, the vertical structure of aerosols—characterized by Aerosol Extinction Coefficient (AEC) profiles (Xiong et al., 2025; Jiang et al., 2024; Zhen et al., 2024) and layer heights (Fan et al., 2025; Kim et al., 2025; Wang et al., 2013; Christian et al., 2019; Yorks et al., 2023)—is the core physical quantity determining their climatic and environmental effects. Specifically, the vertical distribution governs the atmospheric radiative energy budget and thermal structure; the altitude of absorbing aerosols determines heating rate profile (Lu et al., 2020; Xu et al., 2016) and atmospheric stability (Koch and Del Genio, 2010), and failure to resolve this vertical feature leads to significant radiative forcing estimate biases (Zarzycki and Bond, 2010; Myhre et al., 2013; Shi et al., 2021; Wang et al., 2014). Second, the vertical co-existence of aerosols and clouds is a prerequisite for aerosol-cloud interactions (Zarzycki and Bond, 2010; Ge et al., 2014), meaning that relying solely on column totals hinders the accurate quantification of aerosol impacts on cloud microphysics (Lu et al., 2025b; Wilcox, 2012; Christensen et al., 2020; Zhao et al., 2019). Finally, vertical stratification bridges regional pollution with global transport: aerosols uplifted into the free troposphere undergo intercontinental transport (Val Martin et al., 2013; Weinzierl et al., 2017; Choi et al., 2024), whereas those trapped within the Planetary Boundary Layer (PBL) directly determine surface $PM_{2.5}$ concentrations (Chen et al., 2022a; Handschuh et al., 2022) and health outcomes (Diner et al., 2018). Therefore, precisely characterizing AEC profiles is not only a vital supplement to traditional AOD analysis (Lv et al., 2016; Wei et al., 2019) but also a cornerstone for reducing climate model uncertainties and

https://doi.org/10.5194/egusphere-2026-397
Preprint. Discussion started: 17 February 2026

understanding transboundary pollutant transport.

Despite the recognized importance of aerosol vertical distribution, acquiring high-precision, spatiotemporally continuous, and vertically resolved three-dimensional (3D) aerosol information globally remains a formidable challenge. Current characterization methods primarily follow two pathways: passive remote sensing retrieval and active remote sensing observation. With extensive swath widths, passive satellite sensors, such as the TROPOspheric Monitoring Instrument (TROPOMI) and Himawari-8, have made strides in retrieving Aerosol Layer Height (ALH) (Lu et al., 2025a; Nanda et al., 2020). Utilizing differential absorption in oxygen bands (O_2 A/B) (Xu et al., 2017; Sanders et al., 2015; Nanda et al., 2018; Kim et al., 2025; Xu et al., 2024; Xu et al., 2019; Ding et al., 2016; Chen et al., 2021b; Chen et al., 2021a; Chen et al., 2025) or multi-angle stereoscopic techniques (Kahn et al., 2005; Dubovik et al., 2011), these algorithms estimate the effective ALH. However, the effective ALH merely represents the optical centroid or a vertically weighted average of the aerosol column, rather than a resolved vertical profile (Chimot et al., 2018; Lu et al., 2025b). While valuable for tracking plumes (Huang et al., 2024), it fails to resolve complex multi-layer structures (Winker et al., 2013) and cannot provide the vertical gradients of AEC necessary for calculating radiative heating rates (Samset et al., 2013). In contrast, spaceborne active lidars—specifically the Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP)—offer high-resolution Aerosol Extinction Profiles (AEPs) that resolve aerosol stratification (Huang et al., 2013; Liu et al., 2012; Qu et al., 2016; Ma et al., 2013; Wang et al., 2021a). Nevertheless, constrained by narrow nadir-viewing geometry (Winker et al., 2009; Hunt et al., 2009) and sparse sampling, active remote sensing suffers from substantial observational gaps, producing a "curtain-like" dataset rather than planar coverage (Koffi et al., 2012b; Colarco et al., 2014). To bridge these gaps, Chemical Transport Models (CTMs, e.g., GEOS-Chem and WRF-Chem) are indispensable for simulating continuous 3D aerosol fields (Xiong et al., 2025; Jiang et al., 2024). However, these models are often hampered by uncertainties in emission

https://doi.org/10.5194/egusphere-2026-397
Preprint. Discussion started: 17 February 2026

inventories (Bond et al., 2013; Hoesly et al., 2018; Liang et al., 2023) and simplified physical parameterizations (e.g., wet scavenging, boundary layer turbulent mixing) (Xiong et al., 2025; Du et al., 2020; Zeng et al., 2020), leading to systematic biases and misalignment with lidar observations (Koffi et al., 2016; Koffi et al., 2012a; Kim et al., 2015).

To bridge the gap between sparse observations and biased simulations, traditional studies have employed Data Assimilation (DA) techniques, such as 3D-Var or Kalman Filtering (Henze et al., 2007; Henze et al., 2009; Wang et al., 2020a; Wang et al., 2020b; Zhang et al., 2011; Sekiyama et al., 2010). However, operational DA schemes often face prohibitive computational costs (Benedetti et al., 2018) and rely on Gaussian error assumptions that may not match the complex, non-linear characteristics of aerosol processes (Bocquet et al., 2010; Geer, 2021). Driven by the exponential growth of atmospheric big data, Deep Learning (DL) has emerged as a transformative alternative for bias correction and multi-source data fusion (Geer, 2021; Wang et al., 2022; Xing et al., 2022; Fan et al., 2025). Early studies successfully applied Random Forests or Deep Neural Networks to improve AOD retrieval (Sanders et al., 2015) or surface $PM_{2.5}$ estimation (Hu et al., 2014; Wei et al., 2019; Li et al., 2017). More recently, Convolutional Neural Networks (CNNs) have been utilized to estimate AEPs by extracting spatial texture features from passive satellite imagery (Zhen et al., 2024). However, these methods typically simplify profile estimation into a standard regression problem (Pashayi et al., 2025; Li et al., 2020) or rely heavily on two-dimensional (2D) spatial convolutions (Daoud et al., 2021). Consequently, they often neglect the inherent sequential correlation and physical coupling of the aerosol vertical structure. In reality, the aerosol state at a specific altitude is intimately linked to layers above and below through turbulent mixing and sedimentation (Zeng et al., 2020; Xiong et al., 2025; Du et al., 2020). Standard CNNs or pixel-wise models struggle to capture this long-range dependency along the vertical axis (Wang et al., 2021b) and often overlook the structural priors provided by CTMs (Reichstein et al., 2019; Geer, 2021). To overcome

https://doi.org/10.5194/egusphere-2026-397
Preprint. Discussion started: 17 February 2026

these limitations, the Transformer architecture (Vaswani et al., 2017), originally designed for sequence modeling, offers a superior solution. Its self-attention mechanism dynamically weighs information from different altitude layers, effectively capturing global vertical dependencies within the atmospheric column. This capability makes it an ideal tool for fusing the physical consistency of GEOS-Chem simulations with the vertical precision of CALIOP observations to reconstruct high-fidelity 3D AEC fields.

To address the spatiotemporal sparsity of spaceborne lidar observations and the inherent simulation biases of CTMs, this study proposes a physics-informed Transformer DL framework aimed at reconstructing high-precision, spatially continuous 3D AEC fields. Distinct from data-driven retrieval methods that rely solely on passive remote sensing imagery, our approach explicitly embeds the vertical structural priors provided by GEOS-Chem simulations and meteorological constraints from Modern-Era Retrospective analysis for Research and Applications, Version 2 (MERRA-2) reanalysis data. Leveraging the self-attention mechanism unique to the Transformer architecture, the model effectively captures and corrects the complex, non-linear bias patterns between simulated results and CALIOP "ground truth". Focusing on East Asia (EA) from 2017 to 2019, we systematically evaluate the framework's performance in 3D AEC field reconstruction and spatiotemporal generalization. Crucially, to overcome the "black-box" limitations of DL, we integrate SHapley Additive exPlanations (SHAP) (Lundberg and Lee, 2017) with attention weight analysis to quantitatively parse the contributions of meteorological factors and aerosol variables to the AEP reconstruction. This process not only enables an interpretable diagnosis of CTM simulation biases—identifying the dominant drivers of bias within specific altitudes or regions—but also provides a solid physical basis for refining GEOS-Chem's physical parameterization schemes and emission inventories, thereby establishing a feedback loop from "data-driven correction" to "physical mechanism improvement." The remainder of this paper is organized as follows. Section 2

https://doi.org/10.5194/egusphere-2026-397
Preprint. Discussion started: 17 February 2026

introduces the multi-source datasets and model configurations. Section 3 details the physics-informed Transformer architecture and evaluation strategies. Section 4 evaluates the bias correction performance against observations and elucidates the physical drivers of simulation biases through interpretability analysis. Finally, Section 5 summarizes the conclusions.

2. Data and Model Configuration

2.1 Study Region

This study focuses primarily on EA (0°-60°N, 70°-150°E, Fig. 1), a domain characterized by complex aerosol sources and intense aerosol-meteorology interactions, serving as the core region for training and evaluating the bias-correction model. Within the EA domain, six Regions of Interest (ROIs) are selected to evaluate performance across distinct aerosol regimes (Fig. 1): the anthropogenic-dominated North China Plain (NCP) and Indo-Gangetic Plain (IGP), the dust-active Taklamakan Desert, the smoke-impacted Indochina Peninsula, and the cleaner Tibetan Plateau and Western Pacific. To investigate vertical characteristics in the lower troposphere extending from the surface to 6 km Above Ground Level (AGL), we establish three spatial diagnostics: a 38°N transect (70°-150°E) capturing the zonal source-to-sink continuum driven by the Westerlies; a 40°N transect (120°-150°E) targeting the dispersion of continental outflow over marine surfaces; and a domain-wide zonal mean profile (0°-60°N) to reveal the macroscopic latitudinal dependence of aerosol loading and layer heights. To rigorously assess the model's spatial generalization capability, we additionally define an independent validation sub-region over North America (NA, 25°-55°N, 70°-130°W).

2.2 CALIOP

The CALIOP sensor onboard the Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observations (CALIPSO) satellite serves as the observational benchmark for characterizing aerosol vertical structure and quantifying simulation biases. We utilize the Level 2 Version 4.51 Aerosol Profile product, which provides vertical profiles of

https://doi.org/10.5194/egusphere-2026-397
Preprint. Discussion started: 17 February 2026

AEC at 532 nm and 1064 nm. The standard Level 2 product offers a uniform horizontal resolution of 5 km and a vertical resolution of 60 m throughout the troposphere (–0.5 km to 20 km). Given that aerosol loading is predominantly confined to the PBL and lower free troposphere (Xiong et al., 2025; Jiang et al., 2024), our analysis is restricted to the altitude range of 0–6 km AGL.

To ensure data reliability, we implement a rigorous, tiered Quality Control (QC) procedure to exclude profiles contaminated by clouds or compromised by retrieval artifacts (details in Sect. S1). This stringent assessment framework substantially enhances the fidelity of the CALIOP dataset for model training (Hong and Di Girolamo, 2022; Jiang et al., 2024; Xiong et al., 2025; Mehta et al., 2023; Winker et al., 2013). In our physics-informed data-driven framework, the collocated CALIOP AEPs are treated as the "ground truth". The training target for the Transformer model is explicitly defined as the simulation bias, calculated as the deviation of the GEOS-Chem simulated AEC from the corresponding CALIOP observations (details in Section 3.1). However, it is essential to acknowledge the intrinsic uncertainties associated with satellite retrievals. Validations against ground-based AERONET measurements indicate a mean relative bias in CALIOP AOD of approximately -5.1% ± 8.5% (Kim et al., 2018). Furthermore, the agreement between CALIOP 532 nm calibrated attenuated backscatter and airborne High Spectral Resolution Lidar (HSRL) measurements is typically within 1.0% ± 3.5% (Getzewich et al., 2018). These inherent observational uncertainties impose a theoretical upper limit on the precision of the bias correction method proposed herein.

2.3 AERONET

To independently evaluate the model's capability in capturing high-frequency temporal variations—bridging the observational gaps inherent to the sparse sampling of polar-orbiting satellites—we utilize AOD data from the AERONET (Holben et al., 1998). Specifically, the Version 3 Level 2.0 (cloud-screened and quality-assured) AOD data are employed (Giles et al., 2019). Since AERONET instruments do not directly

https://doi.org/10.5194/egusphere-2026-397
Preprint. Discussion started: 17 February 2026

measure AOD at 532 nm, data are interpolated to this wavelength using the Ångström exponent derived from adjacent channels (details in Sect. S2). Two representative sites are selected for detailed case studies, targeting distinct aerosol regimes during high-loading episodes (Fig. 1): (1) Kanpur (26.5°N, 80.2°E): Located in the IGP, this site is characterized by a complex mixture of anthropogenic pollution and transported dust. We focus on a 7-day window centered on 30 April 2019, representing a typical pre-monsoon scenario driven by regional dust and haze activity (Misra et al., 2014; Singh et al., 2004; Chinnam et al., 2006); (2) Nong Khai (17.9°N, 102.7°E): Situated in the Indochina Peninsula, this site is dominated by carbonaceous aerosols resulting from intense agricultural biomass burning (Munroe et al., 2008; Nguyen et al., 2021). The analysis window is centered on 22 February 2019, coinciding with the peak of the regional active fire season (Tsay et al., 2013). Examining these high-loading cases allows us to assess the robustness of the model in reproducing the dynamic evolution of pollution accumulation and dissipation, providing a rigorous test of the temporal consistency of our bias correction approach.

2.4 GEOS-Chem and Meteorological Reanalysis

We employ the GEOS-Chem (Bey et al., 2001) global CTM (version 13.4.0) to simulate the 3D distribution of atmospheric composition, focusing on the EA domain defined in Section 2.1. The simulation spans from 1 January 2017 to 31 December 2019. The model is driven by the MERRA-2 assimilated meteorological fields provided by the NASA Global Modeling and Assimilation Office (GMAO) (Gelaro et al., 2017). To accommodate the spatially sparse sampling inherent to CALIOP's nadir-viewing geometry, the simulation is configured at a horizontal resolution of 2° × 2.5° (latitude × longitude) with 47 vertical hybrid sigma-pressure levels extending from the surface to 0.01 hPa. Anthropogenic emissions are prescribed by the Community Emissions Data System (CEDSv2) (Mcduffie et al., 2020). Biogenic emissions of Volatile Organic Compounds (BVOCs) are prescribed using the offline MEGAN inventory (Guenther et al., 2012). Biomass burning emissions are derived from the Global Fire Emissions

https://doi.org/10.5194/egusphere-2026-397
Preprint. Discussion started: 17 February 2026

Database version 4 (GFED4) (Giglio et al., 2013). GEOS-Chem is operated in the "Standard Full Chemistry" mode, utilizing a fully coupled NO_x-O_x-hydrocarbon-aerosol mechanism. The aerosol simulation encompasses sulfate, nitrate, ammonium, carbonaceous aerosols, mineral dust, and sea salt. Both dry and wet deposition processes are explicitly treated to characterize aerosol sinks. Vertical transport is parameterized using a non-local PBL mixing scheme, which accounts for deep convective updrafts and turbulent diffusion (Lin and Mcelroy, 2010). To facilitate direct comparison with CALIOP observations, the model is configured to diagnose and archive 3D AEC directly at 532 nm with a 1-hour temporal resolution.

MERRA-2 reanalysis data serve a dual purpose in this study: acting as the driving meteorological field for the GEOS-Chem and constituting the key physical input features for the physics-informed Transformer model. Integrating these meteorological state variables into the DL framework aims to explicitly capture the non-linear dependence of simulation biases on synoptic and micro-meteorological conditions. Generated by the GEOS-5 atmospheric general circulation model combined with 3D variational DA, MERRA-2 provides globally consistent physical fields. To maintain consistency within our "physics-informed data-driven" framework, we extract meteorological variables that are strictly consistent with those driving GEOS-Chem. These variables are regridded to the 2° × 2.5° resolution to achieve strict spatial alignment with the aerosol simulation outputs, jointly constructing the "meteorological background" vector in the neural network's input layer.

3. Method

3.1 Input Feature Construction and Target Definition

To physically decouple the local vertical structure that determines the instantaneous AEC magnitude from the macroscopic meteorological background that controls pollutant accumulation and transport, we construct a comprehensive input feature space utilizing a dual-stream architecture: the Vertical Profile Stream and the

https://doi.org/10.5194/egusphere-2026-397
Preprint. Discussion started: 17 February 2026

Global Context Stream. A detailed inventory of all input variables, including physical descriptions, units, and data sources, is provided in Sect. S3.

The Vertical Profile Stream encapsulates the explicit vertical structure of the atmosphere. We integrate three sub-components to provide the necessary physicochemical and spatiotemporal context for each layer. (1) Physicochemical State Profiles: Core inputs include GEOS-Chem simulated aerosol component concentrations and MERRA-2 meteorological profiles. Beyond basic mass concentrations, we explicitly introduce microphysical and precursor variables—such as hygroscopic growth factors, aerosol effective radii, and precursor gas concentrations (SO_2, NO_x, NH_3)—to physically constrain chemical transformations and optical extinction processes. (2) Vertical Coordinate Encoding: Since the self-attention mechanism is permutation-invariant (i.e., it treats the input sequence as a bag of vectors without inherent order), explicit height information is essential for resolving atmospheric stratification. Without this embedding, the model cannot distinguish between surface emissions and high-altitude transport layers, potentially confusing the distinct physical processes governing aerosols at different levels. We embed vertical coordinates into high-dimensional vectors and add them to the feature sequence, enabling the model to distinguish between surface interactions and upper-level transport. (3) Spatiotemporal Embeddings: Although geographical (latitude, longitude) and temporal indices (month, day, night) are scalars, they are critical for characterizing regional emission patterns and seasonal cycles. We project these scalars into high-dimensional vectors matching the vertical dimensions of the profile data, ensuring the model captures location-dependent and time-varying bias patterns.

The Global Context Stream incorporates 2D surface diagnostic variables to capture synoptic-scale constraints affecting the atmospheric column. We select the Planetary Boundary Layer Height (PBLH), surface pressure, and friction velocity to constrain vertical mixing and turbulence intensity. Surface fluxes (latent/sensible heat), soil moisture, and Leaf Area Index (LAI) are included to parameterize dry deposition

https://doi.org/10.5194/egusphere-2026-397
Preprint. Discussion started: 17 February 2026

velocities and biogenic emissions. Additionally, total precipitation and evaporation rates serve as proxies for wet scavenging efficiency. These variables aid the model in distinguishing between local pollution accumulation and large-scale ventilation conditions.

We define the learning target Δ_{AEC}^{target} as the systematic bias of GEOS-Chem simulated AEC (AEC_{GC}) evaluated against CALIOP observation (AEC_{CAL}):

$$\Delta_{AEC}^{target} = AEC_{GC} - AEC_{CAL} \quad (1)$$

By modeling the bias Δ_{AEC}^{target}, the framework preserves the physical patterns captured by the CTM while specifically correcting systematic biases arising from input uncertainties and simplified parameterizations.

3.2 Physically-Informed Transformer Architecture

The overall architecture of the proposed framework is illustrated in Figure 2. To effectively bridge the gap between GEOS-Chem simulations and CALIOP satellite observations, the model is structured into five physically distinct modules: a dual-stream input layer for decoupling the Vertical Profile and Global Context Stream, a Feature Embedding layer equipped with a Gated Feature Fusion mechanism, Transformer encoder stack for capturing vertical dependencies, a Cross-Attention module for incorporating synoptic constraints and a residual output generation layer. Rather than treating all inputs uniformly, the model processes the two input streams using embedding strategies specifically tailored to their physical nature (as detailed in Sect. S4a, b), ensuring that the distinct structural characteristics of atmospheric vertical profiles and global environmental contexts are preserved in the high-dimensional feature representation.

