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)

Abstract

Accurately characterizing aerosol vertical distributions is essential for evaluating radiative forcing and air quality. While Chemical Transport Models (CTMs) simulate spatially continuous Aerosol Extinction Coefficient (AEC, km^{-1}), they exhibit systematic AEC biases. Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP) observations provide precise AEC profiles but are constrained by sparse spatial sampling. To bridge this gap, we propose a physics-informed Transformer framework as a supervised bias-correction model to correct biases in the AEC 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 rectification with synoptic contexts. This approach effectively predicts systematic biases relative to CALIOP 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 precision 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 capturing smoothed biomass burning and dust plumes. Additionally, it exhibits robust cross-continental transferability, reproducing bias patterns over the North American domain (R=0.70) without retraining, confirming the internalization of universal physicochemical relationships linking atmospheric states to simulation biases. Furthermore, interpretability analysis serves as a diagnostic tool to guide physical model improvement. The model identifies temperature and sensible heat flux as primary drivers to constrain boundary layer mixing, pointing to potential uncertainties in vertical eddy diffusion. Additionally, it uses environmental proxies (e.g., vegetation

indices and diffuse radiation) to diagnose potential deficiencies in dust threshold friction velocity and secondary organic aerosol yields. These insights provide a physical basis for refining parameterization schemes in CTMs.

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

1. Introduction

Atmospheric aerosols are key modulators of the Earth's climate radiative forcing (Kahn et al., 2005; Stier et al., 2007; Dong et al., 2023; Wang 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 (Wang et al., 2013; Kim et al., 2025; Yorks et al., 2023; Fan et al., 2025)—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 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) (Ding et al., 2016; Nanda et al., 2018; Chen et al., 2025; Sanders et al., 2015) 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 (Liu et al., 2012; Huang 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 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., 2009; Sekiyama et al., 2010; Wang et al., 2020b; Henze et al., 2007; Wang et al., 2020a; Zhang et al., 2011). 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 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 effective tool for fusing the physical consistency of GEOS-Chem simulations with the vertical precision of CALIOP observations to derive high-precision 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 correcting systematic biases to yield high-precision, spatially continuous 3D AEC fields. Distinct from traditional DA systems that require concurrent observational inputs to iteratively update state variables, our framework operates as a supervised bias-correction model. It captures the intrinsic state-dependent mapping between CTM structural uncertainties and diverse environmental contexts. By conditioning the correction exclusively on the CTM's a priori state and meteorological drivers, the model effectively mitigates systematic biases without relying on CALIOP data during the inference phase. Furthermore, 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 observational benchmarks. Focusing on East Asia (EA) from 2017 to 2019, we systematically evaluate the framework’s performance in bias correction 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 bias correction process. 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

bridges data-driven correction with the targeted refinement of GEOS-Chem's physical parameterizations. The remainder of this paper is organized as follows. Section 2 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 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 (detailed in Sect. S1). This stringent assessment framework substantially enhances the precision 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).

To ensure physical consistency between the GEOS-Chem and CALIOP satellite observations, we employ a strict spatiotemporal collocation strategy. Spatially, the high-resolution CALIOP Level 2 profiles are mapped onto the GEOS-Chem grid. All quality-controlled CALIOP profiles falling within a specific grid cell are spatially averaged to represent the observational mean state of that grid box. Temporally, we adopt a precise nearest-hour collocation approach. The CALIOP overpass times are mathematically rounded to the nearest UTC hour and paired strictly with the GEOS-Chem 1-hourly instantaneous outputs. Aligning the instantaneous model output with the concurrent instantaneous observation minimizes temporal representativeness errors (typically constrained within ±30 minutes) (Ichoku et al., 2002). However, we acknowledge that this approach inherently introduces spatial representativeness errors. Averaging the narrow, "curtain-like" nadir view of CALIOP to represent a bulk 2°×2.5° grid box inevitably suffers from sub-grid heterogeneity, particularly in regions with complex terrain or localized intense emissions.

In our physics-informed data-driven framework, the collocated CALIOP AEPs are

222 treated as the observational benchmarks. The training target for the Transformer model
223 is explicitly defined as the simulation bias, calculated as the deviation of the GEOS-
224 Chem simulated AEC from the corresponding CALIOP observations (detailed in Sect.
225 3.1). However, it is essential to acknowledge the intrinsic uncertainties associated with
226 satellite retrievals. Validations against ground-based AERONET measurements
227 indicate a mean relative bias in CALIOP AOD of approximately -5.1%±8.5% (Kim et
228 al., 2018). Furthermore, the agreement between CALIOP 532 nm calibrated attenuated
229 backscatter and airborne High Spectral Resolution Lidar (HSRL) measurements is
230 typically within 1.0%±3.5% (Getzewich et al., 2018). These inherent observational
231 uncertainties coupled with the spatial representativeness errors impose a theoretical
232 upper limit on the precision of the bias correction method proposed herein.

233 **2.3 AERONET**

234 To independently evaluate the model's capability in capturing high-frequency
235 temporal variations—bridging the observational gaps inherent to the sparse sampling
236 of polar-orbiting satellites—we utilize AOD data from the AERONET (Holben et al.,
237 1998). Specifically, the Version 3 Level 2.0 (cloud-screened and quality-assured) AOD
238 data are employed (Giles et al., 2019). Since AERONET instruments do not directly
239 measure AOD at 532 nm, data are interpolated to this wavelength using the Ångström
240 exponent derived from adjacent channels (detail in Sect. S2). Two representative sites
241 are selected for detailed case studies, targeting distinct aerosol regimes during high-
242 loading episodes (Fig. 1): (1) Kanpur (26.5°N, 80.2°E): Located in the IGP, this site is
243 characterized by a complex mixture of anthropogenic pollution and transported dust.
244 We focus on a 7-day window centered on 30 April 2019, representing a typical pre-
245 monsoon scenario driven by regional dust and haze activity (Misra et al., 2014; Singh
246 et al., 2004; Chinnam et al., 2006); (2) Nong Khai (17.9°N, 102.7°E): Situated in the
247 Indochina Peninsula, this site is dominated by carbonaceous aerosols resulting from
248 intense agricultural biomass burning (Munroe et al., 2008; Nguyen et al., 2021). The
249 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 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

We design a dual-stream input architecture to decouple local vertical atmospheric states from synoptic meteorological forcing. The detailed inventory of all input variables is provided in Section S3 in the supplement.

The Vertical Profile Stream (VPS) resolves the atmospheric column through three sub-components. (1) Physicochemical profiles: This includes GEOS-Chem simulated aerosol species and MERRA-2 meteorological profiles. Beyond basic mass concentrations, we explicitly incorporate precursor gases (SO_2, NO_x, NH_3) and microphysical variables (e.g., hygroscopic growth factors and effective radii) to physically constrain secondary aerosol formation and optical extinction. (2) Height information : To maintain vertical stratification within the attention mechanism, we embed explicit altitude information. This allows the model to correctly differentiate near-surface emission interactions from free-tropospheric long-range transport. (3) Spatiotemporal indices: Geographical coordinates (latitude, longitude) and temporal indices (month, day, night) are projected into high-dimensional vectors to capture regional emission patterns and seasonal cycles.

306 The Synoptic Forcing Stream (SFS) incorporates 2D surface diagnostics to
307 represent synoptic constraints on the atmospheric column. Variables such as Planetary
308 Boundary Layer Height (PBLH) and friction velocity act as indicators for turbulent
309 mixing. Surface fluxes and Leaf Area Index (LAI) parameterize deposition and
310 biogenic emissions, while precipitation rates serve as proxies for wet scavenging.

311 Furthermore, although our architecture does not employ explicit historical time-
312 series modeling, it robustly captures diurnal variability. By rigorously matching the
313 instantaneous MERRA-2 fields with the exact CALIOP overpass time, the model is
314 directly conditioned on the concurrent thermodynamic and dynamic states. Combined
315 with explicit day/night flags, this allows the framework to dynamically resolve
316 meteorology-driven diurnal processes (e.g., boundary layer evolution and
317 photochemistry) without relying on lagged predictors.

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

320 $$\Delta_{AEC}^{target} = AEC_{GC} - AEC_{CAL} \quad (1)$$

321 Predicting the simulation bias Δ_{AEC}^{target}, rather than the absolute AEC magnitude,
322 ensures the framework preserves the fundamental physical transport patterns resolved
323 by the CTM, focusing solely on correcting systematic deviations caused by
324 parameterization or emission uncertainties. It is important to emphasize that while
325 CALIOP observations provide the target during training, they are strictly excluded from
326 the input feature space during inference. Consequently, the framework's corrective
327 capacity is inherently bounded by the information content of the GEOS-Chem and
328 MERRA-2 predictors. The model is designed to rectify state-dependent systematic
329 biases rather than to artificially generate aerosol signals from completely unrepresented
330 physical processes that lack corresponding perturbation signatures in the input fields.

331 It should be noted that using CALIOP retrievals as the baseline inherently
332 propagates its systematic uncertainties (e.g., a mean relative bias of -5.1%, as discussed

in Sect. 2.2) into the learning target. If CALIOP exhibits a systematic negative bias, the model may theoretically learn a tendency to slightly over-compensate the AEC. However, because GEOS-Chem's structural biases are typically an order of magnitude larger than these observational uncertainties, the data-driven correction remains highly beneficial. A detailed quantitative evaluation of the model's sensitivity to these observational uncertainties is presented in Section 4.1.5.

3.2 Physically-Informed Transformer Architecture

The overall architecture of our proposed framework (Figure 2) bridges GEOS-Chem simulations and CALIOP observations. To preserve the distinct structural characteristics of atmospheric profiles and synoptic environmental contexts, the framework processes these two streams through specialized embedding strategies (detailed in Sect. S4a, b). The model comprises an altitude-dependent gated feature fusion mechanism, a Transformer encoder for vertical dependencies, a cross-attention module for synoptic constraints, and an output layer.

3.2.1 Altitude-Dependent Gated Feature Fusion

Physical factors governing AEC vary significantly with altitude. Local emissions and chemical composition dominate near-surface AEC (Xiong et al., 2025; Jiang et al., 2024), whereas long-range transport and regional backgrounds dictate the free troposphere (Uno et al., 2009; Val Martin et al., 2013). To reflect this stratification, we design a gated feature fusion mechanism within the VPS. Instead of statically concatenating inputs, this module dynamically weights the contributions of physicochemical profiles, height information, and spatiotemporal indices for each altitude layer. This allows the model to autonomously prioritize the most relevant physical drivers at specific heights.

The SFS incorporates diverse meteorological parameters with distinct physical units. To prevent the network from treating these distinct physical quantities merely as dimensionless numbers, we implement a variable identity embedding (Eq. S4). This mechanism assigns a unique physical tag to each 2D variable, ensuring the model

accurately distinguishes between different meteorological forcing factors when modulating the AEC simulation bias.

3.2.2 Modeling Vertical Connectivity and Synoptic Modulation

Aerosol layers are inherently coupled through vertical exchange processes such as turbulent mixing, deep convection, and gravitational sedimentation. We employ a Transformer encoder stack to explicitly model this vertical connectivity. Its self-attention mechanism acts as a dynamic vertical covariance operator (detailed in Sect. S4c). It facilitates information flow between near-surface accumulation layers and high-altitude transport layers, ensuring the rectified AEC profile maintains physical continuity.

To constrain this vertical AEC bias correction with synoptic meteorology, we introduce a cross-attention layer. Functionally, this mechanism acts as a dynamic diagnostic process. It allows the aerosol state at each specific altitude to dynamically respond to the prevailing synoptic conditions (e.g., surface wind speed, PBLH), thereby extracting relevant environmental constraints for the local bias adjustment. This design mimics physical reality, where synoptic meteorological backgrounds continuously modulate local microphysical structures.

3.2.3 Output Layer

To predict the final AEC bias, we employ a residual connection that adds the initial baseline state (from the VPS) directly to the output of the cross-attention module (which has already fused the encoded VPS with the SFS). Physically, this residual design serves as a critical prior constraint. It anchors the network to the fundamental atmospheric state provided by GEOS-Chem, ensuring the model computes a meteorology-driven perturbation rather than generating unphysical aerosol signals. Subsequently, the integrated features undergo a progressive dimension-reduction (represented as $D \rightarrow D/2 \rightarrow D/4$ in Fig. 2). This architecture functions as an information funnel, filtering redundant meteorological noise and distilling the non-linear interactions among diverse drivers to accurately quantify the true magnitude of the AEC biases.

3.3 Magnitude-Weighted Loss Function

To address the statistical imbalance between the predominant clean background signals and the physically critical pollution episodes, we propose a Magnitude-Weighted Loss (L_{MW}, detailed in Sect. S4g). This customized loss function dynamically rescales the correction weighting to enhance the model's sensitivity to large simulation AEC biases while strictly suppressing spurious aerosol artifacts in atmospheric regimes where the CTM already performs satisfactorily.

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 five 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" 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 (LOYO) cross-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 test the model's transferability beyond its training distribution, we perform an out-of-domain evaluation on the independent NA 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. Furthermore, to dissect the impact of varying aerosol composition regimes on model transferability, the NA validation results are further stratified using CALIOP aerosol subtype classifications.