3.2.1 Dynamic Embedding and Gated Feature Fusion

To explicitly capture the height-dependent variations in dominant physical drivers, we design a Gated Feature Fusion mechanism within the Vertical Profile Stream. This design is motivated by the physical consensus that the factors governing AEC vary

https://doi.org/10.5194/egusphere-2026-397
Preprint. Discussion started: 17 February 2026

significantly with altitude: near-surface AEC is predominantly driven by local emissions and chemical composition (Xiong et al., 2025; Jiang et al., 2024), whereas AEC in the free troposphere is often dictated by long-range transport pathways and regional background levels (Uno et al., 2009; Val Martin et al., 2013).

Instead of using a static concatenation which assumes fixed importance across the entire column, this mechanism dynamically weighs the contributions of four feature components: physicochemical profiles, vertical height information, spatial coordinates, and temporal indices. For each vertical layer z, the unified local embedding $E_{local}^{(z)}$ is computed as an adaptive weighted sum:

$$E_{local}^{(z)} = \sum_{k=1}^{4} \alpha_k^{(z)} \cdot Norm\left(h_k^{(z)}\right) \; s.t. \; \sum \alpha_k = 1 \tag{2}$$

where $h_k^{(z)}$ represents the embedding of the k-th feature group. The attention weights $\alpha_k^{(z)}$ are generated by a learnable gating network, allowing the model to autonomously identify and prioritize the most physically relevant information source for each specific altitude layer.

For the Global Context Stream, we address the challenge of integrating diverse environmental variables that possess distinct physical natures. Since different physical variables may share similar numerical ranges but possess distinct physical meanings, simple numerical projection inevitably leads to semantic ambiguity. We implement a Variable Identity Embedding strategy (Eq. S4) to encode the unique "physical identity" of each variable into the latent space. This ensures that the model explicitly distinguishes between different physical variables, preserving their distinct physical semantics during the cross-modal interaction.

3.2.2 Transformer Encoder and Cross-Modal Interaction

Following Gated Feature Fusion, the Vertical Profile Stream is processed by a Transformer Encoder stack to explicitly model the vertical connectivity of the atmosphere. Physically, aerosol layers are not isolated but are inherently coupled

https://doi.org/10.5194/egusphere-2026-397
Preprint. Discussion started: 17 February 2026

through vertical exchange processes such as turbulent mixing, deep convection, and gravitational sedimentation. Unlike CNNs which primarily focus on local spatial neighborhoods, the Multi-Head Self-Attention (MSA, detailed in Sect. S4c) mechanism enables the model to capture global vertical dependencies. This facilitates the necessary information flow between near-surface accumulation layers and high-altitude transport layers, ensuring that the reconstructed profile maintains physical continuity along the vertical axis.

To physically constrain the vertical reconstruction using synoptic-scale conditions, we introduce a Cross-Attention module. In this configuration, the encoded vertical profiles serve as Queries (Q), while the Global Context Stream embeddings act as both Keys (K) and Values (V). This mechanism enables the aerosol state at a specific layer (Q) actively scans the prevailing synoptic conditions (K)—such as surface wind speed or humidity—to retrieve the relevant environmental constraints (V) needed to correct its bias. This design effectively mimics the physical reality where local microphysical structures are conditioned by the macro-scale meteorological background, ensuring that the corrected AEC profiles are dynamically consistent with the large-scale weather patterns.

3.2.3 Residual Output Generation with Long Skip Connection

To generate the final bias prediction, we introduce a Long Skip Connection, adding the initial projected physicochemical features (H_{phyche}) directly to the output of the Cross-Attention module (H_{cross}) before the final projection:

$$\Delta_{AEC}^{pred} = MLP\left(H_{cross} + H_{phyche}\right) \cdot s + b \quad (3)$$

Where Δ_{AEC}^{pred} denotes the predicted systematic bias of the AEC. The parameters s and b are learnable scaling and bias terms, respectively, introduced to adaptively map the normalized network outputs to the physical magnitude of extinction biases. This design explicitly preserves physicochemical state information, allowing the regressor to access the baseline atmospheric state while utilizing context-aware features (e.g., thermodynamic stability, turbulent mixing intensity, and surface emission conditions)

https://doi.org/10.5194/egusphere-2026-397
Preprint. Discussion started: 17 February 2026

to disentangle complex bias sources.

3.3 Magnitude-Weighted Loss Function with False Alarm Penalty

The distribution of simulation biases exhibits significant heterogeneity. Standard Mean Squared Error (MSE) loss is often dominated by the numerous samples with small biases, causing the model to under-correct large systematic biases. To address this, we propose a Magnitude-Weighted Loss (L_{MW}):

$$L_{MW} = \frac{1}{N}\sum_{i=1}^{N}\left(\Delta_{AEC}^{pred} - \Delta_{AEC}^{target}\right)^2 \cdot \omega\left(\Delta_{AEC}^{target}\right) \quad (4)$$

$$\omega\left(\Delta_{AEC}^{target}\right) = \left|\Delta_{AEC}^{target}\right|^p + \lambda \cdot \mathrm{e}^{-\beta\left|\Delta_{AEC}^{target}\right|} + \epsilon \quad (5)$$

where Δ_{AEC}^{target} is the target bias and Δ_{AEC}^{pred} is the predicted bias. The weight function $\omega\left(\Delta_{AEC}^{target}\right)$ incorporates three hyperparameters determined through empirical validation: (1) Large-Error Prioritization ($\left|\Delta_{AEC}^{target}\right|^p$): This term ($p$ is a constant term.) amplifies the penalty for samples with large simulation biases, forcing the model to focus on instances where GEOS-Chem fails significantly. (2) False Alarm Penalty ($\lambda \cdot \mathrm{e}^{-\beta\left|\Delta_{AEC}^{target}\right|}$): This term imposes a penalty when the target bias is near zero but the model predicts a non-zero correction, effectively suppressing "false alarms" in layers where the CTM performance is already satisfactory. Both λ and β are constant terms. (3) Base Stability Term (ϵ): A small constant ensures numerical stability.

3.4 Model Evaluation Strategy

To comprehensively assess the robustness and generalization capability of the physics-informed Transformer model, we design a rigorous evaluation framework covering four dimensions.

(1) Spatial Block Cross-Validation: To mitigate information leakage caused by spatial autocorrelation (Geer, 2021) , we implement a spatial block K-fold cross-validation strategy (Sect. S5). The study region is divided into non-overlapping 4° × 5° blocks (aggregating 2 × 2 model grids). In each iteration, the model is trained on four folds and evaluated on the remaining spatially independent fold. This "checkerboard"

https://doi.org/10.5194/egusphere-2026-397
Preprint. Discussion started: 17 February 2026

approach ensures performance metrics reflect the model's ability to extrapolate to unseen geographic locations.

(2) Temporal Transferability: Given the interannual variability in emissions and meteorology (Xiong et al., 2025), we adopt a "Leave-One-Year-Out" validation scheme comprising three independent experiments (Table 1). This tests whether the model learns generalizable physical rules rather than overfitting to specific temporal patterns.

(3) External Spatial Generalization: To rigorously stress-test the model's transferability beyond its training distribution, we perform an out-of-domain evaluation on the independent NA sub-region defined in Section 2.1. By directly applying the model trained on EA data to this unseen continent—which possesses distinct aerosol sources and meteorological backgrounds—we evaluate whether the learned bias-correction mechanism captures universal physical laws rather than region-specific correlations.

(4) Independent Ground-based Validation: We employ ground-based AERONET observations as an independent physical benchmark. Predicted AEPs are vertically integrated to derive column AODs, which are then compared with AERONET data to assess the reproduction of high-frequency temporal evolution.

To quantify the model performance across these dimensions, we employ a comprehensive set of statistical metrics including the Pearson correlation coefficient (R), Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and Normalized Root Mean Square Error (NRMSE). Detailed mathematical definitions are provided in Sect. S6. NRMSE is specifically used to enable fair comparisons across vertical layers by normalizing biases against the exponentially decaying dynamic range of AEC.

3.5 Model Interpretability Framework

To elucidate the inference logic of the DL model and ensure physical consistency, we establish a comprehensive interpretability framework examining internal mechanisms and feature sensitivities to identify the physical drivers governing the bias correction.

https://doi.org/10.5194/egusphere-2026-397
Preprint. Discussion started: 17 February 2026

3.5.1 Dual-Mechanism Attribution

We employ distinct attribution methods for the hybrid inputs. First, for Vertical Profile Stream, we apply Gradient-based Attribution (detailed in Sect. S7a), utilizing the Input × Gradient method (Shrikumar et al., 2017) to quantify the sensitivity of bias correction to local variations in vertical structures. This identifies which specific altitude layers and species dominate the AEC simulation bias. Second, we conduct Attention-based Analysis (detailed in Sect. S7b) by extracting Cross-Attention weights to map the interaction strength between the Global Context Stream and the Vertical Profile Stream at specific vertical layers, revealing how macroscopic meteorology modulates vertical reconstruction biases.

3.5.2 Gated Information Fusion Analysis

We analyze the scalar weights learned by the gated fusion mechanism to understand information prioritization (detailed in Sect. S7c): ① Vertical Dependency visualizes how reliance on different information sources varies with altitude. ② Overall Contribution quantifies the holistic importance of each feature group (Profile, Global, Spatiotemporal, Height) to the final reconstruction.

3.5.3 Feature Sensitivity and Regional Drivers

We assess robustness and diagnose drivers using two perturbation-based methods. ① Permutation Feature Importance: Measures global sensitivity by randomly shuffling feature groups. Importance is defined as the percentage increase in MSE relative to the baseline, identifying fundamental predictors for domain-wide AEC bias correction (detailed in Sect. S7d). ② SHAP Analysis: SHAP is used to dissect regional heterogeneity. We employ a K-means clustering strategy to construct a representative background dataset capturing diverse atmospheric states (detailed in Sect. S8). SHAP values are computed for specific regions to reveal how dominant AEC bias drivers shift under different environmental regimes.

4. Results and Discussion

https://doi.org/10.5194/egusphere-2026-397
Preprint. Discussion started: 17 February 2026

4.1 Evaluation of the Transformer Model

4.1.1 Overall Predictive Performance and Temporal Generalization

To quantitatively assess the model's robustness in capturing the non-linear mapping between GEOS-Chem simulation biases and atmospheric states, we execute a "Leave-One-Year-Out" cross-validation strategy comprising three independent experiments (Table 1). The Transformer achieves a high average R of 0.659 and a low MAE of 0.014 km^{-1} on the independent test years (Table 2). These metrics demonstrate robust predictive skill, confirming that the model has successfully learned to reproduce the systematic component of extinction biases from the input state variables. Notably, a comparison with the internal validation results (Table S2) reveals that the model's performance on the unseen test sets is comparable to—and in some metrics marginally superior to—that on the validation sets. This consistency suggests that the Transformer architecture has extracted time-invariant, physically meaningful representations of aerosol bias mechanisms rather than overfitting to specific temporal anomalies in the training data. The ability to generalize to years with distinct meteorological interactions underscores the model's potential for operational bias correction.

4.1.2 Seasonal Stability and Robustness

We further examine the temporal stability of the model by analyzing monthly variations in the predictive accuracy of the AEC simulation bias (Fig. 3). The model exhibits a distinct seasonal pattern characterized by superior performance in winter and a moderate decline in summer. During the winter months (December–February), R consistently peaks above 0.7. This enhanced performance is attributed to the synergistic effect of favorable meteorology and observational quality. Specifically, the lower solar elevation angle in winter minimizes solar background noise, thereby enhancing the Signal-to-Noise Ratio (SNR) of the CALIOP retrievals compared to the strong background illumination present in summer (Zhen et al., 2024). Additionally, the stable boundary layer in winter confines aerosols to lower altitudes (Xiong et al., 2025), creating sharper vertical gradients that are physically more distinct for the network to

https://doi.org/10.5194/egusphere-2026-397
Preprint. Discussion started: 17 February 2026

capture. In contrast, a discernible decline in performance occurs during the summer (June–August) over EA. This reduction implies a compound mechanism driven by both data scarcity and observational uncertainty. First, the decline coincides with a sharp decrease in the effective sample size (Fig. 3a, gray bars). This is mechanically linked to the Asian Summer Monsoon, where frequent cloud cover necessitates the exclusion of a significant volume of cloud-contaminated CALIOP profiles (Winker et al., 2009; Winker et al., 2013; Vernier et al., 2011), thereby reducing the representativeness of training data for complex convective scenarios. Second, and more critically, the inherently lower SNR in summer observations imposes a theoretical ceiling on point-to-point correlation metrics. Since random noise in the validation target (CALIOP) cannot be physically predicted, it naturally degrades the R, even if the model correctly retrieves the underlying aerosol signal.

Crucially, despite the fluctuations in linear correlation driven by these external sampling and observational constraints, the model demonstrates remarkable stability in correcting systematic biases. Figure 3c illustrates that the monthly mean bias remains tightly constrained within ± 0.01 km^{-1} throughout the year, exhibiting negligible seasonal drift even during the challenging summer months. This decoupling of metrics implies that while random noise (reflected in lower R) increases in summer due to complex meteorology and reduced SNR, the model does not introduce systematic artifacts. This conclusion is further supported by internal validation results (Fig. S2), which confirms that performance fluctuations are a response to data quality rather than intrinsic model instability. Moreover, detailed monthly density scatter plots (Figs. S3–S5) visually corroborate this robustness, revealing that the majority of predictions remain tightly clustered around the 1:1 identity line, independent of the season.

4.1.3 Vertical Profile Fidelity and Height-Dependent Performance

Given the critical importance of vertical stratification in radiative forcing calculations, we further scrutinize the model's performance as a function of altitude. A distinct advantage of this study, unlike traditional bias correction methods limited to

https://doi.org/10.5194/egusphere-2026-397
Preprint. Discussion started: 17 February 2026

column-integrated parameters, lies in its capability to finely resolve vertical aerosol structures. As evidenced by the vertical profiles of evaluation metrics (Fig. 4a), the model achieves peak performance within the bulk of the PBL (0.5–1.5 km), where R consistently exceeds 0.7. This altitude range corresponds to the region with the heaviest aerosol loading and the most complex chemical composition (Xiong et al., 2025; Jiang et al., 2024). This superior skill suggests that the Transformer's self-attention mechanism effectively captures the steep vertical gradients and pollutant accumulation patterns driven by boundary layer dynamics. Furthermore, the NRMSE (Fig. 4b) remains suppressed below 5% throughout the entire 0–6 km column. This low error magnitude indicates that the model maintains consistent relative predictive accuracy even in the cleaner free troposphere, avoiding the generation of spurious artifacts often seen in DL applications on sparse data. Finally, the vertical profile of mean bias (Fig. 4c) fluctuates narrowly around zero at all altitude levels. This confirms the model's low systematic bias in the vertical dimension, ensuring that the correction process mitigates existing simulation errors without introducing new artificial biases.

4.1.4 Spatial Generalization and Bias Reproduction

Beyond capturing vertical structures, the capacity to resolve the spatial heterogeneity of systematic biases is vital for correcting 3D aerosol fields. In the primary study domain (EA), the Transformer demonstrates high fidelity in reproducing the complex spatial bias modes of the original GEOS-Chem simulation (Fig. 5, columns 1–3). Specifically, the model accurately captures the systematic underestimation over major anthropogenic and biomass burning source regions—including the NCP, IGP, and Indochina Peninsula—as well as over natural dust sources like Taklamakan Desert. Conversely, it correctly identifies regions of systematic overestimation, predominantly located over the remote Western Pacific Ocean, the high-altitude Tibetan Plateau, and high-latitude terrestrial regions (e.g., Siberia). These capabilities suggest that the model effectively differentiates between bias regimes associated with distinct environments: it mitigates the systematic overestimation in clean background regions while

https://doi.org/10.5194/egusphere-2026-397
Preprint. Discussion started: 17 February 2026

concurrently compensating for the underestimation of source intensities in high-loading regions.

However, a localized area of strong positive AEC simulation bias (GEOS-Chem overestimation) appears in Central China in the 2019 target map (Fig. 5c), which is not fully reproduced by the prediction (Fig. 5g). We attribute this discrepancy primarily to observational sparsity and the episodic nature of the bias. Specifically, this region corresponds to a significantly lower density of valid CALIOP samples compared to the surrounding domain (Fig. S6), likely resulting from retrieval limitations associated with complex terrain and frequent cloud cover. Furthermore, monthly decomposition reveals that this elevated annual mean AEC simulation bias is predominantly driven by extreme values in January (Fig. S7), representing a transient winter episode specific to the 2019 test year. In such data-sparse regimes, the physics-informed Transformer model yields conservative predictions, suggesting that it prioritizes learning generalized physical laws over overfitting to localized outliers or specific interannual anomalies under-represented in the training distribution.

Crucially, the fourth column of Figure 5 (2018 NA) presents a rigorous "out-of-domain" generalization test, where the model trained exclusively on EA data is directly applied to NA. Despite distinct differences in emission inventories and meteorological backgrounds between the two continents, the model exhibits remarkable spatial transferability. It successfully predicts the systematic underestimation over the Eastern United States and the overestimation over coastal areas (e.g., the Gulf of Mexico and the Atlantic coast), mirroring the actual GEOS-Chem AEC bias patterns (Target). This successful spatial extrapolation strongly suggests that the physics-informed Transformer model has learned the universal physical mapping between comprehensive atmospheric state variables (detailed in Table S1) and CTM simulation biases, rather than merely overfitting to the geographical coordinates or specific emission patterns of the EA training domain.

4.2 Evaluation of Corrected GEOS-Chem Simulations

https://doi.org/10.5194/egusphere-2026-397
Preprint. Discussion started: 17 February 2026

4.2.1 Correction of Systematic Biases in AEC Magnitude

To quantify the efficacy of the GEOS-Chem corrected by Transformer (GC-TF) model in correcting the magnitude of the AEC, we first compare the overall statistical relationship between simulated values and CALIOP observations across three independent test years from 2017 to 2019 (Fig. 6). The original GEOS-Chem simulation AEC exhibits a dispersed distribution around the 1:1 identity line, with R ranging only from 0.50 to 0.53 and RMSE remaining high at 0.052–0.055 km^{-1}. Notably, the low linear regression slopes (0.46–0.51) of the original simulation indicate a tendency to underestimate aerosol extinction intensity under high-loading conditions. In contrast, after correction by the GC-TF model, the AEC data points converge significantly toward the 1:1 line. The R for AEC improves to 0.66–0.73, the RMSE decreases by approximately 25% (to 0.039–0.042 km^{-1}), and the regression slope recovers to 0.55–0.60. These results demonstrate that the framework effectively reduces random biases and realigns the dynamic range of simulated extinction with observations.

It is acknowledged that the high proportion of clean background samples (e.g., in the upper troposphere) contributes to the overall correlation metrics. To rigorously assess the model's capability in capturing effective aerosol signals—rather than merely fitting the zero-value baseline—we conduct a threshold-based sensitivity analysis (Fig. S8). As the extinction threshold increases from 0.00 to 0.20 km^{-1}, effectively filtering out background noise and isolating optically thick aerosol layers, the GC-TF model consistently outperforms the original GEOS-Chem simulation across all three independent test years. Although the R values naturally decline as the sample size shrinks to focus exclusively on extreme pollution events (indicated by the declining gray dashed line in Fig. S8), the corrected results maintain a persistent performance advantage over the original simulation. This confirms that the framework effectively rectifies structural biases in high-AEC regimes and that its performance gains are not merely artifacts of correctly predicting clean background states.

Further analysis of the vertical structure reveals more complex characteristics of

https://doi.org/10.5194/egusphere-2026-397
Preprint. Discussion started: 17 February 2026

the model AEC bias. The annual mean vertical extinction profiles for the three test years (Fig. 7) reveal a phenomenon: although the low slope in the scatter plots implies an "underestimation" of strong signals, the annual mean profiles reveal that the original GEOS-Chem exhibits a systematic "overestimation" relative to CALIOP observations within the boundary layer (<2 km). This apparent contradiction between macroscopic statistical metrics and the vertical mean state actually exposes the typical "excessive diffusion" issue in CTM simulations: the model struggles to capture the peaks of extreme pollution events (leading to regression slopes<1) while systematically overestimating widespread background aerosol concentrations (resulting in higher intercepts and a systematic overestimation of the mean profile). The GC-TF model successfully addresses this by performing a bidirectional correction: mitigating the systemic overestimation in background regions while recovering the high-loading signals attenuated by model diffusion. In particular, the corrected results neither introduce spurious artifacts nor result in over-smoothing, accurately preserving vertical variation trends consistent with observations even in the free troposphere where aerosol loading is low.

It is important to note that while the absolute magnitude of the residual error (i.e., the remaining bias of the corrected AEC relative to observations) remains highest in the near-surface layer (0–1 km) due to the significantly higher aerosol base loading in the PBL (Fig. 7), the GC-TF model demonstrates its most critical contribution in this 0–1 km layer. It effectively bridges the gap between original simulations and observations, reducing the mean AEC simulation bias by 33%–95% across the independent test years (Fig. 7). Seasonal analysis (Fig. S9) further confirms that the model robustly reduces simulation biases in the near-surface layer, regardless of the season. This substantial reduction in near-surface bias is particularly vital for accurately deriving surface $PM_{2.5}$ concentrations and assessing aerosol health impacts.

The model's capacity to capture spatial heterogeneity is further validated through regional analysis (Figs. S10–12) and explicit vertical bias profiles (Fig. S13). Rather

https://doi.org/10.5194/egusphere-2026-397
Preprint. Discussion started: 17 February 2026

than merely learning a globally uniform bias factor, the GC-TF model exhibits significant state-dependent adaptability. In the anthropogenic-dominated NCP and IGP, as well as the dust-dominated Taklamakan Desert, where the original model shows significant underestimation (Figs. S10–12a, b, c), the GC-TF model successfully enhances the AEC to match observations, effectively pulling the negative bias profiles back toward the zero-reference line (Fig. S13a–c). Additionally, in the Indochina Peninsula, GEOS-Chem exhibits a spurious extinction peak near 0.8 km (Fig. S10d) likely due to mischaracterized injection heights of biomass burning smoke. The GC-TF model significantly attenuates this spurious peak. Conversely, over the relatively clean Western Pacific, the model effectively reduces simulated values to address overestimation (Fig. S10f). This ability to adaptively adjust the correction direction—enhancing in polluted regions while suppressing in clean marine environments—confirms the model's sensitivity to diverse underlying surfaces and emission regimes.

4.2.2 Restoration of Zonal Mean and Longitudinal Vertical Structures

To further evaluate the model's fidelity in resolving the macroscopic spatial-vertical structure of AEC, we analyze the zonal mean vertical distribution of AEC over EA (70°E–150°E) for the 2019 test year (Fig. 8). CALIOP observations (Fig. 8a) identify a prominent aerosol high-loading belt concentrated between 20°N and 30°N, corresponding to major anthropogenic sources in South and East Asia. This aerosol layer is predominantly confined to the lower troposphere below 2 km, with a high-extinction core concentrated within the lowest 1 km. In contrast, the original GEOS-Chem simulation (Fig. 8b) exhibits a characteristic "excessive diffusion" bias: the high-extinction layer is vertically over-extended (reaching above 3 km) and meridionally spread into the clean tropical regions south of 10°N, resulting in a southward displacement of the pollution center. The GC-TF model successfully rectifies these biases by re-centering the high-concentration core to the observed 25°N latitude and effectively constraining the vertical extent of the aerosol layer. By constraining the vertical extent of aerosols, the model reduces the spurious diffusion into the free

https://doi.org/10.5194/egusphere-2026-397
Preprint. Discussion started: 17 February 2026

troposphere and restores the peak extinction intensity suppressed by model smoothing.

Figure 9 further illustrates the annual mean longitudinal vertical cross-sections of AEC along two key latitudinal transects (38.0°N and 40.0°N). These transects capture the transition from continental dust sources and anthropogenic centers to downwind marine regions. (1) Along the 38.0°N transect: CALIOP data (Fig. 9a) reveal two distinct high-extinction cores: the Taklamakan Desert (75°E–85°E) and the NCP (115°E–120°E). The original GEOS-Chem model (Fig. 9c) almost completely fails to capture the intense near-surface dust accumulation in the Taklamakan region—likely due to uncertainties in dust emission schemes or terrain smoothing effects in the CTM—and severely underestimates the core intensity over the NCP. The GC-TF model (Fig. 9e) successfully recovers the missing dust signal and sharpens the anthropogenic core over the NCP, restoring high extinction values (>0.16 km^{-1}) within the 0–1 km layer to match observations. (2) Along the 40.0°N transect: This profile highlights model performance over Northeast Asia and the Sea of Japan. CALIOP detects a concentrated high-extinction core over the North Korea (~127°E). While the original model (Fig. 9d) significantly underestimates this peak, the GC-TF (Fig. 9f) accurately reconstructs this local anthropogenic hotspot. Notably, the original GEOS-Chem exhibits unphysical aerosol "blobs" over the Sea of Japan (135°E–138°E) and near 148°E, which are unsupported by observations. The GC-TF model effectively suppresses these model artifacts, ensuring that the extinction distribution in the downwind marine regions maintains physical consistency with CALIOP observations.

4.2.3 Correction of AOD and Assessment of Spatial Generalization

Figure 10 presents a comprehensive evaluation of the GC-TF model's performance in correcting column-integrated AOD biases over both the primary study domain (EA) and an independent generalization domain (NA). By comparing CALIOP observations, original GEOS-Chem simulations, and the GC-TF model results, we elucidate the model's efficacy in reproducing spatial heterogeneity and improving statistical consistency.

https://doi.org/10.5194/egusphere-2026-397
Preprint. Discussion started: 17 February 2026

Within the EA training domain, although the original GEOS-Chem simulation (Figs. 10a2, b2, c2) captures the macroscopic features of aerosol distribution, it exhibits significant systematic biases. Specifically, it tends to underestimate AOD intensity over major anthropogenic source regions (the NCP and the IGP) and biomass burning hotspots (the Indochina Peninsula), while simultaneously introducing spurious background aerosols over the cleaner Tibetan Plateau and Western Pacific. In contrast, the GC-TF model (Figs. 10a3, b3, c3) significantly sharpens spatial gradients. The corrected AOD fields show high agreement with CALIOP observations (Figs. 10a1, b1, c1), effectively recovering high-value centers in polluted regions while suppressing false positives in clean areas. Statistical evaluation further confirms this improvement: while the original simulation shows dispersed scatter plots with low slopes (0.65–0.74), the GC-TF corrected data tightly converge onto the 1:1 identity line. The R improves from 0.80–0.84 to 0.91–0.93, the RMSE decreases by approximately 26%–40%, and the regression slope recovers to 0.82–0.99 (Figs. 10a5, b5, c5). The temporal consistency observed from 2017 to 2019 indicates that the model has learned stable physical mapping relationships rather than overfitting to specific meteorological years.

Crucially, the evaluation over the NA domain (Fig. 10d) provides compelling evidence of the model's spatial generalization capability. Despite the significantly lower AOD magnitude and distinct emission characteristics compared to EA, the GC-TF model demonstrates robust transferability. The original GEOS-Chem simulation over NA (Figs. 10d2, d4) shows poor agreement with observations (R=0.31, slope=0.18), indicating severe deficiencies in capturing regional aerosol variability. Specifically, it underestimates anthropogenic pollution sources in the Eastern US (Fig. 10d1) and exhibits spurious high-AOD trails over the Gulf of Mexico and the East Coast, likely due to excessive transport or overestimated sea salt. Applying the GC-TF model—trained exclusively on EA data—to this unseen region (Figs. 10d3, d5) yields a substantial performance leap: R more than doubles to 0.70, and the slope improves to 0.45. Spatially, the model successfully recovers the smoothed pollution peaks in the

https://doi.org/10.5194/egusphere-2026-397
Preprint. Discussion started: 17 February 2026

Eastern US and corrects the oceanic regions to clean background levels, consistent with CALIOP. This bidirectional correction capability—enhancing underestimated terrestrial signals while suppressing overestimated oceanic backgrounds—strongly suggests that the physics-informed Transformer framework has captured the universal physical linkages between atmospheric states and simulation biases, rather than merely memorizing the geographical features of the EA training domain.

4.3 Mechanisms of Cross-Continental Transferability

The robust generalization capability demonstrated in the NA domain is fundamentally underpinned by the universality of atmospheric physical laws. While EA and NA differ vastly in emission intensities and aerosol composition regimes, the fundamental physicochemical processes governing aerosol lifecycles (e.g., hygroscopic growth, turbulent mixing, and wet deposition) adhere to universal physical laws. The robust performance in the NA domain (R=0.70) indicates that the systematic biases in GEOS-Chem are predominantly state-dependent rather than location-dependent. Specifically, CTMs tend to exhibit consistent "error signatures" under specific meteorological conditions. When the GC-TF model encounters similar synoptic patterns in NA, the self-attention mechanism identifies these familiar meteorological contexts and applies the learned correction rules. In this sense, the meteorological background acts as a bridge, facilitating the transfer of bias-correction capabilities across continents.

The model's capability is further highlighted by its bidirectional correction performance over distinct underlying surfaces. Over the Eastern United States, the model successfully mitigates the "smearing effect" characteristic of Eulerian models (Rastigejev et al., 2010; Eastham and Jacob, 2017), restoring the sharp spatial gradients and high-intensity cores of anthropogenic pollution that were overly smoothed in the original simulation. Conversely, over marine regions such as the Gulf of Mexico and the Atlantic coast, the model effectively suppresses the spurious background aerosols often caused by excessive transport or sea-salt overestimation. This adaptive

https://doi.org/10.5194/egusphere-2026-397
Preprint. Discussion started: 17 February 2026

behavior—enhancing peaks over land while suppressing artifacts over oceans—confirms that the model utilizes multi-dimensional atmospheric variables to constrain the spatial distribution of aerosols physically, rather than applying a uniform bias correction factor. It is noteworthy, however, that the correction performance in NA (R=0.70) remains statistically lower than in the EA training domain (R=0.93). This performance gap is likely attributable to the "domain shift" in aerosol composition. The NA atmosphere is characterized by lower background concentrations and a higher fraction of biogenic secondary organic aerosols (Goldstein et al., 2009), whose optical properties and hygroscopicity differ from the sulfate-nitrate-dust dominated mixtures in EA (Crawford et al., 2021). This suggests that while the physical laws of transport are universal, chemical-specific optical relationships may require broader representativeness in the training data for further refinement.

4.4 Case Studies: Vertical Structure and Temporal Evolution During High Aerosol Loading Periods

To thoroughly evaluate the GC-TF model's capability in resolving complex vertical AEC stratification and capturing high-frequency pollution dynamics, we conduct a composite analysis of representative high-loading episodes identified by CALIOP over three distinct aerosol regimes (details in Section 2.3).

In the Taklimakan Desert (Fig. 11a), the original GEOS-Chem simulation exhibits a nearly vertical extinction profile with extremely low values, indicating that the CTM fails to simulate the significant aerosol extinction layer observed by CALIOP. In contrast, the GC-TF model effectively reconstructs the extinction signal in this region by leveraging multidimensional input atmospheric state variables (including meteorological fields and background chemical composition). Its profile shape shows superior agreement with satellite observations, largely retrieving the "missing signal" of the original simulation. In Kanpur, a region dominated by anthropogenic pollution (Fig. 11b), the original model fails to accurately capture the magnitude and vertical gradient of the pollution layer below 3.5 km. The GC-TF model, while correcting

https://doi.org/10.5194/egusphere-2026-397
Preprint. Discussion started: 17 February 2026

baseline errors, reshapes the vertical distribution pattern, yielding a vertical decrement trend that aligns closely with observational reality. Figure 11c reveals critical structural biases during a biomass burning event in the Indochina Peninsula. The original GEOS-Chem simulation presents an "inverted" vertical gradient, where the extinction peak appears at a height of 2.5 km rather than near the surface. This bias is likely associated with uncertainties in the parameterization of biomass burning Plume Injection Height in the CTM, suggesting the model erroneously transports surface smoke into the free troposphere (Paugam et al., 2016). The GC-TF model successfully identifies this physical process bias; by significantly enhancing AEC below 2.0 km, it restores a near-surface high-loading pattern consistent with observations, while maintaining reasonable simulation magnitudes in the free troposphere.

In addition to the instantaneous vertical structure, verifying the temporal continuity of the correction results is equally crucial. Given the spatiotemporal sparsity of CALIOP observations, we utilize ground-based high-frequency AERONET AOD data to further evaluate model performance. Figure 11d shows that during the study period in Kanpur, the original GEOS-Chem simulation generally underestimates aerosol loading, and the curve exhibits overly smooth characteristics, lacking response to high-frequency fluctuations. The AOD time series output by the GC-TF model demonstrates superior trend consistency with AERONET observations. Notably, between May 1 and May 2, the original simulation shows significant overestimation, whereas the GC-TF model successfully adjusts the predicted values back to levels closer to observations. During the active fire period in Nong Khai (Fig. 11e), the original model severely underestimates the AOD magnitude. The GC-TF model significantly elevates the simulation baseline and captures the phase of diurnal variation trends well (e.g., the fluctuations around 07:00 UTC during February 22–24). However, its predicted peak amplitudes remain significantly lower than the extreme values observed by AERONET. This dampening of extreme peaks likely stems from the spatial scale mismatch between the gridded model (2°×2.5°) and ground-based point

https://doi.org/10.5194/egusphere-2026-397
Preprint. Discussion started: 17 February 2026

observations, which causes local extreme emissions to be averaged out within the grid. Nonetheless, the GC-TF model demonstrates significant improvement over the original CTM in reproducing meteorology-driven pollution accumulation and dispersion processes.

4.5 Interpretability Analysis

To move beyond statistical performance and elucidate the internal decision-making logic of the GC-TF model, we employ a multi-perspective interpretability framework to quantify feature contributions and interaction mechanisms. This section and the subsequent analysis are based entirely on the independent test dataset, ensuring that the interpretations reflect the model's generalized physical representations on unseen data.

4.5.1 Macro-Physical Logic: Physicochemical Baseline and Contextual Correction

To understand how the DL model reconstructs 3D aerosol fields, we first examine the macroscopic contribution of different information sources. Global sensitivity tests (Fig. S14) unambiguously establish the foundational role of physical priors: randomly permuting the physicochemical profile features induces a dramatic 196.1% surge in MSE. This confirms that while the Transformer architecture excels at extracting spatiotemporal patterns, the physicochemical state features provided by the CTM remain the indispensable physical foundation, essentially determining the "Base Magnitude" of the AEC. In this framework, spatial, temporal, and height features function as "Correction Factors," refining this baseline by integrating multidimensional environmental contexts.

Further analysis of the gated fusion weights (Fig. 12) reveals that the model effectively learns to differentiate its correction strategy based on the dominant physical processes at different atmospheric stratifications: (1) In the near-surface layer (<0.5 km), the model assigns the highest weight to profile features. This aligns with the physical reality that extinction near the surface is predominantly controlled by local instantaneous states—specifically emission intensities and hygroscopic growth (Jiang

https://doi.org/10.5194/egusphere-2026-397
Preprint. Discussion started: 17 February 2026

et al., 2024; Xiong et al., 2025). The model accurately identifies that these physicochemical variables are the most critical determinants of extinction magnitude in the lowest layer. (2) As height increases to the boundary layer transition zone (0.5–1.0 km), a strategic shift occurs. The reliance on raw simulation profiles diminishes, while the weight of spatial features increases significantly. This region typically corresponds to the entrainment zone or the top of the nocturnal boundary layer, where CTMs are prone to vertical diffusion errors (Eastham and Jacob, 2017; Rastigejev et al., 2010; Lin and Mcelroy, 2010). The model appears to mitigate these uncertainties by leveraging spatial priors to constrain potential diffusion biases, effectively identifying "regions of uncertainty" in the simulation. (3) In the mid-lower troposphere (1.0–3.5 km), the contribution of temporal features exhibits distinct peaks, particularly around 1.5 km and 2.8 km. The peak near 1.5 km corresponds to the typical maximum daytime PBLH in EA (Guo et al., 2016; Kim, 2022), while the signal at 2.8 km aligns with the active layer for long-range transport (Uno et al., 2009). This suggests the model utilizes temporal cues to capture the diurnal evolution of the boundary layer and the seasonality of trans-regional transport events. (4) Finally, in the free troposphere (>3.5 km), the dependence on local profile features fades, and spatial features become the dominant driving factor. In this region, aerosol variability is decoupled from local surface processes and driven primarily by large-scale advection (Weinzierl et al., 2017; Val Martin et al., 2013; Uno et al., 2009). The model successfully captures this mechanistic decoupling, shifting its strategy to perform position-dependent background correction using geographical coordinates. Throughout the entire column, explicit height encoding maintains a consistently high contribution, serving as a critical vertical positioning anchor.

4.5.2 Micro-Drivers of Bias Correction: Vertical Attribution of State Variables

Following the macroscopic establishment of feature importance, we utilize Integrated Gradients analysis (Fig. 13) to dissect the specific physical variables driving extinction bias correction across different altitudes. The analysis reveals that the model

https://doi.org/10.5194/egusphere-2026-397
Preprint. Discussion started: 17 February 2026

has constructed a sophisticated decision-making logic that effectively decouples thermodynamic constraints from compositional variables.

First, thermodynamic variables serve as the primary constraints for structural correction. Temperature (T) consistently acts as the dominant driver within the PBL, with attribution scores exceeding 0.20. This heavy reliance suggests that the model implicitly diagnoses atmospheric stability and the vertical lapse rate—key factors often mischaracterized in CTMs—to target biases associated with turbulent mixing intensities (Lin and Mcelroy, 2010). Acting in concert with temperature, Relative Humidity (RH) functions as a critical optical modulator. Although its absolute attribution score is lower than that of temperature, its contribution is coupled with explicit hygroscopic growth factors (e.g., $\mathrm{AerHygroscopicGrowth\_SO_4^{2-}}$). This enables the Transformer to fine-tune the optical AEC, effectively correcting biases arising from non-linear hygroscopic parameterization errors (Burgos et al., 2020; Zhai et al., 2021), particularly under high-humidity conditions.

Second, particulate mass concentrations act as the fundamental anchors for pollution intensity. In the lower troposphere (<1.5 km), $PM_{2.5}$ and PM_{10} consistently rank among the top drivers, confirming their role in determining the baseline magnitude of AEC. The model also demonstrates a physically stratified recognition of aerosol types; for instance, the importance of sea salt aerosol (AerMassSAL) is confined strictly to the marine boundary layer (<1.5 km) and decays rapidly aloft, accurately reflecting the vertical distribution of coarse-mode marine aerosols (Bian et al., 2019; Murphy et al., 2019).

Finally, dynamic variables exhibit a distinct "long-tail effect" in the vertical dimension. Unlike precursor concentrations which decay sharply with height, the importance of wind components (U, V) remains relatively stable in the free troposphere (>2.0 km). This indicates that as height increases, the model shifts its focus from local accumulation to large-scale advection, utilizing wind fields to identify and correct background biases induced by long-range transport.

https://doi.org/10.5194/egusphere-2026-397
Preprint. Discussion started: 17 February 2026

4.5.3 Modulation by Environmental Context: Synoptic and Surface Constraints

While vertical profiles determine the baseline extinction, the Cross-Attention mechanism enables the model to utilize 2D global variables as "Environmental Constraints" to modulate the vertical reconstruction (Fig. 14). This process operates through two distinct physical pathways: dynamic ventilation and radiative-surface forcing.