(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.

(5) Methodological benchmarking: We evaluate the proposed Transformer against conventional machine learning baselines and conduct ablation studies to justify the architectural complexity and isolate the sources of performance improvements.

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 correction framework and ensure physical consistency, we establish a hierarchical diagnostic approach. This framework addresses potential functional overlaps by characterizing model behavior across three scales: micro-scale local sensitivity, domain-wide feature ranking, and regional heterogeneity.

3.5.1 Micro-Scale Local Sensitivity

We employ distinct attribution methods tailored to the hybrid inputs to capture micro-scale responses. For the VPS, we apply gradient-based attribution, utilizing the Input×Gradient method (Shrikumar et al., 2017) to quantify the sensitivity of AEC bias correction to physicochemical profiles. Simultaneously, Cross-Attention weights are extracted to map the interaction strength between the SFS and the VPS, revealing how

synoptic forcing modulates vertical profile rectifications. Furthermore, to understand the model's internal decision-making, we analyze the learnable weights of the gated feature fusion mechanism (detailed in Sect. 3.2.1). This analysis visualizes the altitude-dependent prioritization among the four VPS components: physicochemical profiles, height information, spatial coordinates, and temporal indices.

3.5.2 Domain-wide Feature Ranking

To assess the model's reliance on the overarching input feature groups (the VPS and SFS), we perform permutation feature importance analysis (detailed in Sect. S7d). By measuring the percentage increase in Mean Squared Error (MSE) when specific groups are randomly shuffled, this method provides a domain-wide to identify the fundamental predictors essential for AEC bias correction.

3.5.3 Regional Heterogeneity

Considering the spatial heterogeneity of aerosol sources, SHAP Analysis is used to dissect regional dependencies and feature interactions. 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 the designated ROIs to reveal how dominant AEC bias drivers shift under different environmental regimes.

4. Results and Discussion

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 LOYO 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 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 Precision 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 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 precision 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. Over regions like the IGP, this negative simulation bias is primarily driven by the underrepresentation of local biofuel and agricultural emissions in traditional inventories (Mcduffie et al., 2020), coupled with simplified aerosol mixing state assumptions that underestimate extinction enhancement under high humidity (Burgos et al., 2020; Zhai et al., 2021). Furthermore, the model's excessive numerical diffusion, a common limitation in CTMs, artificially dilutes the strong near-surface pollutant accumulation bounded by local topography (e.g., the Himalayas) (Eastham and Jacob, 2017). The GC-TF framework effectively identifies and mitigates these state-dependent underestimations. Additionally, it rectifies the biases over natural dust sources like the 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 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.

4.1.5 Sensitivity to Observational Uncertainties

As discussed in Section 3.1, using satellite retrievals as the learning target inherently absorbs CALIOP's systematic uncertainties. To quantify how this observational limitation impacts the reliability of our framework, we conduct a perturbation-based sensitivity analysis (detailed in Sect. S13). We retrain the GC-TF model by artificially injecting a $\pm$5% systematic perturbation into the CALIOP AEC learning target.

Table S3 and Figure S17 demonstrate that this systematic perturbation induces only a narrow envelope of variation in the corrected AEC profiles. The absolute shift in the Mean Bias fluctuates tightly between 0.001 and 0.004 km^{-1}, and the perturbed predictive RMSE (0.040 km^{-1}) consistently outperforms the original GEOS-Chem simulation (0.052 km^{-1}) by a large margin. This confirms that while observational uncertainties theoretically bound the absolute precision, the physics-informed Transformer does not uncontrollably amplify these errors, ensuring the robustness of the data-driven correction.

4.1.6 Methodological Benchmarking and Structural Necessity

To justify the architectural complexity and isolate the sources of performance gains, we conduct comprehensive benchmarking and ablation studies using the independent 2017 test dataset (Table S5). When trained with identical GEOS-Chem and MERRA-2 predictors, the proposed Transformer significantly outperforms conventional machine learning baselines. A standard Multilayer Perceptron (MLP) and a 1-Dimensional Convolutional Neural Network (1D-CNN) yielded substantially lower R (R=0.083 and 0.540, respectively) compared to the Transformer (R=0.666). This performance gap confirms that global sequence modeling via self-attention is critical for capturing the long-range vertical coupling of atmospheric aerosols—such as boundary layer-to-free troposphere exchange—which localized convolutions or point-wise networks fail to resolve.

Furthermore, ablation experiments confirm that the performance enhancements are intrinsically linked to our structural designs. Removing the Gated Feature Fusion or

613 the Cross-Attention module noticeably degrades predictive accuracy (Table S5). More
614 importantly, beyond statistical improvements, these modules are physically
615 indispensable. They transition the framework from a black-box predictor into a
616 diagnostic tool, providing the explicit attention weights necessary to quantify height-
617 dependent physical drivers (Sect. 4.5.1) and surface environmental modulations (Sect.
618 4.5.3).

619 **4.2 Evaluation of Corrected GEOS-Chem Simulations**

620 **4.2.1 Correction of Systematic Biases in AEC Magnitude**

621 To quantify the efficacy of the GEOS-Chem corrected by Transformer (GC-TF)
622 model in correcting the magnitude of the AEC, we first compare the overall statistical
623 relationship between simulated values and CALIOP observations across three
624 independent test years from 2017 to 2019 (Fig. 6). The original GEOS-Chem simulation
625 AEC exhibits a dispersed distribution around the 1:1 identity line, with R ranging only
626 from 0.50 to 0.53 and RMSE remaining high at 0.052–0.055 km^{-1}. Notably, the low
627 linear regression slopes (0.46–0.51) of the original simulation indicate a tendency to
628 underestimate aerosol extinction intensity under high-loading conditions. In contrast,
629 after correction by the GC-TF model, the AEC data points converge significantly
630 toward the 1:1 line. The R for AEC improves to 0.66–0.73, the RMSE decreases by
631 approximately 25% (to 0.039–0.042 km^{-1}), and the regression slope recovers to 0.55–
632 0.60. These results demonstrate that the framework effectively reduces random biases
633 and realigns the dynamic range of simulated extinction with observations.