First, the model employs a sophisticated dynamic diagnosis mechanism to constrain transport and mixing efficiency. The consistently high attention weights of 10 m wind components (U10M, V10M) throughout the column reflect the model's use of near-surface wind speed as a proxy for synoptic flow patterns. Notably, the model identifies the dominance of meridional transport in the EA monsoon region, assigning slightly higher importance to meridional winds (V10M) to capture dominant pollutant exchange pathways (Ding et al., 2009; Choi et al., 2024; Uno et al., 2009). By identifying prevailing dynamic regimes, the model effectively addresses common CTM biases related to pollutant accumulation under stagnant conditions (Kim et al., 2024; An et al., 2019) and numerical diffusion under strong advection (Rastigejev et al., 2010; Eastham and Jacob, 2017). More critically, the cross-attention weights reveal a deep decoupling of boundary layer turbulence processes. Thermodynamic drivers (HFLUX, LAI) exhibit a "surface-peak, upward-decay" pattern, suggesting the model uses them to determine the initial buoyancy potential. In sharp contrast, mechanical drivers (USTAR, Z0M) display an "increase-and-plateau" trend with height. This distinction implies that the model successfully diagnoses the "vertical permeability" of the atmosphere—assessing whether mechanical shear is sufficient to pump pollutants through capping inversions into the free troposphere.

Second, radiative and surface boundary conditions are leveraged to correct biases in chemical formation and stability parameterization. The high attention assigned to photosynthetically active radiation—distinguishing between direct (PARDR) and diffuse (PARDF) components—indicates the model's implicit capture of photochemical

https://doi.org/10.5194/egusphere-2026-397
Preprint. Discussion started: 17 February 2026

variation. By sensing solar input, the model dynamically compensates for biases related to secondary aerosol formation rates or biogenic emissions (Guenther et al., 2012). Furthermore, snow mass (SNOMAS) emerges as a key predictor in the lowest layers. The model identifies snow-covered surfaces as a proxy for "suppressed diffusion states", characterizing conditions prone to deep near-surface inversions. This allows for the targeted correction of the "over-dilution" bias often found in Eulerian models (Lin and Mcelroy, 2010; Holtslag et al., 2013), effectively restoring high-concentration signals in stable boundary layers that are typically smoothed out by minimal diffusion threshold constraints.

4.5.4 Regional Heterogeneity: Spatially Distinct Physical Fingerprints

To reveal the model's decision-making logic under diverse environmental contexts, we conducted a "physical fingerprint" analysis (Fig. 15) based on SHAP importance rankings. The results demonstrate that the model autonomously establishes three distinct physical response modes adapted to regional underlying surfaces and emission characteristics.

In high-emission anthropogenic and biomass burning regions (NCP, IGP, and Indochina), the model primarily leverages radiative components for optical identification. A striking commonality across these regions is the high ranking of diffuse photosynthetically active radiation (PARDF), often surpassing the direct component (PARDR). This reflects the model's capability to capture the physical phenomenon where high aerosol loading enhances scattering and increases the diffuse fraction of solar radiation (Mercado et al., 2009; Che et al., 2018), utilizing it as a robust proxy for optical turbidity. Furthermore, the model accurately differentiates surface energy partitioning regimes: in the humid, vegetated IGP and Indochina, it prioritizes latent heat flux (EFLUX) to gauge hygroscopic growth and wet removal potential; conversely, in the urbanized NCP, it relies more heavily on sensible heat flux (HFLUX), consistent with the high Bowen ratio of urban surfaces (Miao et al., 2009) where thermal turbulence dominates vertical dispersion.

https://doi.org/10.5194/egusphere-2026-397
Preprint. Discussion started: 17 February 2026

In the dust-dominated Taklamakan Desert, the model reconstructs a closed-loop mechanism driven by "Thermodynamic Forcing and Dynamic Uplift". Incident shortwave flux (SWGDN) and direct radiation (PARDR) play an overwhelmingly dominant role, indicating that the model identifies clear-sky, high-solar-input conditions as the prerequisite thermodynamic background for inducing thermal instability. Crucially, this thermodynamic setup is coupled with dynamic and surface descriptors: the model combines low vegetation indices (GRN, identifying erodible bare soil), high 10m wind speeds (providing surface shear stress) (Shao et al., 2011), and a preference for coarse-mode PM_{10} over $PM_{2.5}$. This combination confirms that the Transformer has successfully learned the physical process of wind-driven dust emission in arid source regions.

Finally, over the marine environment of the Western Pacific, the model shifts to a "Latent Heat-Driven" mode. Latent heat flux (EFLUX) contributes nearly half of the total feature importance, defining the moisture supply at the air-sea interface that controls marine aerosol hygroscopicity. Additionally, the model captures the mechanics of sea spray generation by linking near-surface wind speeds (V10M) with land-sea masking indicators (GRN). This suggests the model recognizes that under high-wind conditions over water, mechanically generated sea salt aerosols (Grythe et al., 2014)—and their subsequent hygroscopic evolution—are the primary drivers of AEC variability.

4.6 Feedback from Data-Driven Correction to Physical Model Improvement

The interpretability analysis in Section 4.5 demonstrates that the GC-TF model does not merely fit statistical biases but successfully reconstructs a closed-loop physical logic governing aerosol vertical structure. This "physics-informed" transparency offers a unique opportunity to diagnose the root causes of CTM biases and establish a feedback loop for model improvement.

4.6.1 Diagnosing Thermodynamic Parameterization Deficiencies

The model's dominant reliance on temperature and sensible heat flux (Sections 4.5.2, 4.5.4) to correct vertical profiles strongly implies that the original GEOS-Chem

https://doi.org/10.5194/egusphere-2026-397
Preprint. Discussion started: 17 February 2026

simulation suffers from structural deficiencies in diagnosing PBLH and mixing intensity. Specifically, the "excessive diffusion" biases observed in the lower troposphere suggest that the CTM may misinterpret thermodynamic stability over complex underlying surfaces. In highly urbanized regions like the NCP, the model's high sensitivity to sensible heat flux indicates that the current scheme may not fully resolve the distinct energy budget of urban canopies. This suggests that future improvements should prioritize refining stability-dependent diffusion coefficients or implementing more explicit parameterizations for urban surface energy balance to better constrain vertical mixing.

4.6.2 Refining Emission and Formation Schemes via Environmental Proxies

The "context-aware" behavior of the Transformer reveals missing process-level links in the current CTM emission and formation modules. In the Taklamakan Desert, the model explicitly combines GRN (identifying bare soil) and Wind Speed to reconstruct dust extinction. The fact that the DL model requires these specific inputs to correct the simulation implies that the original dust emission scheme likely fails to accurately capture the non-linear threshold between wind shear and erodible surfaces, potentially miscalculating the threshold friction velocity or wind-erosion efficiency. Similarly, in Indochina, the high sensitivity to diffuse radiation (PARDF) suggests that secondary organic aerosol formation associated with biomass burning plumes is under-represented. The model compensates for this by using radiation as a proxy for photochemical aging rates, pointing to a need for optimizing secondary organic aerosol formation mechanisms and injection height parameterizations in the CTM.

4.6.3 The Role of Machine Learning in Earth System Modeling

Beyond bias correction, this study exemplifies a paradigm shift from "black-box" prediction to "glass-box" diagnosis. By decoupling the contributions of meteorology and composition, the physics-informed Transformer serves as a diagnostic tool. It verifies that while CTMs provide a solid physicochemical baseline, they struggle with the non-linear modulation of local environmental contexts. The correction strategies

https://doi.org/10.5194/egusphere-2026-397
Preprint. Discussion started: 17 February 2026

learned by the model—such as using GRN as a dynamic land mask or SNOMAS as a stability indicator—provide traceable scientific evidence for incorporating these neglected environmental constraints into future deterministic parameterization schemes.

5. Conclusions

This study proposes a physics-informed Transformer DL framework that reconstructs high-precision 3D AEC fields by bridging the gap between CTM simulations and CALIOP satellite observations. Distinct from traditional bias correction methods that rely solely on passive remote sensing imagery or treat the model as a "black box," this framework explicitly integrates physicochemical vertical structural priors (e.g., AEPs and species concentrations) from GEOS-Chem and meteorological constraints from MERRA-2. By constructing a dual-stream architecture featuring Gated Feature Fusion and Cross-Attention mechanisms, the model effectively captures the complex, non-linear dependencies between atmospheric states and AEC simulation biases. The primary conclusions are as follows:

First, the framework significantly improves the fidelity of AEC vertical profiles, systematically correcting the systematic biases inherent in Eulerian CTMs. Extensive "Leave-One-Year-Out" cross-validation over EA (2017–2019) demonstrates that the R for AEC evaluated against CALIOP observations increases from 0.49–0.53 in the original GEOS-Chem simulation to 0.66–0.73, while the RMSE is reduced by approximately 25%. Vertically, the model achieves optimal performance within the PBL (0.5–1.5 km), with R consistently exceeding 0.7 and NRMSE remaining below 5%. The framework effectively mitigates the "excessive diffusion" issue common in CTMs, rectifying the spurious southward displacement of pollution centers, restoring smoothed high-intensity extinction signals in anthropogenic and dust source regions, and suppressing spurious overestimations in the clean free troposphere and over marine regions. Crucially, threshold-based sensitivity analysis confirms that these performance gains are robust even when focusing exclusively on optically thick aerosol layers.

https://doi.org/10.5194/egusphere-2026-397
Preprint. Discussion started: 17 February 2026

Consequently, this vertical rectification leads to a substantial improvement in column-integrated AOD, with the R relative to CALIOP increasing from 0.80–0.84 to 0.91–0.93 across the EA domain, ensuring consistency between resolved AEC vertical structures and the total aerosol load. Moreover, validation against high-frequency ground-based AERONET measurements confirms the framework's superiority over the original GEOS-Chem simulation in capturing the dynamic temporal evolution of pollution episodes, bridging the observational gaps inherent to the sparse sampling of polar-orbiting satellites.

Second, the model exhibits robust cross-continental transferability, indicating that it has internalized universal physical laws rather than overfitting to regional training characteristics. When applied directly to the NA domain—a region with distinct emission structures and lower aerosol loading—without any retraining, the model successfully reproduces the spatial AEC simulation bias patterns of GEOS-Chem. The R for column-integrated AOD compared with CALIOP retrievals in this region improves significantly from 0.31 (original) to 0.70 (corrected), further verifying the physical consistency and transferability of the established mapping between atmospheric states and AEC simulation biases.

Third, by integrating interpretable DL techniques, this study establishes a closed-loop feedback pathway from "data-driven correction" to "physical mechanism diagnosis." Attribution analysis reveals that the model autonomously identifies AEC simulation bias drivers with clear physical significance: (1) In the PBL, the model relies primarily on temperature and sensible heat flux to correct vertical mixing biases, pointing to deficiencies in stability-dependent diffusion schemes within the CTM; (2) The synergistic diagnosis of dust emission thresholds using vegetation indices and wind speed suggests potential biases in the non-linear response of existing dust parameterizations to wind erosion mechanisms; (3) In biomass burning regions, the sensitivity to diffuse radiation reveals missing links in secondary organic aerosol formation processes or plume injection height parameterizations; (4) In marine regions,

https://doi.org/10.5194/egusphere-2026-397
Preprint. Discussion started: 17 February 2026

the model utilizes latent heat flux and wind patterns to target and correct potential overestimation biases in sea salt aerosol generation functions or hygroscopic growth parameterizations.

In summary, this study not only confirms the effectiveness of the physics-informed Transformer in reconstructing high-fidelity 3D aerosol fields to support radiative forcing assessments but also highlights its potential as a diagnostic tool for Earth system modeling. The interpretability analysis provides developers with actionable physical insights, highlighting a shift in the integration of deep learning with physical modeling. Rather than serving solely as a post-processing tool for bias correction, the framework functions as a diagnostic mechanism to identify structural model deficiencies and guide the refinement of parameterization schemes in Earth system models.

Code and Data availability. The GEOS-Chem model is available at http://www.geos-chem.org. The CALIOP satellite data can be accessed via https://subset.larc.nasa.gov/calipso. The MERRA-2 data are available from the NASA Goddard Earth Sciences Data and Information Services Center (https://disc.gsfc.nasa.gov/). AERONET data can be obtained from https://aeronet.gsfc.nasa.gov/. The source code of the physics-informed Transformer model proposed in this study and the trained weights are available upon request from the corresponding author.

Author contribution. JX designed the research, developed the model methodology, conducted the simulations, performed the data analysis, and prepared the original draft. YW supervised the study, provided the computing resources, and revised the manuscript. JW revised the manuscript. YYW, MZ, MT, WD, JK, and LW provided technical support and reviewed the manuscript. All authors contributed to the interpretation of the results and the improvement of the paper.

https://doi.org/10.5194/egusphere-2026-397
Preprint. Discussion started: 17 February 2026

Competing Interests. The contact author has declared that none of the authors has any competing interests.

Acknowledgements. This research was funded by the National Natural Science Foundation of China (Grant No. 42571377). Jun Wang's participation is made possible via the in-kind support (Lichtenberger Family Chair in Chemical Engineering) from the University of Iowa.

References

An, Z., Huang, R. J., Zhang, R., Tie, X., Li, G., Cao, J., Zhou, W., Shi, Z., Han, Y., Gu, Z., and Ji, Y.: Severe haze in northern China: A synergy of anthropogenic emissions and atmospheric processes, Proc Natl Acad Sci U S A, 116, 8657-8666, 10.1073/pnas.1900125116, 2019.

Benedetti, A., Reid, J. S., Knippertz, P., Marsham, J. H., Di Giuseppe, F., Rémy, S., Basart, S., Boucher, O., Brooks, I. M., Menut, L., Mona, L., Laj, P., Pappalardo, G., Wiedensohler, A., Baklanov, A., Brooks, M., Colarco, P. R., Cuevas, E., da Silva, A., Escribano, J., Flemming, J., Huneeus, N., Jorba, O., Kazadzis, S., Kinne, S., Popp, T., Quinn, P. K., Sekiyama, T. T., Tanaka, T., and Terradellas, E.: Status and future of numerical atmospheric aerosol prediction with a focus on data requirements, Atmospheric Chemistry and Physics, 18, 10615-10643, 10.5194/acp-18-10615-2018, 2018.

Bey, I., Jacob, D. J., Yantosca, R. M., Logan, J. A., Field, B. D., Fiore, A. M., Li, Q. B., Liu, H. G. Y., Mickley, L. J., and Schultz, M. G.: Global modeling of tropospheric chemistry with assimilated meteorology: Model description and evaluation, Journal of Geophysical Research-Atmospheres, 106, 23073-23095, 10.1029/2001jd000807, 2001.

Bian, H. S., Froyd, K., Murphy, D. M., Dibb, J., Darmenov, A., Chin, M., Colarco, P. R., da Silva, A., Kucsera, T. L., Schill, G., Yu, H. B., Bui, P., Dollner, M., Weinzierl,

https://doi.org/10.5194/egusphere-2026-397
Preprint. Discussion started: 17 February 2026

B., and Smirnov, A.: Observationally constrained analysis of sea salt aerosol in the marine atmosphere, Atmospheric Chemistry and Physics, 19, 10773-10785, 10.5194/acp-19-10773-2019, 2019.

Bocquet, M., Pires, C. A., and Wu, L.: Beyond Gaussian Statistical Modeling in Geophysical Data Assimilation, Monthly Weather Review, 138, 2997-3023, 10.1175/2010mwr3164.1, 2010.

Bond, T. C., Doherty, S. J., Fahey, D. W., Forster, P. M., Berntsen, T., DeAngelo, B. J., Flanner, M. G., Ghan, S., Kärcher, B., Koch, D., Kinne, S., Kondo, Y., Quinn, P. K., Sarofim, M. C., Schultz, M. G., Schulz, M., Venkataraman, C., Zhang, H., Zhang, S., Bellouin, N., Guttikunda, S. K., Hopke, P. K., Jacobson, M. Z., Kaiser, J. W., Klimont, Z., Lohmann, U., Schwarz, J. P., Shindell, D., Storelvmo, T., Warren, S. G., and Zender, C. S.: Bounding the role of black carbon in the climate system: A scientific assessment, Journal of Geophysical Research: Atmospheres, 118, 5380-5552, 10.1002/jgrd.50171, 2013.

Burgos, M. A., Andrews, E., Titos, G., Benedetti, A., Bian, H. S., Buchard, V., Curci, G., Kipling, Z., Kirkevåg, A., Kokkola, H., Laakso, A., Letertre-Danczak, J., Lund, M. T., Matsui, H., Myhre, G., Randles, C., Schulz, M., van Noije, T., Zhang, K., Alados-Arboledas, L., Baltensperger, U., Jefferson, A., Sherman, J., Sun, J. Y., Weingartner, E., and Zieger, P.: A global model-measurement evaluation of particle light scattering coefficients at elevated relative humidity, Atmospheric Chemistry and Physics, 20, 10231-10258, 10.5194/acp-20-10231-2020, 2020.

Che, H. Z., Qi, B., Zhao, H. J., Xia, X. G., Eck, T. F., Goloub, P., Dubovik, O., Estelles, V., Cuevas-Agulló, E., Blarel, L., Wu, Y. F., Zhu, J., Du, R. G., Wang, Y. Q., Wang, H., Gui, K., Yu, J., Zheng, Y., Sun, T. Z., Chen, Q. L., Shi, G. Y., and Zhang, X. Y.: Aerosol optical properties and direct radiative forcing based on measurements from the China Aerosol Remote Sensing Network (CARSNET) in eastern China, Atmospheric Chemistry and Physics, 18, 405-425, 10.5194/acp-18-405-2018, 2018.

https://doi.org/10.5194/egusphere-2026-397
Preprint. Discussion started: 17 February 2026

Chen, B., Song, Z., Pan, F., and Huang, Y.: Obtaining vertical distribution of $PM_{2.5}$ from CALIOP data and machine learning algorithms, Sci Total Environ, 805, 150338, 10.1016/j.scitotenv.2021.150338, 2022a.

Chen, J., Jiang, Z., Li, R., Liao, C., Miyazaki, K., and Jones, D. B. A.: Large discrepancy between observed and modeled wintertime tropospheric NO2 variabilities due to COVID-19 controls in China, Environmental Research Letters, 17, 10.1088/1748-9326/ac4ec0, 2022b.

Chen, X., Wang, J., Xu, X. G., and Zhou, M.: Dust Aerosol Optical Centroid Height (AOCH) Over Bright Surface: First Retrieval From TROPOMI Oxygen A and B Absorption Bands, Ieee Geoscience and Remote Sensing Letters, 22, 10.1109/lgrs.2025.3601046, 2025.

Chen, X., Xu, X. G., Wang, J., and Diner, D. J.: Can multi-angular polarimetric measurements in the oxygen-A and B bands improve the retrieval of aerosol vertical distribution?, Journal of Quantitative Spectroscopy & Radiative Transfer, 270, 10.1016/j.jqsrt.2021.107679, 2021a.

Chen, X., Wang, J., Xu, X., Zhou, M., Zhang, H., Castro Garcia, L., Colarco, P. R., Janz, S. J., Yorks, J., McGill, M., Reid, J. S., de Graaf, M., and Kondragunta, S.: First retrieval of absorbing aerosol height over dark target using TROPOMI oxygen B band: Algorithm development and application for surface particulate matter estimates, Remote Sensing of Environment, 265, 10.1016/j.rse.2021.112674, 2021b.

Chimot, J., Veefkind, J. P., Vlemmix, T., and Levelt, P. F.: Spatial distribution analysis of the OMI aerosol layer height: a pixel-by-pixel comparison to CALIOP observations, Atmospheric Measurement Techniques, 11, 2257-2277, 10.5194/amt-11-2257-2018, 2018.

Chinnam, N., Dey, S., Tripathi, S. N., and Sharma, M.: Dust events in Kanpur, northern India: Chemical evidence for source and implications to radiative forcing, Geophysical Research Letters, 33, 10.1029/2005gl025278, 2006.

https://doi.org/10.5194/egusphere-2026-397
Preprint. Discussion started: 17 February 2026

Choi, M., Park, J., Sung, M., and Ying, Q.: Long-Range Transport of Secondary Inorganic Aerosol from China to South Korea, Environmental Science & Technology Letters, 11, 1233-1238, 10.1021/acs.estlett.4c00852, 2024.

Christensen, M. W., Jones, W. K., and Stier, P.: Aerosols enhance cloud lifetime and brightness along the stratus-to-cumulus transition, Proceedings of the National Academy of Sciences of the United States of America, 117, 17591-17598, 10.1073/pnas.1921231117, 2020.