634 Despite these substantial statistical improvements, visual scatter remains in the
635 density plots. To rigorously quantify these discrepancies, an error envelope of ± 0.15
636 km^{-1} is introduced in Figure 6(d–f). Statistical analysis indicates that outliers exceeding
637 this threshold account for only 1.20% of the total dataset. Further diagnostic analysis
638 (detailed in Sect. S10) reveals that these extreme deviations are not random noise but
639 exhibit distinct spatial clustering over major emission hotspots (e.g., the IGP, the NCP,
640 and the Indochina Peninsula), and are vertically confined within the PBL (<1.5 km

AGL). These residuals are primarily driven by representativeness errors: CALIOP's narrow footprint captures transient, highly concentrated sub-grid aerosol plumes, which are inherently smoothed out during the spatial averaging process across the coarse 2°×2.5° grid of GEOS-Chem. Consequently, the GC-TF model captures the systematic, state-dependent biases of the grid mean, rather than fitting stochastic sub-grid extremes.

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. S9). 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. S9), 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 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. S10) 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. S11–13) and explicit vertical bias profiles (Fig. S14). Rather 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. S11–13a, 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. S14a–c). Additionally, in the Indochina Peninsula, GEOS-Chem exhibits a spurious extinction peak near 0.8 km (Fig. S11d) 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. S11f). 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 precision 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 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) effectively captures 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.

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 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 relies on the universality of atmospheric physical laws. While EA and NA differ vastly in emission

intensities and aerosol regimes, core physicochemical processes governing aerosol lifecycles remain consistent. The GC-TF model recognizes familiar synoptic patterns and successfully transfers these learned physical constraints across continents. This capability is well evidenced by its bidirectional correction over distinct underlying surfaces. Over the Eastern United States, the model mitigates the Eulerian smearing effect (Rastigejev et al., 2010; Eastham and Jacob, 2017). It restores sharp spatial gradients and high-intensity anthropogenic pollution cores. Over marine regions like the Gulf of Mexico, it effectively suppresses spurious background aerosols often driven by excessive transport or sea-salt overestimation. The model dynamically utilizes multidimensional atmospheric variables to physically constrain spatial distributions rather than applying a simple uniform correction factor.

Despite these capabilities, the overall correction performance in NA (R=0.70) remains statistically lower than in the EA training domain (R=0.93). We hypothesize that this performance gap originates from a fundamental domain shift in aerosol composition. The NA atmosphere features lower background concentrations and a significantly higher fraction of biogenic Secondary Organic Aerosols (SOA) (Goldstein et al., 2009). The optical properties and hygroscopicity of these organic species differ fundamentally from the sulfate-nitrate-dust mixtures dominating EA (Crawford et al., 2021).

To quantitatively verify this impact, we further evaluate the model performance stratified by CALIOP aerosol subtypes over NA (Table S4). The results reveal a clear divergence in the model's corrective capability across distinct aerosol regimes. For dust-dominated regimes, the model effectively mitigates GEOS-Chem's systematic underestimation, enhancing the regression slope from 0.21 to 0.32 and increasing R from 0.41 to 0.50. This confirms that the physical constraints governing dust extinction and vertical transport learned in EA translate reliably to the NA domain. In contrast, the model yields negligible improvements for the SOA-dominated continental regime. Although total mass concentrations are provided as predictors, the specific

809 thermodynamic-to-optical mapping learned in EA—typically dominated by the high
810 hygroscopic growth of inorganic salts—is less applicable to the complex, weakly
811 hygroscopic biogenic SOA prevalent in NA. The stagnant RMSE and R in the SOA
812 group suggest that without locally representative training samples to capture the unique
813 mass extinction efficiencies and refractive indices of NA-specific organic species, the
814 data-driven framework maintains a conservative correction. This ultimately limits the
815 overall precision gain across the NA background.

816 **4.4 Case Studies: Vertical Structure and Temporal Evolution During High Aerosol**
817 **Loading Periods**

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

822 In the Taklimakan Desert (Fig. 11a), the original GEOS-Chem simulation exhibits
823 a nearly vertical extinction profile with extremely low values, indicating that the CTM
824 fails to simulate the significant aerosol extinction layer observed by CALIOP. In
825 contrast, the GC-TF model effectively recovers the extinction signal in this region by
826 leveraging multidimensional input atmospheric state variables (including
827 meteorological fields and background chemical composition). Its profile shape shows
828 superior agreement with satellite observations, largely retrieving the "missing signal"
829 of the original simulation. In Kanpur, a region dominated by anthropogenic pollution
830 (Fig. 11b), the original model fails to accurately capture the magnitude and vertical
831 gradient of the pollution layer below 3.5 km. The GC-TF model, while correcting
832 baseline errors, reshapes the vertical distribution pattern, yielding a vertical decrement
833 trend that aligns closely with observational reality. Figure 11c reveals critical structural
834 biases during a biomass burning event in the Indochina Peninsula. The original GEOS-
835 Chem simulation presents an "inverted" vertical gradient, where the extinction peak
836 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. A critical concern regarding DL models trained on polar-orbiting satellite data is the potential overfitting to the sensor's limited twice-daily sampling timing (e.g., ~01:30 and 13:30 local time for CALIOP). However, the continuous AOD time series output by the GC-TF model demonstrates superior trend consistency with the high-frequency AERONET observations throughout all daylight hours. Figure 11d-e show the corrected model successfully captures the dynamic diurnal evolution and phase fluctuations of aerosols. Specifically, 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. In contrast, the GC-TF model closely tracks these dynamic variations. 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). This empirical evidence confirms that by conditioning the bias correction on meteorology-driven diurnal processes (as discussed in Sect. 3.1), the framework robustly generalizes across the

entire diurnal cycle and avoids overfitting to specific CALIOP overpass times.

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 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 underlying physical mappings learned by the correction framework, we employ a hierarchical diagnostic approach to quantify feature contributions. 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 Domain-Wide Reliance and Altitude-Dependent Prioritization

To understand how the model resolves 3D aerosol fields, we first examine its overarching reliance on different information streams. Domain-wide permutation feature importance (Fig. S15) unambiguously establishes the foundational role of physical priors: randomly permuting the physicochemical profile induces a dramatic 196.1% increase in MSE. This confirms that the physicochemical profiles provided by the CTM remain the indispensable physical foundation, determining the absolute Magnitude of the AEC. In contrast, spatial, temporal, and height components function as modulating variables, refining this baseline across diverse environmental regimes.