Christian, K., Wang, J., Ge, C., Peterson, D., Hyer, E., Yorks, J., and McGill, M.: Radiative Forcing and Stratospheric Warming of Pyrocumulonimbus Smoke Aerosols: First Modeling Results With Multisensor (EPIC, CALIPSO, and CATS) Views from Space, Geophysical Research Letters, 46, 10061-10071, 10.1029/2019gl082360, 2019.

Colarco, P. R., Kahn, R. A., Remer, L. A., and Levy, R. C.: Impact of satellite viewing-swath width on global and regional aerosol optical thickness statistics and trends, Atmospheric Measurement Techniques, 7, 2313-2335, 10.5194/amt-7-2313-2014, 2014.

Crawford, J. H., Ahn, J. Y., Al-Saadi, J., Chang, L., Emmons, L. K., Kim, J., Lee, G., Park, J. H., Park, R. J., Woo, J. H., Song, C. K., Hong, J. H., Hong, Y. D., Lefer, B. L., Lee, M., Lee, T., Kim, S., Min, K. E., Yum, S. S., Shin, H. J., Kim, Y. W., Choi, J. S., Park, J. S., Szykman, J. J., Long, R. W., Jordan, C. E., Simpson, I. J., Fried, A., Dibb, J. E., Cho, S., and Kim, Y. P.: The Korea-United States Air Quality (KORUS-AQ) field study, Elementa-Science of the Anthropocene, 9, 10.1525/elementa.2020.00163, 2021.

Daoud, N., Eltahan, M., and Elhennawi, A.: Aerosol Optical Depth Forecast over Global Dust Belt Based on LSTM, CNN-LSTM, CONV-LSTM and FFT Algorithms, 19th International Conference on Smart Technologies (IEEE EUROCON), Lviv, UKRAINE, Jul 06-08 2021, WOS:000728121700034, 186-191, 10.1109/eurocon52738.2021.9535571, 2021.

https://doi.org/10.5194/egusphere-2026-397
Preprint. Discussion started: 17 February 2026

Diner, D. J., Boland, S. W., Brauer, M., Bruegge, C., Burke, K. A., Chipman, R., Di Girolamo, L., Garay, M. J., Hasheminassab, S., Hyer, E., Jerrett, M., Jovanovic, V., Kalashnikova, O. V., Liu, Y., Lyapustin, A. I., Martin, R. V., Nastan, A., Ostro, B. D., Ritz, B., Schwartz, J., Wang, J., and Xu, F.: Advances in multiangle satellite remote sensing of speciated airborne particulate matter and association with adverse health effects: from MISR to MAIA, Journal of Applied Remote Sensing, 12, 10.1117/1.Jrs.12.042603, 2018.

Ding, A. J., Wang, T., Xue, L. K., Gao, J., Stohl, A., Lei, H. C., Jin, D. Z., Ren, Y., Wang, X. Z., Wei, X. L., Qi, Y. B., Liu, J., and Zhang, X. Q.: Transport of north China air pollution by midlatitude cyclones: Case study of aircraft measurements in summer 2007 (vol 114, D08304, 2007), Journal of Geophysical Research-Atmospheres, 114, 10.1029/2009jd012339, 2009.

Ding, S. G., Wang, J., and Xu, X. G.: Polarimetric remote sensing in oxygen A and B bands: sensitivity study and information content analysis for vertical profile of aerosols, Atmospheric Measurement Techniques, 9, 2077-2092, 10.5194/amt-9-2077-2016, 2016.

Dong, W., Tao, M., Xu, X., Wang, J., Wang, Y., Wang, L., Song, Y., Fan, M., and Chen, L.: Satellite Aerosol Retrieval From Multiangle Polarimetric Measurements: Information Content and Uncertainty Analysis, IEEE Transactions on Geoscience and Remote Sensing, 61, 1-13, 10.1109/tgrs.2023.3264554, 2023.

Du, Q., Zhao, C., Zhang, M., Dong, X., Chen, Y., Liu, Z., Hu, Z., Zhang, Q., Li, Y., Yuan, R., and Miao, S.: Modeling diurnal variation of surface $PM_{2.5}$ concentrations over East China with WRF-Chem: impacts from boundary-layer mixing and anthropogenic emission, Atmospheric Chemistry and Physics, 20, 2839-2863, 10.5194/acp-20-2839-2020, 2020.

Dubovik, O., Herman, M., Holdak, A., Lapyonok, T., Tanré, D., Deuzé, J. L., Ducos, F., Sinyuk, A., and Lopatin, A.: Statistically optimized inversion algorithm for enhanced retrieval of aerosol properties from spectral multi-angle polarimetric

https://doi.org/10.5194/egusphere-2026-397
Preprint. Discussion started: 17 February 2026

satellite observations, Atmospheric Measurement Techniques, 4, 975-1018, 10.5194/amt-4-975-2011, 2011.

Eastham, S. D. and Jacob, D. J.: Limits on the ability of global Eulerian models to resolve intercontinental transport of chemical plumes, Atmospheric Chemistry and Physics, 17, 2543-2553, 10.5194/acp-17-2543-2017, 2017.

Fan, Y., Sun, L., Wang, Z., Pang, S., and Wei, J.: Unveiling diurnal aerosol layer height variability from space using deep learning, ISPRS Journal of Photogrammetry and Remote Sensing, 229, 211-222, 10.1016/j.isprsjprs.2025.08.021, 2025.

Ge, C., Wang, J., and Reid, J. S.: Mesoscale modeling of smoke transport over the Southeast Asian Maritime Continent: coupling of smoke direct radiative effect below and above the low-level clouds, Atmospheric Chemistry and Physics, 14, 159-174, 10.5194/acp-14-159-2014, 2014.

Geer, A. J.: Learning earth system models from observations: machine learning or data assimilation?, Philosophical Transactions of the Royal Society a-Mathematical Physical and Engineering Sciences, 379, 10.1098/rsta.2020.0089, 2021.

Gelaro, R., McCarty, W., Suárez, M. J., Todling, R., Molod, A., Takacs, L., Randles, C. A., Darmenov, A., Bosilovich, M. G., Reichle, R., Wargan, K., Coy, L., Cullather, R., Draper, C., Akella, S., Buchard, V., Conaty, A., da Silva, A. M., Gu, W., Kim, G. K., Koster, R., Lucchesi, R., Merkova, D., Nielsen, J. E., Partyka, G., Pawson, S., Putman, W., Rienecker, M., Schubert, S. D., Sienkiewicz, M., and Zhao, B.: The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2), Journal of Climate, 30, 5419-5454, 10.1175/jcli-d-16-0758.1, 2017.

Getzewich, B. J., Vaughan, M. A., Hunt, W. H., Avery, M. A., Powell, K. A., Tackett, J. L., Winker, D. M., Kar, J., Lee, K. P., and Toth, T. D.: CALIPSO lidar calibration at 532 nm: version 4 daytime algorithm, Atmospheric Measurement Techniques, 11, 6309-6326, 10.5194/amt-11-6309-2018, 2018.

Giglio, L., Randerson, J. T., and van der Werf, G. R.: Analysis of daily, monthly, and annual burned area using the fourth-generation global fire emissions database

https://doi.org/10.5194/egusphere-2026-397
Preprint. Discussion started: 17 February 2026

(GFED4), Journal of Geophysical Research-Biogeosciences, 118, 317-328, 10.1002/jgrg.20042, 2013.

Giles, D. M., Sinyuk, A., Sorokin, M. G., Schafer, J. S., Smirnov, A., Slutsker, I., Eck, T. F., Holben, B. N., Lewis, J. R., Campbell, J. R., Welton, E. J., Korkin, S. V., and Lyapustin, A. I.: Advancements in the Aerosol Robotic Network (AERONET) Version 3 database - automated near-real-time quality control algorithm with improved cloud screening for Sun photometer aerosol optical depth (AOD) measurements, Atmospheric Measurement Techniques, 12, 169-209, 10.5194/amt-12-169-2019, 2019.

Goldstein, A. H., Koven, C. D., Heald, C. L., and Fung, I. Y.: Biogenic carbon and anthropogenic pollutants combine to form a cooling haze over the southeastern United States, Proceedings of the National Academy of Sciences of the United States of America, 106, 8835-8840, 10.1073/pnas.0904128106, 2009.

Grythe, H., Ström, J., Krejci, R., Quinn, P., and Stohl, A.: A review of sea-spray aerosol source functions using a large global set of sea salt aerosol concentration measurements, Atmospheric Chemistry and Physics, 14, 1277-1297, 10.5194/acp-14-1277-2014, 2014.

Guenther, A. B., Jiang, X., Heald, C. L., Sakulyanontvittaya, T., Duhl, T., Emmons, L. K., and Wang, X.: The Model of Emissions of Gases and Aerosols from Nature version 2.1 (MEGAN2.1): an extended and updated framework for modeling biogenic emissions, Geoscientific Model Development, 5, 1471-1492, 10.5194/gmd-5-1471-2012, 2012.

Guo, J., He, J., Liu, H., Miao, Y., Liu, H., and Zhai, P.: Impact of various emission control schemes on air quality using WRF-Chem during APEC China 2014, Atmospheric Environment, 140, 311-319, 10.1016/j.atmosenv.2016.05.046, 2016.

Handschuh, J., Erbertseder, T., Schaap, M., and Baier, F.: Estimating $PM_{2.5}$ surface concentrations from AOD: A combination of SLSTR and MODIS, Remote Sensing Applications: Society and Environment, 26, 10.1016/j.rsase.2022.100716,

https://doi.org/10.5194/egusphere-2026-397
Preprint. Discussion started: 17 February 2026

2022.

Hayasaka, T., Satake, S., Shimizu, A., Sugimoto, N., Matsui, I., Aoki, K., and Muraji, Y.: Vertical distribution and optical properties of aerosols observed over Japan during the Atmospheric Brown Clouds–East Asia Regional Experiment 2005, Journal of Geophysical Research: Atmospheres, 112, 10.1029/2006jd008086, 2007.

Henze, D. K., Hakami, A., and Seinfeld, J. H.: Development of the adjoint of GEOS-Chem, Atmospheric Chemistry and Physics, 7, 2413-2433, 10.5194/acp-7-2413-2007, 2007.

Henze, D. K., Seinfeld, J. H., and Shindell, D. T.: Inverse modeling and mapping US air quality influences of inorganic $PM_{2.5}$ precursor emissions using the adjoint of GEOS-Chem, Atmospheric Chemistry and Physics, 9, 5877-5903, 10.5194/acp-9-5877-2009, 2009.

Hoesly, R. M., Smith, S. J., Feng, L., Klimont, Z., Janssens-Maenhout, G., Pitkanen, T., Seibert, J. J., Vu, L., Andres, R. J., Bolt, R. M., Bond, T. C., Dawidowski, L., Kholod, N., Kurokawa, J.-i., Li, M., Liu, L., Lu, Z., Moura, M. C. P., O'Rourke, P. R., and Zhang, Q.: Historical (1750–2014) anthropogenic emissions of reactive gases and aerosols from the Community Emissions Data System (CEDS), Geoscientific Model Development, 11, 369-408, 10.5194/gmd-11-369-2018, 2018.

Holben, B. N., Eck, T. F., Slutsker, I., Tanre, D., Buis, J. P., Setzer, A., Vermote, E., Reagan, J. A., Kaufman, Y. J., Nakajima, T., Lavenu, F., Jankowiak, I., and Smirnov, A.: AERONET - A federated instrument network and data archive for aerosol characterization, Remote Sensing of Environment, 66, 1-16, 10.1016/s0034-4257(98)00031-5, 1998.

Holtslag, A. A. M., Svensson, G., Baas, P., Basu, S., Beare, B., Beljaars, A. C. M., Bosveld, F. C., Cuxart, J., Lindvall, J., Steeneveld, G. J., Tjernström, M., and Van De Wiel, B. J. H.: Stable Atmospheric Boundary Layers and Diurnal Cycles: Challenges for Weather and Climate Models, Bulletin of the American

https://doi.org/10.5194/egusphere-2026-397
Preprint. Discussion started: 17 February 2026

Meteorological Society, 94, 1691-1706, 10.1175/bams-d-11-00187.1, 2013.

Hong, Y. and Di Girolamo, L.: An Overview of Aerosol Properties in Clear and Cloudy Sky Based on CALIPSO Observations, Earth and Space Science, 9, 10.1029/2022ea002287, 2022.

Hu, X. F., Waller, L. A., Lyapustin, A., Wang, Y. J., Al-Hamdan, M. Z., Crosson, W. L., Estes, M. G., Estes, S. M., Quattrochi, D. A., Puttaswamy, S. J., and Liu, Y.: Estimating ground-level $PM_{2.5}$ concentrations in the Southeastern United States using MAIAC AOD retrievals and a two-stage model, Remote Sensing of Environment, 140, 220-232, 10.1016/j.rse.2013.08.032, 2014.

Huang, J. T., Loría-Salazar, S., Deng, M., Lee, J., and Holmes, H. A.: Assessment of smoke plume height products derived from multisource satellite observations using lidar-derived height metrics for wildfires in the western US, Atmospheric Chemistry and Physics, 24, 3673-3698, 10.5194/acp-24-3673-2024, 2024.

Huang, L., Jiang, J. H., Tackett, J. L., Su, H., and Fu, R.: Seasonal and diurnal variations of aerosol extinction profile and type distribution from CALIPSO 5-year observations, Journal of Geophysical Research: Atmospheres, 118, 4572-4596, 10.1002/jgrd.50407, 2013.

Hunt, W. H., Winker, D. M., Vaughan, M. A., Powell, K. A., Lucker, P. L., and Weimer, C.: CALIPSO Lidar Description and Performance Assessment, Journal of Atmospheric and Oceanic Technology, 26, 1214-1228, 10.1175/2009jtecha1223.1, 2009.

Jiang, X., Wang, Y., Wang, L., Tao, M., Wang, J., Zhou, M., Bai, X., and Gui, L.: Characteristics of Daytime-And-Nighttime AOD Differences Over China: A Perspective From CALIOP Satellite Observations and GEOS-Chem Model Simulations, Journal of Geophysical Research: Atmospheres, 129, 10.1029/2023jd039158, 2024.

Kahn, R. A., Gaitley, B. J., Martonchik, J. V., Diner, D. J., Crean, K. A., and Holben, B.: Multiangle Imaging Spectroradiometer (MISR) global aerosol optical depth

https://doi.org/10.5194/egusphere-2026-397
Preprint. Discussion started: 17 February 2026

validation based on 2 years of coincident Aerosol Robotic Network (AERONET) observations - art. no. D10S04, Journal of Geophysical Research-Atmospheres, 110, 10.1029/2004jd004706, 2005.

Kaufman, Y. J., Tanré, D., and Boucher, O.: A satellite view of aerosols in the climate system, Nature, 419, 215-223, 10.1038/nature01091, 2002.

Kim, H., Chen, X., Wang, J., Lu, Z. D., Zhou, M., Carmichael, G. R., Park, S. S., and Kim, J.: Aerosol layer height (ALH) retrievals from oxygen absorption bands: intercomparison and validation among different satellite platforms, GEMS, EPIC, and TROPOMI, Atmospheric Measurement Techniques, 18, 327-349, 10.5194/amt-18-327-2025, 2025.

Kim, H., Park, R. J., Hong, S. Y., Park, D. H., Kim, S. W., Oak, Y. J., Feng, X., Lin, H., and Fu, T. M.: A mixed layer height parameterization in a 3-D chemical transport model: Implications for gas and aerosol simulations, Sci Total Environ, 955, 176838, 10.1016/j.scitotenv.2024.176838, 2024.

Kim, K.-Y.: Diurnal and seasonal variation of planetary boundary layer height over East Asia and its climatic change as seen in the ERA-5 reanalysis data, SN Applied Sciences, 4, 10.1007/s42452-021-04918-5, 2022.

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, Atmospheric Measurement Techniques, 11, 6107-6135, 10.5194/amt-11-6107-2018, 2018.

Kim, P. S., Jacob, D. J., Fisher, J. A., Travis, K., Yu, K., Zhu, L., Yantosca, R. M., Sulprizio, M. P., Jimenez, J. L., Campuzano-Jost, P., Froyd, K. D., Liao, J., Hair, J. W., Fenn, M. A., Butler, C. F., Wagner, N. L., Gordon, T. D., Welti, A., Wennberg, P. O., Crounse, J. D., St. Clair, J. M., Teng, A. P., Millet, D. B., Schwarz, J. P., Markovic, M. Z., and Perring, A. E.: Sources, seasonality, and trends of southeast US aerosol: an integrated analysis of surface, aircraft, and satellite observations

https://doi.org/10.5194/egusphere-2026-397
Preprint. Discussion started: 17 February 2026

with the GEOS-Chem chemical transport model, Atmospheric Chemistry and Physics, 15, 10411-10433, 10.5194/acp-15-10411-2015, 2015.

Koch, D. and Del Genio, A. D.: Black carbon semi-direct effects on cloud cover: review and synthesis, Atmospheric Chemistry and Physics, 10, 7685-7696, 10.5194/acp-10-7685-2010, 2010.

Koffi, B., Schulz, M., Bréon, F. M., Griesfeller, J., Winker, D., Balkanski, Y., Bauer, S., Berntsen, T., Chin, M., Collins, W. D., Dentener, F., Diehl, T., Easter, R., Ghan, S., Ginoux, P., Gong, S., Horowitz, L. W., Iversen, T., Kirkevåg, A., Koch, D., Krol, M., Myhre, G., Stier, P., and Takemura, T.: Application of the CALIOP layer product to evaluate the vertical distribution of aerosols estimated by global models: AeroCom phase I results, Journal of Geophysical Research: Atmospheres, 117, 10.1029/2011jd016858, 2012a.

Koffi, B., Schulz, M., Bréon, F. M., Griesfeller, J., Winker, D., Balkanski, Y., Bauer, S., Berntsen, T., Chin, M. A., Collins, W. D., Dentener, F., Diehl, T., Easter, R., Ghan, S., Ginoux, P., Gong, S. L., Horowitz, L. W., Iversen, T., Kirkevåg, A., Koch, D., Krol, M., Myhre, G., Stier, P., and Takemura, T.: Application of the CALIOP layer product to evaluate the vertical distribution of aerosols estimated by global models: AeroCom phase I results, Journal of Geophysical Research-Atmospheres, 117, 10.1029/2011jd016858, 2012b.

Koffi, B., Schulz, M., Breon, F. M., Dentener, F., Steensen, B. M., Griesfeller, J., Winker, D., Balkanski, Y., Bauer, S. E., Bellouin, N., Berntsen, T., Bian, H., Chin, M., Diehl, T., Easter, R., Ghan, S., Hauglustaine, D. A., Iversen, T., Kirkevag, A., Liu, X., Lohmann, U., Myhre, G., Rasch, P., Seland, O., Skeie, R. B., Steenrod, S. D., Stier, P., Tackett, J., Takemura, T., Tsigaridis, K., Vuolo, M. R., Yoon, J., and Zhang, K.: Evaluation of the aerosol vertical distribution in global aerosol models through comparison against CALIOP measurements: AeroCom phase II results, J Geophys Res Atmos, 121, 7254-7283, 10.1002/2015JD024639, 2016.

Koren, I., Kaufman, Y. J., Remer, L. A., and Martins, J. V.: Measurement of the effect

https://doi.org/10.5194/egusphere-2026-397
Preprint. Discussion started: 17 February 2026

of Amazon smoke on inhibition of cloud formation, Science, 303, 1342-1345, 10.1126/science.1089424, 2004.

Li, T. W., Shen, H. F., Yuan, Q. Q., Zhang, X. C., and Zhang, L. P.: Estimating Ground-Level $PM_{2.5}$ by Fusing Satellite and Station Observations: A Geo-Intelligent Deep Learning Approach, Geophysical Research Letters, 44, 11985-11993, 10.1002/2017gl075710, 2017.

Li, X. N., Cheng, X., Wu, W. J., Wang, Q. H., Tong, Z. Y., Zhang, X. Q., Deng, D. H., and Li, Y. H.: Forecasting of bioaerosol concentration by a Back Propagation neural network model, Science of the Total Environment, 698, 10.1016/j.scitotenv.2019.134315, 2020.

Liang, M., Han, Z., Li, J., Sun, Y., Liang, L., and Li, Y.: Radiative effects and feedbacks of anthropogenic aerosols on boundary layer meteorology and fine particulate matter during the COVID-19 lockdown over China, Sci Total Environ, 862, 160767, 10.1016/j.scitotenv.2022.160767, 2023.

Lin, J. T. and McElroy, M. B.: Impacts of boundary layer mixing on pollutant vertical profiles in the lower troposphere: Implications to satellite remote sensing, Atmospheric Environment, 44, 1726-1739, 10.1016/j.atmosenv.2010.02.009, 2010.

Liu, J., Zheng, Y., Li, Z., Flynn, C., and Cribb, M.: Seasonal variations of aerosol optical properties, vertical distribution and associated radiative effects in the Yangtze Delta region of China, Journal of Geophysical Research: Atmospheres, 117, 10.1029/2011jd016490, 2012.

Liu, Y., Huang, J., Shi, G., Takamura, T., Khatri, P., Bi, J., Shi, J., Wang, T., Wang, X., and Zhang, B.: Aerosol optical properties and radiative effect determined from sky-radiometer over Loess Plateau of Northwest China, Atmospheric Chemistry and Physics, 11, 11455-11463, 10.5194/acp-11-11455-2011, 2011.

Lu, Q., Liu, C., Zhao, D. L., Zeng, C., Li, J., Lu, C. S., Wang, J. D., and Zhu, B.: Atmospheric heating rate due to black carbon aerosols: Uncertainties and impact

https://doi.org/10.5194/egusphere-2026-397
Preprint. Discussion started: 17 February 2026

1415 factors, Atmospheric Research, 240, 10.1016/j.atmosres.2020.104891, 2020.
1416 Lu, Z., Wang, J., Chen, X., Xu, X., Zhou, M., Fu, D., and Jiang, J. H.: First Retrieval of
1417 Aerosol Vertical Profile With Passive Remote Sensing: Part 1. Development of
1418 Algorithm Theoretical Basis, Journal of Geophysical Research: Atmospheres, 130,
1419 10.1029/2025jd044332, 2025a.
1420 Lu, Z. D., Wang, J., Chen, X., Xu, X. G., Zhou, M., Fu, D. J., and Jiang, J. H.: First
1421 Retrieval of Aerosol Vertical Profile With Passive Remote Sensing: Part 1.
1422 Development of Algorithm Theoretical Basis, Journal of Geophysical Research-
1423 Atmospheres, 130, 10.1029/2025jd044332, 2025b.
1424 Lundberg, S. M. and Lee, S. I.: A Unified Approach to Interpreting Model Predictions,
1425 31st Annual Conference on Neural Information Processing Systems (NIPS), Long
1426 Beach, CA, Dec 04-09 2017, WOS:000452649404081, 2017.
1427 Lv, B., Hu, Y., Chang, H. H., Russell, A. G., and Bai, Y.: Improving the Accuracy of
1428 Daily $PM_{2.5}$ Distributions Derived from the Fusion of Ground-Level
1429 Measurements with Aerosol Optical Depth Observations, a Case Study in North
1430 China, Environmental Science & Technology, 50, 4752-4759,
1431 10.1021/acs.est.5b05940, 2016.
1432 Ma, X., Bartlett, K., Harmon, K., and Yu, F.: Comparison of AOD between CALIPSO
1433 and MODIS: significant differences over major dust and biomass burning regions,
1434 Atmospheric Measurement Techniques, 6, 2391-2401, 10.5194/amt-6-2391-2013,
1435 2013.
1436 McDuffie, E. E., Smith, S. J., O'Rourke, P., Tibrewal, K., Venkataraman, C., Marais, E.
1437 A., Zheng, B., Crippa, M., Brauer, M., and Martin, R. V.: A global anthropogenic
1438 emission inventory of atmospheric pollutants from sector- and fuel-specific
1439 sources (1970–2017): an application of the Community Emissions Data System
1440 (CEDS), Earth System Science Data, 12, 3413-3442, 10.5194/essd-12-3413-2020,
1441 2020.
1442 Mehta, S. K., Ananthavel, A., Velu, V., Prabhakaran, T., Pandithurai, G., and Rao, D.

https://doi.org/10.5194/egusphere-2026-397
Preprint. Discussion started: 17 February 2026

N.: Characteristics of elevated aerosol layer over the Indian east coast, Kattankulathur (12.82°N, 80.04°E): A northeast monsoon region, Science of The Total Environment, 886, 10.1016/j.scitotenv.2023.163917, 2023.

Mercado, L. M., Bellouin, N., Sitch, S., Boucher, O., Huntingford, C., Wild, M., and Cox, P. M.: Impact of changes in diffuse radiation on the global land carbon sink, Nature, 458, 1014-U1087, 10.1038/nature07949, 2009.

Miao, S. G., Chen, F., Lemone, M. A., Tewari, M., Li, Q. C., and Wang, Y. C.: An Observational and Modeling Study of Characteristics of Urban Heat Island and Boundary Layer Structures in Beijing, Journal of Applied Meteorology and Climatology, 48, 484-501, 10.1175/2008jamc1909.1, 2009.

Misra, A., Gaur, A., Bhattu, D., Ghosh, S., Dwivedi, A. K., Dalai, R., Paul, D., Gupta, T., Tare, V., Mishra, S. K., Singh, S., and Tripathi, S. N.: An overview of the physico-chemical characteristics of dust at Kanpur in the central Indo-Gangetic basin, Atmospheric Environment, 97, 386-396, 10.1016/j.atmosenv.2014.08.043, 2014.

Munroe, D. K., Wolfinbarger, S. R., Calder, C. A., Shi, T., Xiao, N., Lam, C. Q., and Li, D.: The relationships between biomass burning, land-cover/-use change, and the distribution of carbonaceous aerosols in mainland Southeast Asia: a review and synthesis, Journal of Land Use Science, 3, 161-183, 10.1080/17474230802332241, 2008.

Murphy, D. M., Froyd, K. D., Bian, H. S., Brock, C. A., Dibb, J. E., DiGangi, J. P., Diskin, G., Dollner, M., Kupc, A., Scheuer, E. M., Schill, G. P., Weinzierf, B., Williamson, C. J., and Yu, P. F.: The distribution of sea-salt aerosol in the global troposphere, Atmospheric Chemistry and Physics, 19, 4093-4104, 10.5194/acp-19-4093-2019, 2019.

Myhre, G., Samset, B. H., Schulz, M., Balkanski, Y., Bauer, S., Berntsen, T. K., Bian, H., Bellouin, N., Chin, M., Diehl, T., Easter, R. C., Feichter, J., Ghan, S. J., Hauglustaine, D., Iversen, T., Kinne, S., Kirkevåg, A., Lamarque, J. F., Lin, G.,

https://doi.org/10.5194/egusphere-2026-397
Preprint. Discussion started: 17 February 2026

Liu, X., Lund, M. T., Luo, G., Ma, X., van Noije, T., Penner, J. E., Rasch, P. J., Ruiz, A., Seland, O., Skeie, R. B., Stier, P., Takemura, T., Tsigaridis, K., Wang, P., Wang, Z., Xu, L., Yu, H., Yu, F., Yoon, J. H., Zhang, K., Zhang, H., and Zhou, C.: Radiative forcing of the direct aerosol effect from AeroCom Phase II simulations, Atmospheric Chemistry and Physics, 13, 1853-1877, 10.5194/acp-13-1853-2013, 2013.

Nanda, S., de Graaf, M., Veefkind, J. P., Sneep, M., ter Linden, M., Sun, J. Y. T., and Levelt, P. F.: A first comparison of TROPOMI aerosol layer height (ALH) to CALIOP data, Atmospheric Measurement Techniques, 13, 3043-3059, 10.5194/amt-13-3043-2020, 2020.

Nanda, S., Veefkind, J. P., de Graaf, M., Sneep, M., Stammes, P., de Haan, J. F., Sanders, A. F. J., Apituley, A., Tuinder, O., and Levelt, P. F.: A weighted least squares approach to retrieve aerosol layer height over bright surfaces applied to GOME-2 measurements of the oxygen A band for forest fire cases over Europe, Atmospheric Measurement Techniques, 11, 3263-3280, 10.5194/amt-11-3263-2018, 2018.

Nguyen, D.-L., Czech, H., Pieber, S. M., Schnelle-Kreis, J., Steinbacher, M., Orasche, J., Henne, S., Popovicheva, O. B., Abbaszade, G., Engling, G., Bukowiecki, N., Nguyen, N.-A., Nguyen, X.-A., and Zimmermann, R.: Carbonaceous aerosol composition in air masses influenced by large-scale biomass burning: a case study in northwestern Vietnam, Atmospheric Chemistry and Physics, 21, 8293-8312, 10.5194/acp-21-8293-2021, 2021.

Pashayi, M., Satari, M., and Shahraki, M. M.: Multi-layer retrieval of aerosol optical depth in the troposphere using SEVIRI data: a case study of the European continent, Atmospheric Measurement Techniques, 18, 1415-1439, 10.5194/amt-18-1415-2025, 2025.

Paugam, R., Wooster, M., Freitas, S., and Martin, M. V.: A review of approaches to estimate wildfire plume injection height within large-scale atmospheric chemical transport models, Atmospheric Chemistry and Physics, 16, 907-925, 10.5194/acp-

https://doi.org/10.5194/egusphere-2026-397
Preprint. Discussion started: 17 February 2026

16-907-2016, 2016.

Qu, W., Wang, J., Zhang, X., Sheng, L., and Wang, W.: Opposite seasonality of the aerosol optical depth and the surface particulate matter concentration over the north China Plain, Atmospheric Environment, 127, 90-99, 10.1016/j.atmosenv.2015.11.061, 2016.

Rastigejev, Y., Park, R., Brenner, M. P., and Jacob, D. J.: Resolving intercontinental pollution plumes in global models of atmospheric transport, Journal of Geophysical Research: Atmospheres, 115, 10.1029/2009jd012568, 2010.

Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., and Prabhat: Deep learning and process understanding for data-driven Earth system science, Nature, 566, 195-204, 10.1038/s41586-019-0912-1, 2019.

Samset, B. H., Myhre, G., Schulz, M., Balkanski, Y., Bauer, S., Berntsen, T. K., Bian, H., Bellouin, N., Diehl, T., Easter, R. C., Ghan, S. J., Iversen, T., Kinne, S., Kirkevåg, A., Lamarque, J. F., Lin, G., Liu, X., Penner, J. E., Seland, O., Skeie, R. B., Stier, P., Takemura, T., Tsigaridis, K., and Zhang, K.: Black carbon vertical profiles strongly affect its radiative forcing uncertainty, Atmospheric Chemistry and Physics, 13, 2423-2434, 10.5194/acp-13-2423-2013, 2013.

Sanders, A. F. J., de Haan, J. F., Sneep, M., Apituley, A., Stammes, P., Vieitez, M. O., Tilstra, L. G., Tuinder, O. N. E., Koning, C. E., and Veefkind, J. P.: Evaluation of the operational Aerosol Layer Height retrieval algorithm for Sentinel-5 Precursor: application to O_2 A band observations from GOME-2A, Atmospheric Measurement Techniques, 8, 4947-4977, 10.5194/amt-8-4947-2015, 2015.

Sekiyama, T. T., Tanaka, T. Y., Shimizu, A., and Miyoshi, T.: Data assimilation of CALIPSO aerosol observations, Atmospheric Chemistry and Physics, 10, 39-49, 10.5194/acp-10-39-2010, 2010.

Shao, Y. P., Wyrwoll, K. H., Chappell, A., Huang, J. P., Lin, Z. H., McTainsh, G. H., Mikami, M., Tanaka, T. Y., Wang, X. L., and Yoon, S.: Dust cycle: An emerging core theme in Earth system science, Aeolian Research, 2, 181-204,

https://doi.org/10.5194/egusphere-2026-397
Preprint. Discussion started: 17 February 2026

10.1016/j.aeolia.2011.02.001, 2011.

Shi, S. S., Zhu, B., Lu, W., Yan, S. Q., Fang, C. W., Liu, X. H., Liu, D. Y., and Liu, C.: Estimation of radiative forcing and heating rate based on vertical observation of black carbon in Nanjing, China, Science of the Total Environment, 756, 10.1016/j.scitotenv.2020.144135, 2021.

Shrikumar, A., Greenside, P., and Kundaje, A.: Learning Important Features Through Propagating Activation Differences, 34th International Conference on Machine Learning, Sydney, AUSTRALIA, Aug 06-11 2017, WOS:000683309503025, 2017.

Singh, R. P., Dey, S., Tripathi, S. N., Tare, V., and Holben, B.: Variability of aerosol parameters over Kanpur, northern India, Journal of Geophysical Research: Atmospheres, 109, 10.1029/2004jd004966, 2004.

Song, X. W., Wu, D., Jin, L. N., Xu, Y. Y., Chen, X., and Li, Q.: Aerosol Toxicokinetics: A Framework for Unraveling Toxicological Dynamics from Air to the Body, Environmental Science & Technology, 59, 6379-6386, 10.1021/acs.est.5c00751, 2025.

Stier, P., Seinfeld, J. H., Kinne, S., and Boucher, O.: Aerosol absorption and radiative forcing, Atmospheric Chemistry and Physics, 7, 5237-5261, 10.5194/acp-7-5237-2007, 2007.

Tsay, S.-C., Hsu, N. C., Lau, W. K. M., Li, C., Gabriel, P. M., Ji, Q., Holben, B. N., Judd Welton, E., Nguyen, A. X., Janjai, S., Lin, N.-H., Reid, J. S., Boonjawat, J., Howell, S. G., Huebert, B. J., Fu, J. S., Hansell, R. A., Sayer, A. M., Gautam, R., Wang, S.-H., Goodloe, C. S., Miko, L. R., Shu, P. K., Loftus, A. M., Huang, J., Kim, J. Y., Jeong, M.-J., and Pantina, P.: From BASE-ASIA toward 7-SEAS: A satellite-surface perspective of boreal spring biomass-burning aerosols and clouds in Southeast Asia, Atmospheric Environment, 78, 20-34, 10.1016/j.atmosenv.2012.12.013, 2013.

Uno, I., Eguchi, K., Yumimoto, K., Takemura, T., Shimizu, A., Uematsu, M., Liu, Z. Y.,

https://doi.org/10.5194/egusphere-2026-397
Preprint. Discussion started: 17 February 2026

Wang, Z. F., Hara, Y., and Sugimoto, N.: Asian dust transported one full circuit around the globe, Nature Geoscience, 2, 557-560, 10.1038/ngeo583, 2009.

Val Martin, M., Heald, C. L., Ford, B., Prenni, A. J., and Wiedinmyer, C.: A decadal satellite analysis of the origins and impacts of smoke in Colorado, Atmospheric Chemistry and Physics, 13, 7429-7439, 10.5194/acp-13-7429-2013, 2013.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., and Polosukhin, I.: Attention Is All You Need, 31st Annual Conference on Neural Information Processing Systems (NIPS), Long Beach, CA, Dec 04-09 2017, WOS:000452649406008, 2017.

Vernier, J. P., Thomason, L. W., Pommereau, J. P., Bourassa, A., Pelon, J., Garnier, A., Hauchecorne, A., Blanot, L., Trepte, C., Degenstein, D., and Vargas, F.: Major influence of tropical volcanic eruptions on the stratospheric aerosol layer during the last decade, Geophysical Research Letters, 38, n/a-n/a, 10.1029/2011gl047563, 2011.

Wang, J., Park, S., Zeng, J., Ge, C., Yang, K., Carn, S., Krotkov, N., and Omar, A. H.: Modeling of 2008 Kasatochi volcanic sulfate direct radiative forcing: assimilation of OMI SO_2 plume height data and comparison with MODIS and CALIOP observations, Atmospheric Chemistry and Physics, 13, 1895-1912, 10.5194/acp-13-1895-2013, 2013.

Wang, L., Lyu, B., and Bai, Y.: Global aerosol vertical structure analysis by clustering gridded CALIOP aerosol profiles with fuzzy k-means, Sci Total Environ, 761, 144076, 10.1016/j.scitotenv.2020.144076, 2021a.

Wang, Q., Zhou, C., Zhuge, X. Y., Liu, C., Weng, F. Z., and Wang, M. H.: Retrieval of cloud properties from thermal infrared radiometry using convolutional neural network, Remote Sensing of Environment, 278, 10.1016/j.rse.2022.113079, 2022.

Wang, Q. Q., Jacob, D. J., Spackman, J. R., Perring, A. E., Schwarz, J. P., Moteki, N., Marais, E. A., Ge, C., Wang, J., and Barrett, S. R. H.: Global budget and radiative forcing of black carbon aerosol: Constraints from pole-to-pole (HIPPO)

https://doi.org/10.5194/egusphere-2026-397
Preprint. Discussion started: 17 February 2026

observations across the Pacific, Journal of Geophysical Research-Atmospheres, 119, 195-206, 10.1002/2013jd020824, 2014.

Wang, Y., Wang, J., Xu, X. G., Henze, D. K., Qu, Z., and Yang, K.: Inverse modeling of SO_2 and NO_x emissions over China using multisensor satellite data - Part 1: Formulation and sensitivity analysis, Atmospheric Chemistry and Physics, 20, 6631-6650, 10.5194/acp-20-6631-2020, 2020a.

Wang, Y., Wang, J., Zhou, M., Henze, D. K., Ge, C., and Wang, W.: Inverse modeling of SO_2 and NO_x emissions over China using multisensor satellite data - Part 2: Downscaling techniques for air quality analysis and forecasts, Atmospheric Chemistry and Physics, 20, 6651-6670, 10.5194/acp-20-6651-2020, 2020b.

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, 2023.

Wang, Y. L., Huang, R., Song, S. J., Huang, Z. Y., and Huang, G.: Not All Images are Worth 16x16 Words: Dynamic Transformers for Efficient Image Recognition, 35th Annual Conference on Neural Information Processing Systems (NeurIPS), null, ELECTR NETWORK, Dec 06-14 2021, WOS:000922928400032, 2021.

Wei, J., Li, Z., Guo, J., Sun, L., Huang, W., Xue, W., Fan, T., and Cribb, M.: Satellite-Derived 1-km-Resolution PM_1 Concentrations from 2014 to 2018 across China, Environmental Science & Technology, 53, 13265-13274, 10.1021/acs.est.9b03258, 2019.

Weinzierl, B., Ansmann, A., Prospero, J. M., Althausen, D., Benker, N., Chouza, F., Dollner, M., Farrell, D., Fomba, W. K., Freudenthaler, V., Gasteiger, J., Gross, S., Haarig, M., Heinold, B., Kandler, K., Kristensen, T. B., Mayol-Bracero, O. L., Müller, T., Reitebuch, O., Sauer, D., Schäfler, A., Schepanski, K., Spanu, A., Tegen, I., Toledano, C., and Walser, A.: THE SAHARAN AEROSOL LONG-RANGE

https://doi.org/10.5194/egusphere-2026-397
Preprint. Discussion started: 17 February 2026

TRANSPORT AND AEROSOL-CLOUD-INTERACTION EXPERIMENT Overview and Selected Highlights, Bulletin of the American Meteorological Society, 98, 1427-1451, 10.1175/bams-d-15-00142.1, 2017.

Wilcox, E. M.: Direct and semi-direct radiative forcing of smoke aerosols over clouds, Atmospheric Chemistry and Physics, 12, 139-149, 10.5194/acp-12-139-2012, 2012.

Winker, D. M., Tackett, J. L., Getzewich, B. J., Liu, Z., Vaughan, M. A., and Rogers, R. R.: The global 3-D distribution of tropospheric aerosols as characterized by CALIOP, Atmospheric Chemistry and Physics, 13, 3345-3361, 10.5194/acp-13-3345-2013, 2013.