Further analysis of the gated feature fusion weights (Fig. 12) reveals that the model effectively adapts its prioritization mechanism based on atmospheric stratification. In the near-surface layer (<0.5 km), the model assigns the highest weight to physicochemical profiles, aligning with the physical reality that extinction near the surface is predominantly controlled by local emissions and immediate thermodynamic

states (Jiang et al., 2024; Xiong et al., 2025). As height increases to the boundary layer transition zone (0.5–1.0 km), a strategic shift occurs. The reliance on physicochemical profiles diminishes while the weight of spatial coordinates increases significantly. This region typically corresponds to the entrainment zone or the top of the PBL, where CTMs are prone to vertical diffusion errors (Eastham and Jacob, 2017; Rastigejev et al., 2010; Lin and Mcelroy, 2010). The model mitigates these uncertainties by leveraging spatial priors to constrain potential diffusion biases. In the mid-lower troposphere (1.0–3.5 km), the contribution of temporal indices exhibits distinct peaks around 1.5 km and 2.8 km. The 1.5 km peak corresponds to the typical maximum daytime PBLH in EA (Guo et al., 2016; Kim, 2022), while the 2.8 km peak aligns with the active layer for long-range transport (Uno et al., 2009). This demonstrates the model's utilization of temporal cues to capture the diurnal evolution of the PBL and seasonal transport events. In the free troposphere (>3.5 km), aerosol variability is vertically decoupled from 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 decoupling, shifting its strategy to rely heavily on spatial coordinates to constrain background aerosol fields. Explicit height encoding maintains a consistently high contribution throughout the entire column, serving as a critical vertical positioning anchor.

4.5.2 Vertical Attribution of Local State Variables

Following the domain-wide ranking, we utilize gradient-based attribution (Fig. 13) to dissect the specific variables within the VPS driving AEC bias correction across altitudes. These drivers are organized into thermodynamic constraints, particulate compositions, and dynamic factors.

First, thermodynamic variables serve as the primary constraints for rectifying the vertical AEC structure. Temperature (T) consistently acts as the dominant driver within the PBL, with attribution scores exceeding 0.20. This suggests that the model implicitly diagnoses atmospheric stability and vertical lapse rate—factors often mischaracterized

in CTMs—to rectify biases associated with turbulent mixing (Lin and Mcelroy, 2010). Concurrently, Relative Humidity (RH) functions as a critical driver for aerosol optical properties. Its contribution is coupled with explicit hygroscopic growth factors (e.g., AerHygroscopicGrowth_SO_4^{2-}), enabling the model to fine-tune the AEC and correct non-linear hygroscopic parameterization errors under high-humidity conditions (Burgos et al., 2020; Zhai et al., 2021).

Second, particulate mass concentrations act as the fundamental determinants of aerosol loading. In the lower troposphere (<1.5 km), $PM_{2.5}$ and PM_{10} consistently rank among the top drivers. 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 persistent influence aloft. 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 rectify background biases induced by long-range transport.

4.5.3 Synoptic Modulation via Surface Constraints

While the VPS determine the baseline extinction, the cross-attention mechanism enables the model to utilize the SFS to modulate the vertical bias correction (Fig. 14). This process operates through distinct physical pathways.

First, the model employs a robust dynamic representation to constrain transport and mixing efficiency. The consistently high attention weights of 10 m wind components (U10M, V10M) reflect the use of near-surface wind speed as a proxy for synoptic flow patterns. 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 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 vertical partitioning of turbulence drivers. Thermodynamic drivers (HFLUX, LAI) peak at the surface and decay upward, diagnosing surface buoyancy fluxes. In contrast, mechanical drivers (USTAR, Z0M) increase and plateau with height. This distinction implies the model successfully evaluates vertical entrainment potential, assessing whether mechanical shear is sufficient to transport pollutants across capping inversions.

Second, radiative and surface boundary conditions are leveraged to correct parameterization biases. The model senses solar input by distinguishing between direct (PARDR) and diffuse (PARDF) photosynthetically active radiation, capturing variations in photochemistry and secondary aerosol formation (Guenther et al., 2012). Furthermore, snow mass (SNOMAS) emerges as a key predictor in the lowest layers. The model identifies snow-covered surfaces as indicators of strongly stable boundary layer prone to temperature inversions. This allows for the targeted correction of over-dilution biases often found in CTMs (Lin and Mcelroy, 2010; Holtslag et al., 2013), effectively capturing high-concentration signals that are typically artificially smoothed by minimal diffusion threshold constraints.

4.5.4 Regional Heterogeneity of Bias Drivers

To reveal the model's predictive behavior under diverse environmental contexts, we conduct SHAP analysis (Fig. 15) to identify region-specific correction patterns adapted to distinct underlying surfaces and emissions.

In regions dominated by anthropogenic and biomass burning emissions (NCP, IGP, and Indochina), the model leverages radiative components to diagnose atmospheric turbidity. A striking commonality is the high ranking of diffuse radiation (PARDF),

often surpassing the direct component (PARDR). This reflects 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). Furthermore, the model differentiates surface energy partitioning: 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.

In the dust-dominated Taklamakan Desert, the model captures a coupled mechanism driven by thermodynamic instability and dynamic uplift. Incident shortwave flux (SWGDN) and direct radiation (PARDR) play a dominant role, indicating that the model identifies clear-sky, high-solar-input conditions as prerequisites for thermal instability. Crucially, this is coupled with dynamic descriptors: combining 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 confirms the model has learned the physical prerequisites for wind-driven dust emission in arid regions.

Finally, over the marine environment of the Western Pacific, the model shifts to a latent heat-driven mode. Latent heat flux (EFLUX) defining the moisture supply at the air-sea interface, controlling marine aerosol hygroscopicity. Additionally, the model captures sea spray generation by linking near-surface wind speeds (V10M) with land-sea masking indicators (GRN), recognizing that mechanically generated sea salt aerosols (Grythe et al., 2014) and their subsequent hygroscopic evolution are the primary drivers of AEC variability over open water.

4.6 Diagnostic Insights for Refining GEOS-Chem Parameterizations

The interpretability analysis in Section 4.5 demonstrates that the GC-TF model captures physically meaningful relationships rather than merely fitting statistical noise. While directly translating data-driven feature sensitivities into concrete code

modifications remains challenging without further sensitivity simulations, this transparency allows the framework to serve as a valuable hypothesis-generation tool. It highlights potential structural uncertainties in CTMs and points toward targeted refinements in physical parameterizations.

4.6.1 Diagnosing Thermodynamic Parameterization Deficiencies

The model heavily relies on temperature and HFLUX to correct AEC profiles (Sect. 4.5.2, 4.5.4), which suggests potential uncertainties in diagnosing PBLH and turbulent mixing intensity within the GEOS-Chem non-local boundary layer scheme. Given that HFLUX drives surface buoyancy and directly modulates the vertical eddy diffusion coefficient, the widespread excessive diffusion biases observed in the lower troposphere indicate that the model may overestimate thermal turbulence under certain stability conditions. In highly urbanized regions like the NCP, the acute sensitivity to HFLUX implies that current surface energy balance calculations struggle to resolve the distinct thermodynamic properties of urban canopies. Future model development could benefit from constraining stability functions within the vertical diffusion module, or alternatively, coupling a dedicated urban canopy model to better represent sensible heat partitioning.