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

Winker, D. M., Pelon, J., Coakley, J. A., Ackerman, S. A., Charlson, R. J., Colarco, P. R., Flamant, P., Fu, Q., Hoff, R. M., Kittaka, C., Kubar, T. L., Le Treut, H., McCormick, M. P., Mégie, G., Poole, L., Powell, K., Trepte, C., Vaughan, M. A., and Wielicki, B. A.: THE CALIPSO MISSION A Global 3D View of Aerosols and Clouds, Bulletin of the American Meteorological Society, 91, 1211-1229, 10.1175/2010bams3009.1, 2010.

Xing, J., Zheng, S. X., Li, S. W., Huang, L., Wang, X. C., Wang, S. X., Liu, C., Jang, C., Zhu, Y., Zhang, J., Bian, J., Liu, T. Y., Hao, J. M., and Kelly, J. T.: Mimicking atmospheric photochemical modeling with a deep neural network, Atmospheric Research, 265, 10.1016/j.atmosres.2021.105919, 2022.

Xiong, J., Wang, Y., Tao, M., Dong, W., Zhou, L., and Wang, L.: Vertical structure of the aerosols in the troposphere over the North China Plain: An analysis based on observations and simulations from 2007 to 2022, Atmospheric Research, 10.1016/j.atmosres.2025.108348, 2025.

https://doi.org/10.5194/egusphere-2026-397
Preprint. Discussion started: 17 February 2026

Xu, X. G., Chen, X., Wang, J., and Remer, L. A.: Potential of NASA's Plankton, Aerosol, Cloud, and Ocean Ecosystem (PACE) Satellite Observations in the Oxygen Bands for Aerosol over Ocean, Journal of Remote Sensing, 4, 10.34133/remotesensing.0167, 2024.

Xu, X. G., Wang, J., Wang, Y., Zeng, J., Torres, O., Reid, J. S., Miller, S. D., Martins, J. V., and Remer, L. A.: Detecting layer height of smoke aerosols over vegetated land and water surfaces via oxygen absorption bands: hourly results from EPIC/DSCOVR in deep space, Atmospheric Measurement Techniques, 12, 3269-3288, 10.5194/amt-12-3269-2019, 2019.

Xu, X. G., Wang, J., Wang, Y., Zeng, J., Torres, O., Yang, Y. K., Marshak, A., Reid, J., and Miller, S.: Passive remote sensing of altitude and optical depth of dust plumes using the oxygen A and B bands: First results from EPIC/DSCOVR at Lagrange-1 point, Geophysical Research Letters, 44, 7544-7554, 10.1002/2017gl073939, 2017.

Xu, Y., Ramanathan, V., and Washington, W. M.: Observed high-altitude warming and snow cover retreat over Tibet and the Himalayas enhanced by black carbon aerosols, Atmospheric Chemistry and Physics, 16, 1303-1315, 10.5194/acp-16-1303-2016, 2016.

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.

Zarzycki, C. M. and Bond, T. C.: How much can the vertical distribution of black carbon affect its global direct radiative forcing?, Geophysical Research Letters, 37,

https://doi.org/10.5194/egusphere-2026-397
Preprint. Discussion started: 17 February 2026

10.1029/2010gl044555, 2010.

Zeng, Y., Wang, M., Zhao, C., Chen, S., Liu, Z., Huang, X., and Gao, Y.: WRF-Chem v3.9 simulations of the East Asian dust storm in May 2017: modeling sensitivities to dust emission and dry deposition schemes, Geoscientific Model Development, 13, 2125-2147, 10.5194/gmd-13-2125-2020, 2020.

Zhai, S. X., Jacob, D. J., Brewer, J. F., Li, K., Moch, J. M., Kim, J., Lee, S., Lim, H., Lee, H. C., Kuk, S. K., Park, R. J., Jeong, J., Wang, X., Liu, P. F., Luo, G., Yu, F. Q., Meng, J., Martin, R., Travis, K. R., Hair, J. W., Anderson, B. E., Dibb, J. E., Jimenez, J. L., Campuzano-Jost, P., Nault, B. A., Woo, J. H., Kim, Y., Zhang, Q., and Liao, H.: Relating geostationary satellite measurements of aerosol optical depth (AOD) over East Asia to fine particulate matter ($PM_{2.5}$): insights from the KORUS-AQ aircraft campaign and GEOS-Chem model simulations, Atmospheric Chemistry and Physics, 21, 16775-16791, 10.5194/acp-21-16775-2021, 2021.

Zhang, J. L., Campbell, J. R., Reid, J. S., Westphal, D. L., Baker, N. L., Campbell, W. F., and Hyer, E. J.: Evaluating the impact of assimilating CALIOP-derived aerosol extinction profiles on a global mass transport model, Geophysical Research Letters, 38, 10.1029/2011gl047737, 2011.

Zhao, B., Wang, Y., Gu, Y., Liou, K.-N., Jiang, J. H., Fan, J., Liu, X., Huang, L., and Yung, Y. L.: Ice nucleation by aerosols from anthropogenic pollution, Nature Geoscience, 12, 602-607, 10.1038/s41561-019-0389-4, 2019.

Zhen, Y., Yang, X., Tang, H., Shi, H., and Liu, Z.: CALIPSO-based aerosol extinction profile estimation from MODIS and MERRA-2 data using a hybrid model of Transformer and CNN, Science of The Total Environment, 954, 10.1016/j.scitotenv.2024.176423, 2024.

Zhu, H. H., Martin, R. V., van Donkelaar, A., Hammer, M. S., Li, C., Meng, J., Oxford, C. R., Liu, X., Li, Y. S., Zhang, D. D., Singh, I., and Lyapustin, A.: Importance of aerosol composition and aerosol vertical profiles in global spatial variation in the relationship between $PM_{2.5}$ and aerosol optical depth, Atmospheric Chemistry and

https://doi.org/10.5194/egusphere-2026-397
Preprint. Discussion started: 17 February 2026

1695 Physics, 24, 11565-11584, 10.5194/acp-24-11565-2024, 2024.

1696

https://doi.org/10.5194/egusphere-2026-397
Preprint. Discussion started: 17 February 2026

Table 1. Experimental design based on the Leave-One-Year-Out cross-validation strategy.

Experiment ID	Training & Validation Data	Test Data	Evaluation Objective
Exp-2017	Years 2018, 2019	Year 2017	Hindcasting
Exp-2018	Years 2017, 2019	Year 2018	Temporal Interpolation
Exp-2019	Years 2017, 2018	Year 2019	Forecasting

https://doi.org/10.5194/egusphere-2026-397
Preprint. Discussion started: 17 February 2026

Table 2. Quantitative evaluation of the physics-informed Transformer model's predictive performance for AEC simulation biases over EA across independent test years based on the Leave-One-Year-Out strategy.

Experiment ID	R (Test)	MAE (Test)	RMSE (Test)
Exp-2017	0.666	0.014	0.039
Exp-2018	0.659	0.014	0.039
Exp-2019	0.651	0.015	0.041
Average	0.659	0.014	0.040

https://doi.org/10.5194/egusphere-2026-397
Preprint. Discussion started: 17 February 2026

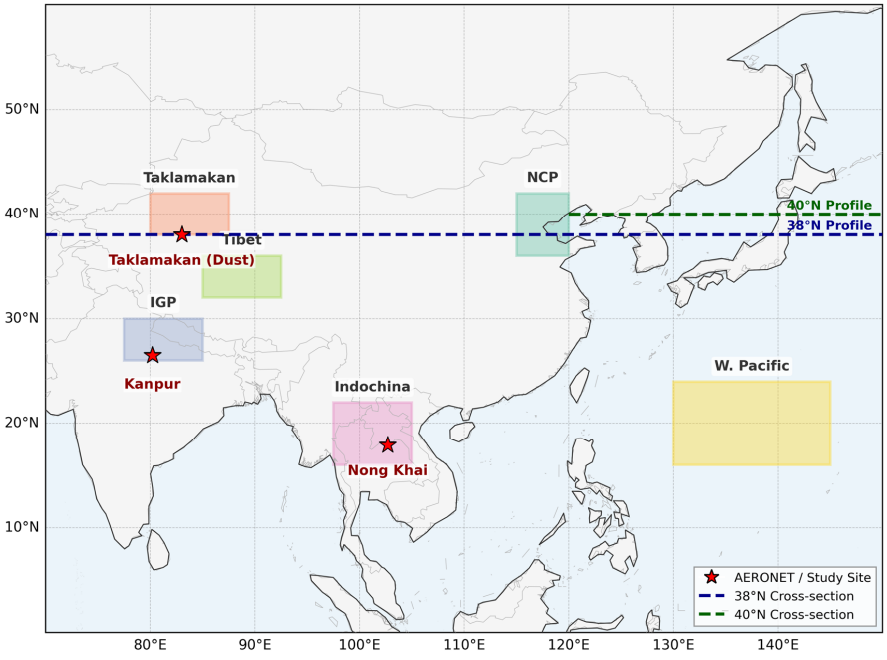


Figure 1. Map of the study domain covering East and South Asia (0°-60°N, 70°-150°E). The colored rectangles delimitate the six regions of interest selected for regional analysis. The dashed lines represent the latitudinal transects at 38°N (blue) and 40°N (green) used for vertical profile cross-section analysis. Red stars indicate the specific study sites selected to represent distinct aerosol regimes: the Kanpur and Nong Khai sites are AERONET ground-based stations selected to represent regions dominated by anthropogenic pollution and biomass burning, respectively; the Taklamakan site marks a reference location selected to investigate the characteristics of natural mineral dust.

https://doi.org/10.5194/egusphere-2026-397
Preprint. Discussion started: 17 February 2026

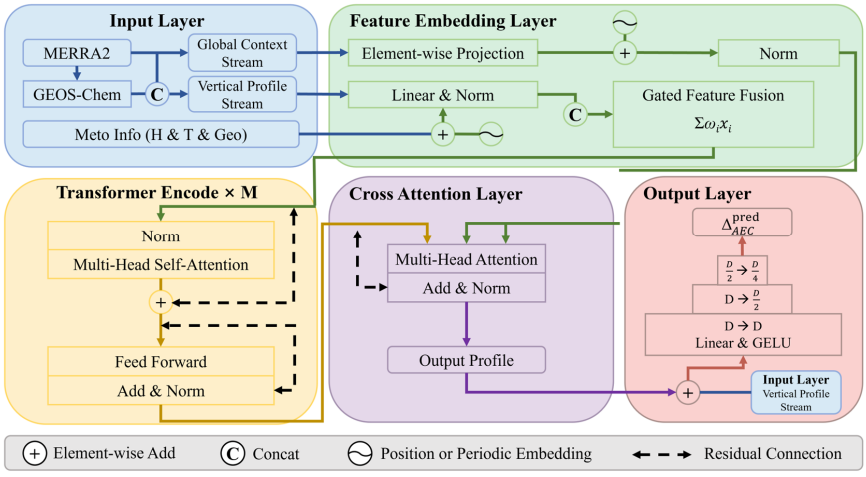


Figure 2. The overall architecture of the physics-informed Transformer model. The network leverages Gated Feature Fusion and Cross-Attention mechanisms within a residual learning framework to correct AEPs.

https://doi.org/10.5194/egusphere-2026-397
Preprint. Discussion started: 17 February 2026

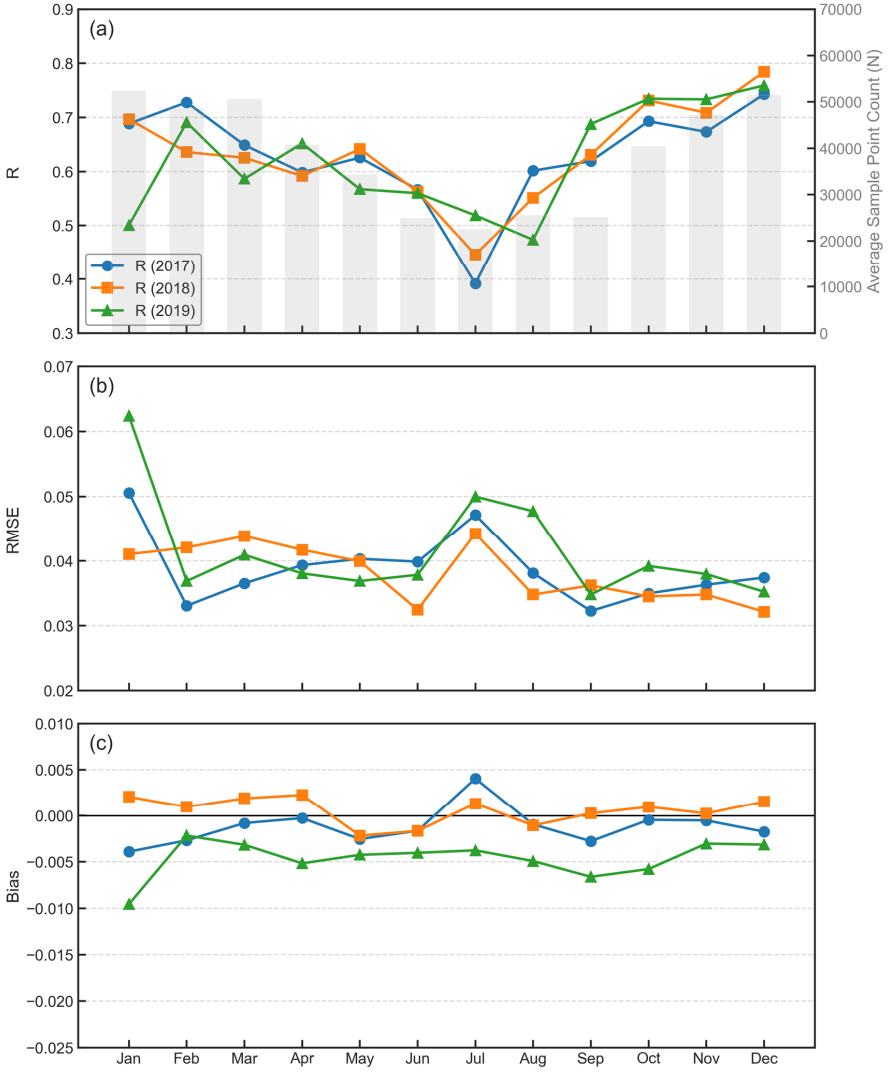


Figure 3. Monthly variations of statistical metrics evaluating the physics-informed Transformer model's predictive performance for the AEC simulation bias over EA across the three test years (2017, 2018, and 2019). The panels display the time series of R (a) with the multi-year average monthly sample size (N, gray bars), RMSE (b), and mean bias (c).

https://doi.org/10.5194/egusphere-2026-397
Preprint. Discussion started: 17 February 2026

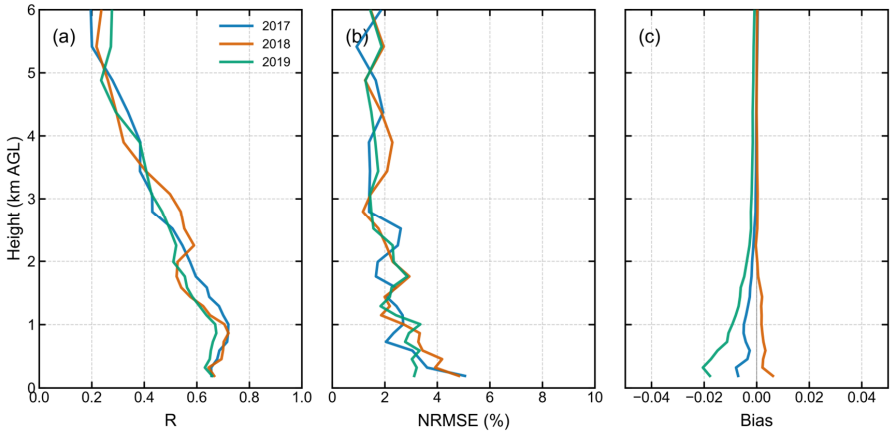


Figure 4. Vertical profiles of the physics-informed Transformer model's predictive performance for AEC simulation bias over EA across three independent test years (2017, 2018, and 2019). The panels show the height-resolved R (a), NRMSE (b), and mean bias (c). The performance is evaluated against CALIOP observations from the surface up to 6 km AGL.

https://doi.org/10.5194/egusphere-2026-397
Preprint. Discussion started: 17 February 2026

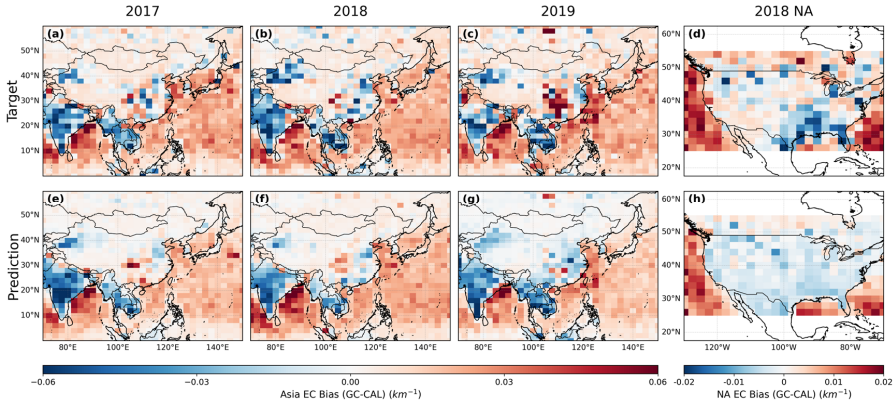


Figure 5. Spatial comparison of the vertically averaged (0–6 km AGL) AEC simulation bias over EA and NA. The top row (a–d) represents the Target systematic bias. The bottom row (e–h) displays the corresponding bias Predicted by the physics-informed Transformer model. Columns 1–3 show results for the independent test years (2017, 2018, 2019) over the primary EA domain (a–c, e–g). Column 4 presents the generalization test over the NA domain for 2018 (d, h), where the model trained on EA data is directly applied to predict biases in an unseen continent.

https://doi.org/10.5194/egusphere-2026-397
Preprint. Discussion started: 17 February 2026

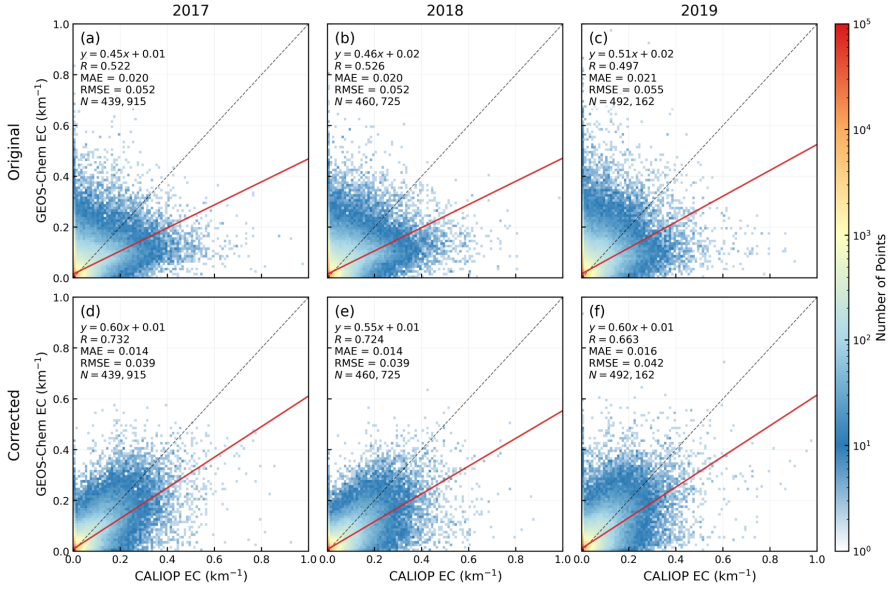


Figure 6. Density scatter plots comparing the simulated AEC against CALIOP observations over EA for the three test years: 2017 (a, d), 2018 (b, e), and 2019 (c, f). The top row (a, b, c) displays the validation results for the original GEOS-Chem simulation, while the bottom row (d, e, f) shows the results after correction by the physics-informed Transformer model.

https://doi.org/10.5194/egusphere-2026-397
Preprint. Discussion started: 17 February 2026

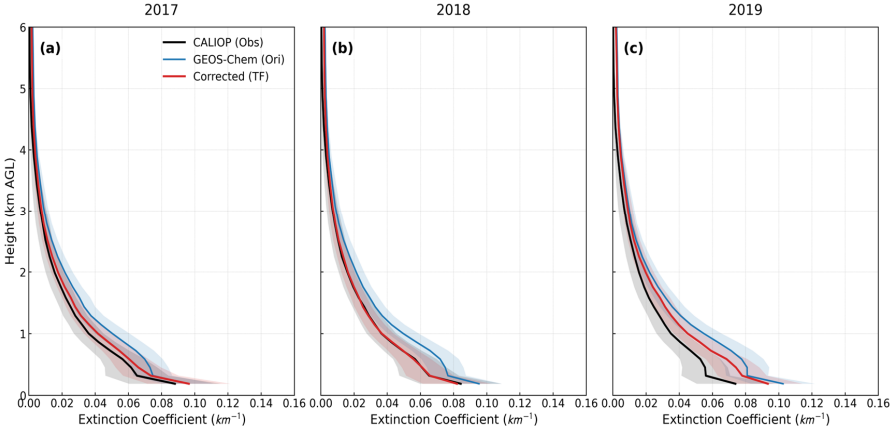


Figure 7. Annual mean vertical profiles of AEC for the three test years (2017–2019). The profiles are averaged over the entire study domain (EA). The shaded areas indicate the standard deviation ($\pm 0.5\sigma$) of the vertical distribution, representing the spatial and temporal variability.

https://doi.org/10.5194/egusphere-2026-397
Preprint. Discussion started: 17 February 2026

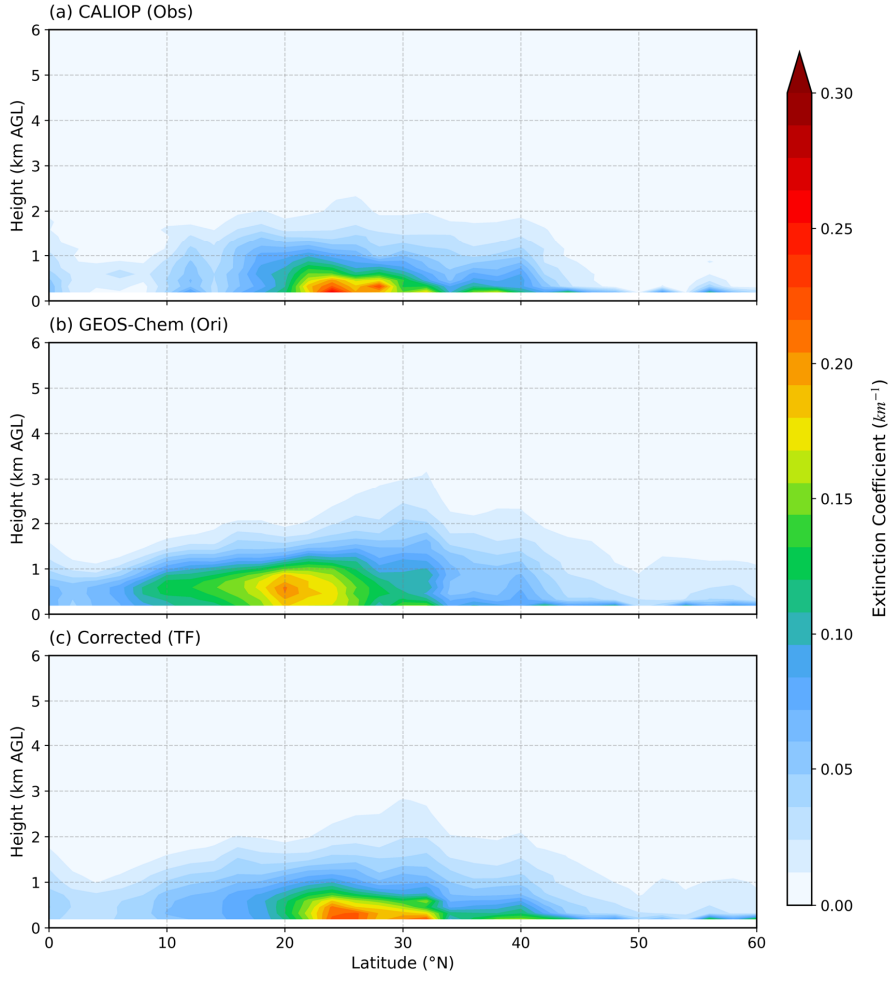


Figure 8. Zonal mean vertical distributions of AEC averaged over the longitude range (70°E–150°E) for the test year 2019. The panels display the latitude-altitude cross-sections for CALIOP observations (a), original GEOS-Chem simulations (b), and the corrected GC-TF results (c).

https://doi.org/10.5194/egusphere-2026-397
Preprint. Discussion started: 17 February 2026

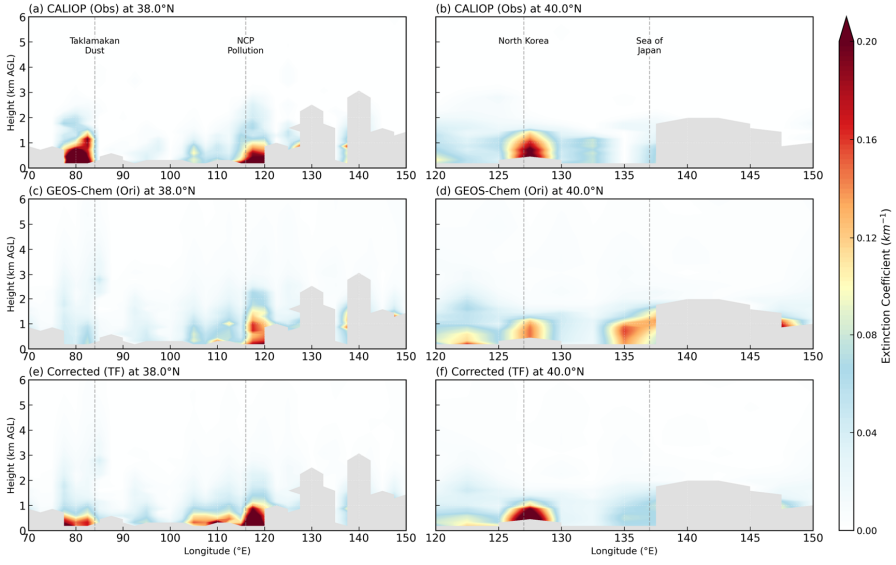


Figure 9. Longitudinal vertical cross-sections of annual mean AEC along two selected latitudes for the test year 2019. The left column displays the profiles at 38.0°N, highlighting the Taklamakan Desert and the NCP. The right column displays the profiles at 40.0°N, capturing the aerosol transport over the Korean Peninsula and the Sea of Japan. Panels show CALIOP observations (a, b), original GEOS-Chem simulations (c, d), and the corrected GC-TF results (e, f). The vertical dashed lines mark the approximate locations of key geographical features. Gray areas indicate missing data.

https://doi.org/10.5194/egusphere-2026-397
Preprint. Discussion started: 17 February 2026

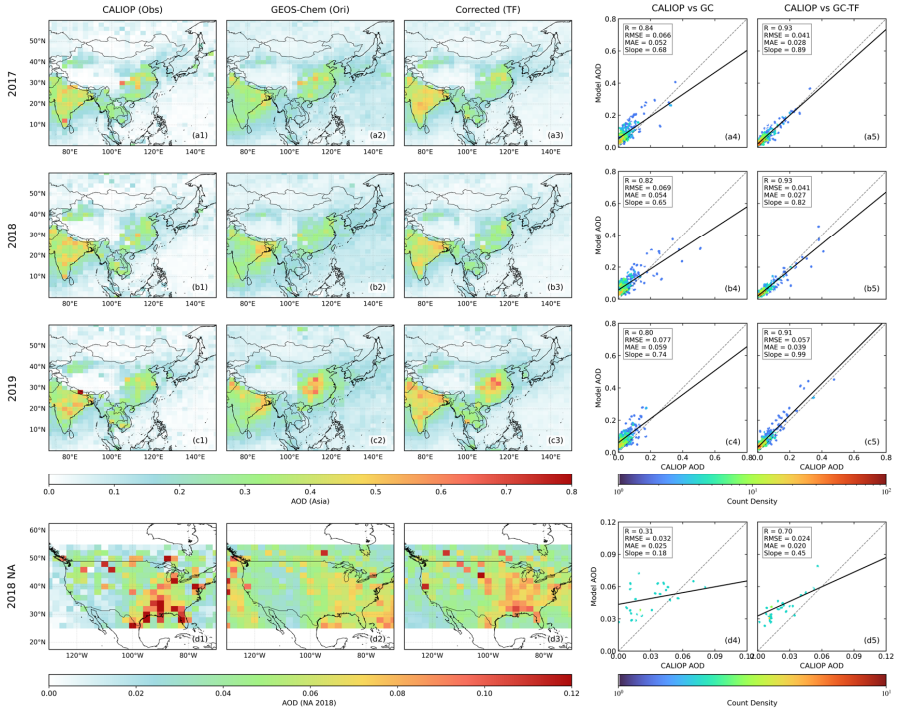


Figure 10. Spatial distributions and statistical evaluations of the annual mean AOD derived from CALIOP observations, GEOS-Chem simulations, and the GC-TF model corrections. Results for the EA domain during the test years 2017 (a), 2018 (b), and 2019 (c), respectively. Columns 1–3 display the spatial patterns of AOD from CALIOP, the original GEOS-Chem (Ori), and the corrected GC-TF results (TF). Columns 4–5 show the corresponding density scatter plots comparing the model predictions (y-axis) against CALIOP observations (x-axis). Generalization test over the NA domain for the year 2018 (d).

https://doi.org/10.5194/egusphere-2026-397
Preprint. Discussion started: 17 February 2026

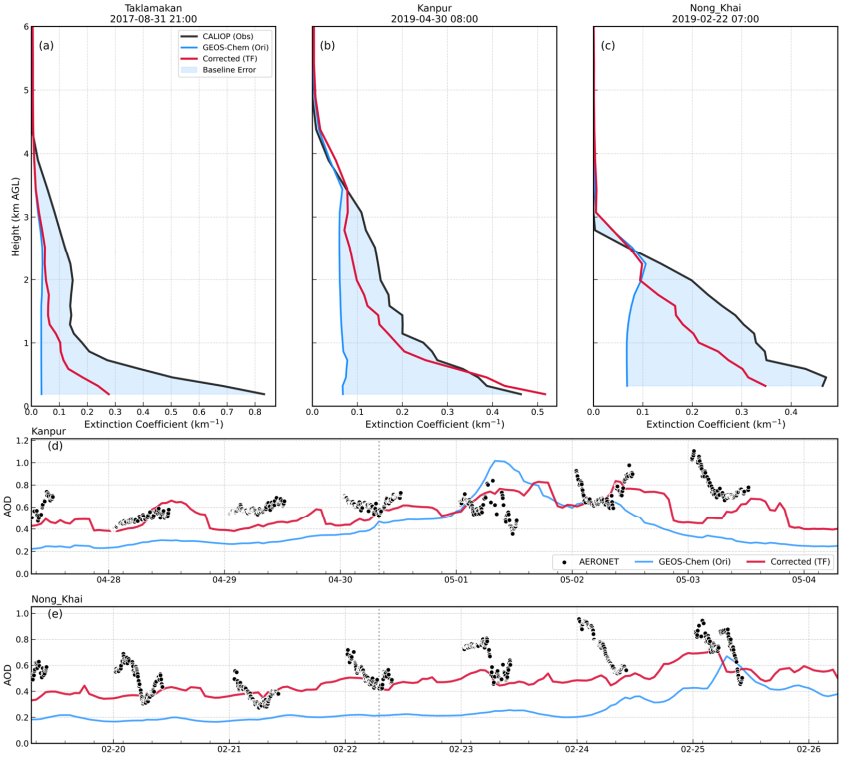
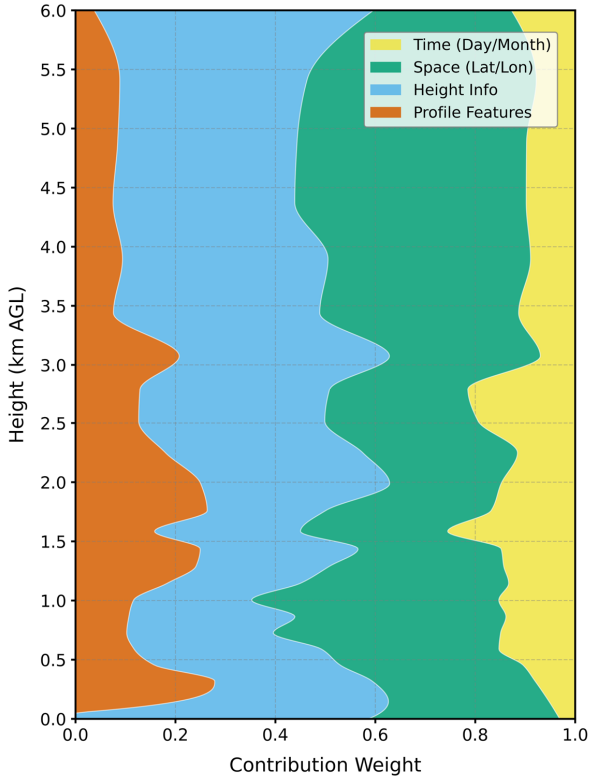


Figure 11. Composite analysis of aerosol vertical structures and temporal evolution during selected pollution episodes. Vertical profiles of AEC at three representative sites: Taklamakan (Dust, a), Kanpur (Anthropogenic Pollution, b), and Nong Khai (Biomass Burning, c). Time series of AOD at the Kanpur (d) and Nong Khai (e) AERONET sites during the corresponding pollution events. The vertical dotted lines mark the CALIOP overpass times (UTC) shown in the top panels.

https://doi.org/10.5194/egusphere-2026-397
Preprint. Discussion started: 17 February 2026



1777

1778 Figure 12. Vertically resolved feature importance analysis. The stacked plot displays
1779 the altitude-dependent contribution weights of the four input feature groups (Profile
1780 Features, Height Info, Space, and Time) learned by the model's Gated Fusion
1781 mechanism.

https://doi.org/10.5194/egusphere-2026-397
Preprint. Discussion started: 17 February 2026

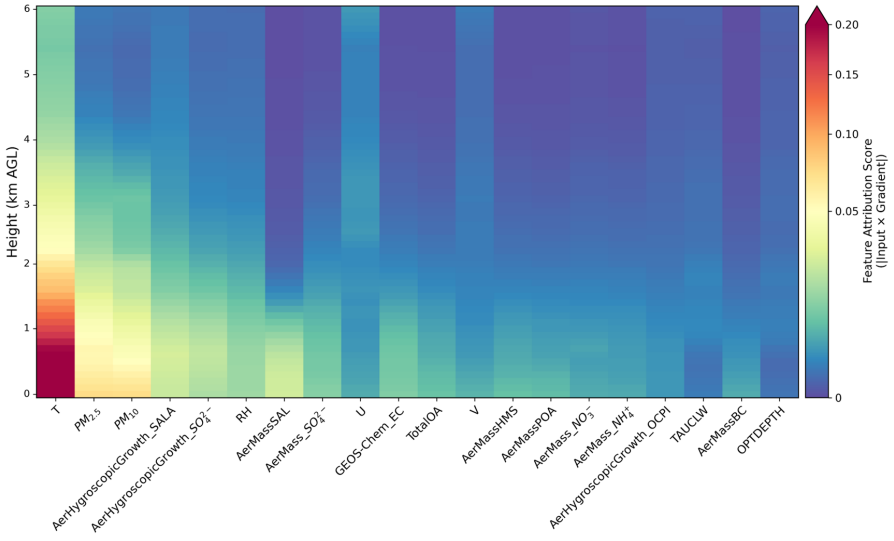


Figure 13. Vertical distribution of feature attribution scores for the top-20 profile variables, calculated using the Integrated Gradients method. The heatmap displays the contribution strength of each variable to the model's prediction at different altitude levels. The variables are sorted by their total contribution.

https://doi.org/10.5194/egusphere-2026-397
Preprint. Discussion started: 17 February 2026

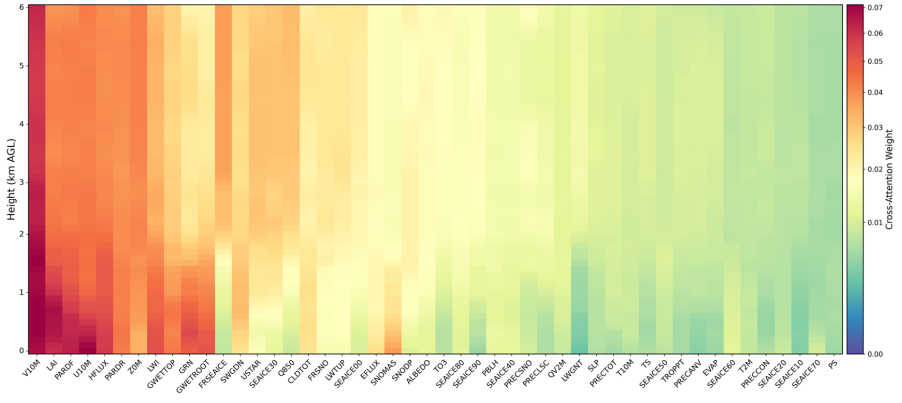


Figure 14. Vertical distribution of the attention weights for global variables learned by the Cross-Attention module. The heatmap displays the interaction strength between the Vertical Profile Stream and the Global Context Stream. The variables are sorted by their total contribution from left to right.

https://doi.org/10.5194/egusphere-2026-397
Preprint. Discussion started: 17 February 2026

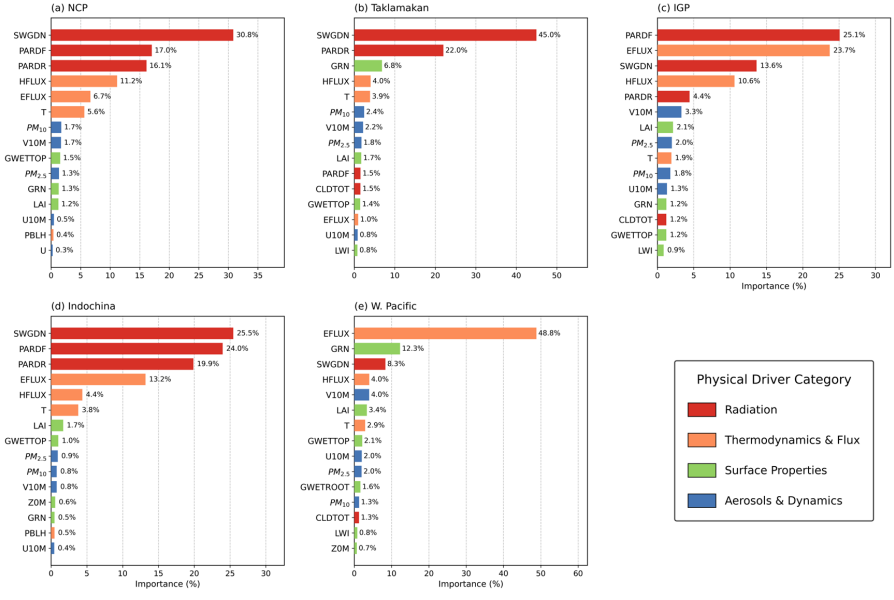


Figure 15. Regional variability in feature importance drivers identified by SHAP analysis for the test year 2019. The panels display the top 15 most influential features for predicting AEC simulation biases in five representative regions: NCP (a), Taklamakan Desert (b), IGP (c), Indochina (d), and Western Pacific (e).