4.6.2 Refining Emission and Formation Schemes via Environmental Proxies

The cross-attention weights, which reveal how synoptic forcing modulates vertical aerosol profiles (Sect. 4.5.3, 4.5.4), highlight potentially inadequately parameterized mechanisms in current emission and chemical modules. Over the Taklamakan Desert, the model explicitly pairs greenness fraction with surface wind speed to capture dust extinction. This suggests that the GEOS-Chem dust emission scheme might struggle to accurately parameterize threshold friction velocity over complex bare soils, indicating that the non-linear response of wind-blown dust to surface shear stress and soil erodibility likely requires recalibration. Similarly, high sensitivity to diffuse radiation in the biomass burning region of Indochina points to potentially under-represented SOA formation. Given that high aerosol loading enhances diffuse radiation and alters

photolysis rates, the data-driven model likely leverages diffuse radiation as a proxy for accelerated photochemical aging. This highlights a need to optimize SOA yield parameterizations and refine biomass burning plume injection heights to capture rapid aerosol evolution in dense smoke.

4.6.3 Bridging Data-Driven Interpretation with CTM Development

Beyond statistical bias correction, this study highlights the utility of physics-informed DL for model diagnosis. By decoupling the contributions of meteorology and aerosol composition, the framework verifies that CTMs provide a robust physicochemical baseline, yet exhibit uncertainties in representing the complex, non-linear interactions between aerosols and meteorology. The correction strategies derived from the data-driven model offer valuable diagnostic clues. Identifying specific environmental proxies that govern simulation biases bridges the gap between data-driven retrieval and deterministic modeling, ultimately guiding the targeted integration of neglected physical constraints into future parameterization schemes.

4.7 Model Limitations and Scope of Application

As a supervised bias-correction framework, the model relies on state-dependent mapping, meaning its performance is fundamentally constrained by the predictive signals available in the input features. The framework excels at correcting systematic, parameterization-driven bias. For instance, it successfully restores the underestimated dust plumes in the Taklamakan Desert by leveraging wind speed, clear-sky radiation, and vegetation indices as physical proxies for actual dust emission conditions (Sect. 4.5.4).

However, the model possesses limited capacity to compensate for entirely missing physical processes. If a highly localized or stochastic event is completely absent from the prescribed emission inventory and produces no corresponding anomalies in the input meteorological or chemical precursor fields, the model lacks the necessary physical constraints to capture the resulting aerosol plume. In such scenarios, the correction remains strictly bounded by the prior information provided by the GEOS-

Chem and MERRA2.

5. Conclusions

This study proposes a physics-informed Transformer DL framework that generates high-precision 3D AEC fields by bridging the gap between CTM simulations and CALIOP satellite observations. Functioning as a supervised bias-correction model rather than a DA system, this framework learns a state-dependent mapping to rectify systematic simulation AEC bias. 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 precision of AEC vertical profiles, systematically correcting the systematic biases inherent in Eulerian CTMs. Extensive LOYO 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.

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 advances beyond standard bias correction to serve as a diagnostic framework for physical mechanisms. Attribution analysis reveals that the model identifies AEC simulation bias drivers with clear physical significance: (1) In the PBL, the heavy reliance on temperature and HFLUX highlights potential uncertainties in vertical eddy diffusion coefficients within stability-dependent mixing schemes; (2) Over dust source regions, the paired use of vegetation indices and wind speed suggests the need to recalibrate threshold friction velocity and soil erodibility; (3) In biomass burning regions, the sensitivity to diffuse radiation points to under-represented SOA yields and photochemical aging processes; (4) In marine regions, the utilization of latent heat flux and surface wind implies uncertainties in sea-salt generation functions and hygroscopic growth .

In summary, this study not only confirms the effectiveness of the physics-informed Transformer in deriving high-precision 3D aerosol fields to support radiative forcing assessments but also highlights its potential as a diagnostic tool for CTMs. The interpretability analysis provides developers with actionable physical insights, highlighting a shift in the integration of DL 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 CTMs.

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.

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, Atmos. Chem. Phys., 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, J. Geophys. Res.:Atmos., 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, B., and Smirnov, A.: Observationally constrained analysis of sea salt aerosol in the marine atmosphere, Atmos. Chem. Phys., 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, Mon. Wea. Rev., 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, J. Geophys. Res.:Atmos., 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, Atmos. Chem. Phys., 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, Atmos. Chem. Phys., 18, 405-425, 10.5194/acp-18-405-2018, 2018.

Chen, B., Song, Z., Pan, F., and Huang, Y.: Obtaining vertical distribution of PM2.5 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, Environ. Res. Lett., 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 Geosci. Remote Sensing Lett., 22, 10.1109/lgrs.2025.3601046, 2025.

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, Atmos. Meas. Tech., 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, Geophys. Res. Lett., 33, 10.1029/2005gl025278, 2006.

Choi, M., Park, J., Sung, M., and Ying, Q.: Long-Range Transport of Secondary Inorganic Aerosol from China to South Korea, Environ. Sci. Technol. Lett., 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, Proc. Natl. Acad. Sci. U.S.A., 117, 17591-17598, 10.1073/pnas.1921231117, 2020.

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, Atmos. Meas. Tech., 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, Elem. Sci. Anth., 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.

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, J. Appl. Remote Sens, 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), J. Geophys. Res.:Atmos., 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, Atmos. Meas. Tech., 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 Trans. Geosci. 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 PM2.5 concentrations over East China with WRF-Chem: impacts from boundary-layer

mixing and anthropogenic emission, Atmos. Chem. Phys., 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 satellite observations, Atmos. Meas. Tech., 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, Atmos. Chem. Phys., 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. J. Photogramm. Remote. Sens., 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, Atmos. Chem. Phys., 14, 159-174, 10.5194/acp-14-159-2014, 2014.

Geer, A. J.: Learning earth system models from observations: machine learning or data assimilation?, Philos. Trans. R. Soc. A Math. Phys. Eng., 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), J. 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, Atmos. Meas. Tech., 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 (GFED4), J. Geophys. Res. Biogeosci., 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, Atmos. Meas. Tech., 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, Proc. Natl. Acad. Sci. U.S.A., 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, Atmos. Chem. Phys., 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, Geosci. Model Dev., 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, Atmos. Environ., 140, 311-319, 10.1016/j.atmosenv.2016.05.046, 2016.

Handschuh, J., Erbertseder, T., Schaap, M., and Baier, F.: Estimating PM2.5 surface concentrations from AOD: A combination of SLSTR and MODIS, Remote. Sens. Appl., 26, 10.1016/j.rsase.2022.100716, 2022.

Henze, D. K., Hakami, A., and Seinfeld, J. H.: Development of the adjoint of GEOS-Chem, Atmos. Chem. Phys., 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 PM2.5 precursor emissions using the adjoint of GEOS-Chem, Atmos. Chem. Phys., 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), Geosci. Model Dev., 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 Sens. Environ., 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, Bull. Am. Meteorol. Soc., 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 Space Sci., 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 PM2.5 concentrations in the Southeastern United States using MAIAC AOD retrievals and a two-stage model, Remote Sens. Environ., 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, Atmos. Chem. Phys., 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, J. Geophys. Res.:Atmos., 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, J. Atmos. Oceanic Technol., 26, 1214-1228, 10.1175/2009jtecha1223.1, 2009.

Ichoku, C., Chu, D. A., Mattoo, S., Kaufman, Y. J., Remer, L. A., Tanré, D., Slutsker, I., and Holben, B. N.: A spatio-temporal approach for global validation and analysis of MODIS aerosol products, Geophys. Res. Lett., 29, 10.1029/2001gl013206, 2002.

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, J. Geophys. Res.:Atmos., 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 validation based on 2 years of coincident Aerosol Robotic Network (AERONET) observations - art. no. D10S04, J. Geophys. Res.:Atmos., 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, Atmos. Meas. Tech., 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 Appl. Sci., 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, Atmos. Meas. Tech., 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 with the GEOS-Chem chemical transport model, Atmos. Chem. Phys., 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, Atmos. Chem. Phys., 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, J. Geophys. Res.:Atmos., 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, J. Geophys. Res.:Atmos., 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.

Li, T. W., Shen, H. F., Yuan, Q. Q., Zhang, X. C., and Zhang, L. P.: Estimating Ground-Level PM2.5 by Fusing Satellite and Station Observations: A Geo-Intelligent Deep Learning Approach, Geophys. Res. Lett., 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, Sci. Total Environ., 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, Atmos. Environ., 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, J. Geophys. Res.:Atmos., 117, 10.1029/2011jd016490, 2012.

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 factors, Atmos. Res., 240, 10.1016/j.atmosres.2020.104891, 2020.

Lu, Z., Wang, J., Chen, X., Xu, X., Zhou, M., Fu, D., and Jiang, J. H.: First Retrieval of Aerosol Vertical Profile With Passive Remote Sensing: Part 1. Development of Algorithm Theoretical Basis, J. Geophys. Res.:Atmos., 130, 10.1029/2025jd044332, 2025a.

Lu, Z. D., Wang, J., Chen, X., Xu, X. G., Zhou, M., Fu, D. J., and Jiang, J. H.: First Retrieval of Aerosol Vertical Profile With Passive Remote Sensing: Part 1. Development of Algorithm Theoretical Basis, J. Geophys. Res.:Atmos., 130, 10.1029/2025jd044332, 2025b.

Lundberg, S. M. and Lee, S. I.: A Unified Approach to Interpreting Model Predictions, 31st Annual Conference on Neural Information Processing Systems (NIPS), Long

Beach, CA, Dec 04-09 2017, WOS:000452649404081, 2017.

Lv, B., Hu, Y., Chang, H. H., Russell, A. G., and Bai, Y.: Improving the Accuracy of Daily PM2.5 Distributions Derived from the Fusion of Ground-Level Measurements with Aerosol Optical Depth Observations, a Case Study in North China, Environ. Sci. Technol., 50, 4752-4759, 10.1021/acs.est.5b05940, 2016.

McDuffie, E. E., Smith, S. J., O'Rourke, P., Tibrewal, K., Venkataraman, C., Marais, E. A., Zheng, B., Crippa, M., Brauer, M., and Martin, R. V.: A global anthropogenic emission inventory of atmospheric pollutants from sector- and fuel-specific sources (1970–2017): an application of the Community Emissions Data System (CEDS), Earth Syst. Sci. Data, 12, 3413-3442, 10.5194/essd-12-3413-2020, 2020.

Mehta, S. K., Ananthavel, A., Velu, V., Prabhakaran, T., Pandithurai, G., and Rao, D. N.: Characteristics of elevated aerosol layer over the Indian east coast, Kattankulathur (12.82°N, 80.04°E): A northeast monsoon region, Sci. Total Environ., 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, J. Appl. Meteorol. Climatol., 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, Atmos. Environ., 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, J. Land Use Sci., 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, Atmos. Chem. Phys., 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., 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, Atmos. Chem. Phys., 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, Atmos. Meas. Tech., 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, Atmos. Meas. Tech., 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, Atmos. Chem. Phys., 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, Atmos. Meas. Tech., 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, Atmos. Chem. Phys., 16, 907-925, 10.5194/acp-16-907-2016, 2016.

Rastigejev, Y., Park, R., Brenner, M. P., and Jacob, D. J.: Resolving intercontinental pollution plumes in global models of atmospheric transport, J. Geophys. Res.:Atmos., 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, Atmos. Chem. Phys., 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 O2 A band observations from GOME-2A, Atmos. Meas. Tech., 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, Atmos. Chem. Phys., 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 Res., 2, 181-204, 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, Sci. Total Environ., 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, J. Geophys. Res.:Atmos., 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, Environ. Sci. Technol., 59, 6379-6386, 10.1021/acs.est.5c00751, 2025.

Stier, P., Seinfeld, J. H., Kinne, S., and Boucher, O.: Aerosol absorption and radiative forcing, Atmos. Chem. Phys., 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, Atmos. Environ., 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., Wang, Z. F., Hara, Y., and Sugimoto, N.: Asian dust transported one full circuit around the globe, Nat. Geosci., 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, Atmos. Chem. Phys., 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, Geophys. Res. Lett., 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 SO2 plume height data and comparison with MODIS and CALIOP observations, Atmos. Chem. Phys., 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 Sens. Environ., 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)

observations across the Pacific, J. Geophys. Res.:Atmos., 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 SO2 and NOx emissions over China using multisensor satellite data - Part 1: Formulation and sensitivity analysis, Atmos. Chem. Phys., 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 SO2 and NOx emissions over China using multisensor satellite data - Part 2: Downscaling techniques for air quality analysis and forecasts, Atmos. Chem. Phys., 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, Atmos. Chem. Phys., 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 PM1 Concentrations from 2014 to 2018 across China, Environ. Sci. Technol., 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 Transport and Aerosol–Cloud-Interaction Experiment: Overview and Selected Highlights, Bull.

Am. Meteorol. Soc., 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, Atmos. Chem. Phys., 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, Atmos. Chem. Phys., 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, J. Atmos. Oceanic Technol., 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, Bull. Am. Meteorol. Soc., 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, Atmos. Res., 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, Atmos. Res., 10.1016/j.atmosres.2025.108348, 2025.

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, Atmos. Chem. Phys., 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.,

1649 Weckwerth, T. M., Li, Z., Campbell, J. R., Yang, P., Diao, M., Noel, V., Meyer, K.
1650 G., Carr, J. L., Garay, M., Christian, K., Bennedetti, A., Ring, A. M., Crawford, A.,
1651 Pavolonis, M. J., Aquila, V., Kim, J., and Kondragunta, S.: A SmallSat Concept to
1652 Resolve Diurnal and Vertical Variations of Aerosols, Clouds, and Boundary Layer
1653 Height, Bull. Am. Meteorol. Soc., 104, E815-E836, 10.1175/bams-d-21-0179.1,
1654 2023.
1655 Zarzycki, C. M. and Bond, T. C.: How much can the vertical distribution of black carbon
1656 affect its global direct radiative forcing?, Geophys. Res. Lett., 37,
1657 10.1029/2010gl044555, 2010.
1658 Zeng, Y., Wang, M., Zhao, C., Chen, S., Liu, Z., Huang, X., and Gao, Y.: WRF-Chem
1659 v3.9 simulations of the East Asian dust storm in May 2017: modeling sensitivities
1660 to dust emission and dry deposition schemes, Geosci. Model Dev.s, 13, 2125-2147,
1661 10.5194/gmd-13-2125-2020, 2020.
1662 Zhai, S. X., Jacob, D. J., Brewer, J. F., Li, K., Moch, J. M., Kim, J., Lee, S., Lim, H.,
1663 Lee, H. C., Kuk, S. K., Park, R. J., Jeong, J., Wang, X., Liu, P. F., Luo, G., Yu, F.
1664 Q., Meng, J., Martin, R., Travis, K. R., Hair, J. W., Anderson, B. E., Dibb, J. E.,
1665 Jimenez, J. L., Campuzano-Jost, P., Nault, B. A., Woo, J. H., Kim, Y., Zhang, Q.,
1666 and Liao, H.: Relating geostationary satellite measurements of aerosol optical
1667 depth (AOD) over East Asia to fine particulate matter (PM2.5): insights from the
1668 KORUS-AQ aircraft campaign and GEOS-Chem model simulations, Atmos.
1669 Chem. Phys., 21, 16775-16791, 10.5194/acp-21-16775-2021, 2021.
1670 Zhang, J. L., Campbell, J. R., Reid, J. S., Westphal, D. L., Baker, N. L., Campbell, W.
1671 F., and Hyer, E. J.: Evaluating the impact of assimilating CALIOP-derived aerosol
1672 extinction profiles on a global mass transport model, Geophys. Res. Lett., 38,
1673 10.1029/2011gl047737, 2011.
1674 Zhao, B., Wang, Y., Gu, Y., Liou, K.-N., Jiang, J. H., Fan, J., Liu, X., Huang, L., and
1675 Yung, Y. L.: Ice nucleation by aerosols from anthropogenic pollution, Nat. Geosci.,
1676 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, Sci. Total Environ., 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 PM2.5 and aerosol optical depth, Atmos. Chem. Phys., 24, 11565-11584, 10.5194/acp-24-11565-2024, 2024.

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

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

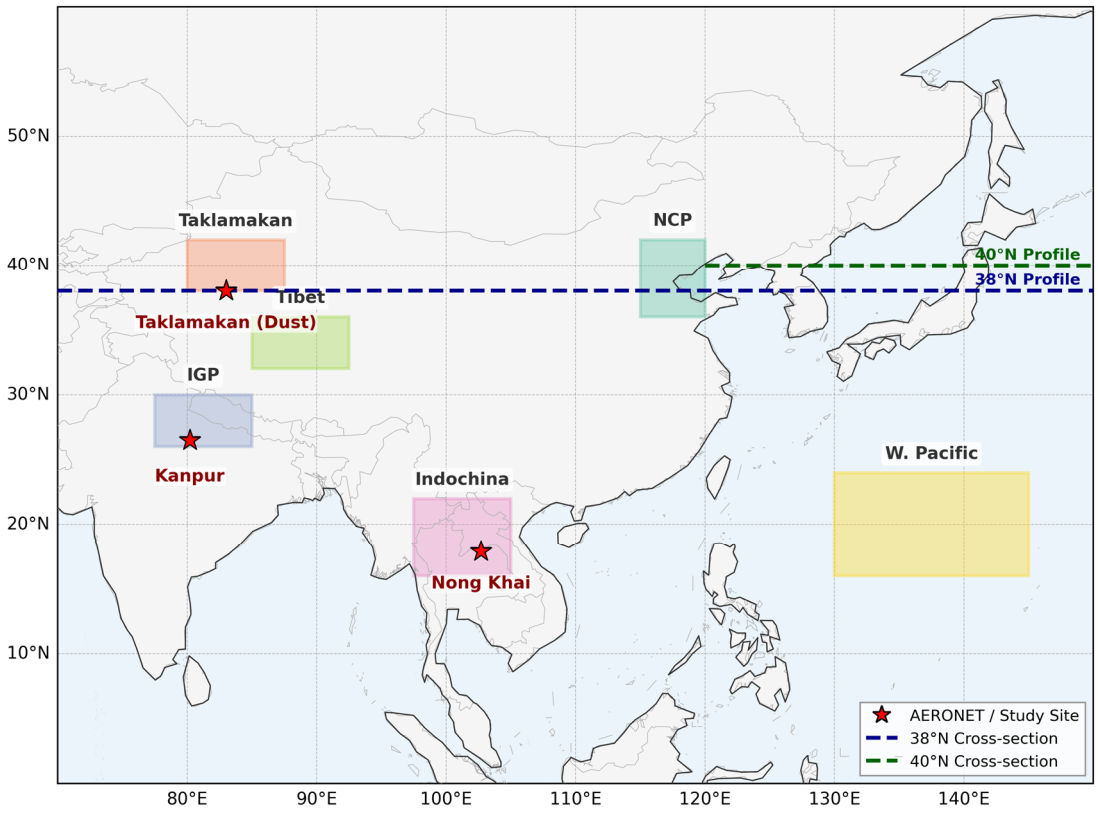

Figure 1. Map of the study domain covering East and South Asia (0°-60°N, 70°-150°E). The colored rectangles delimitate the six regions of interest selected for regional analysis. The dashed lines represent the latitudinal transects at 38°N (blue) and 40°N (green) used for vertical profile cross-section 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.

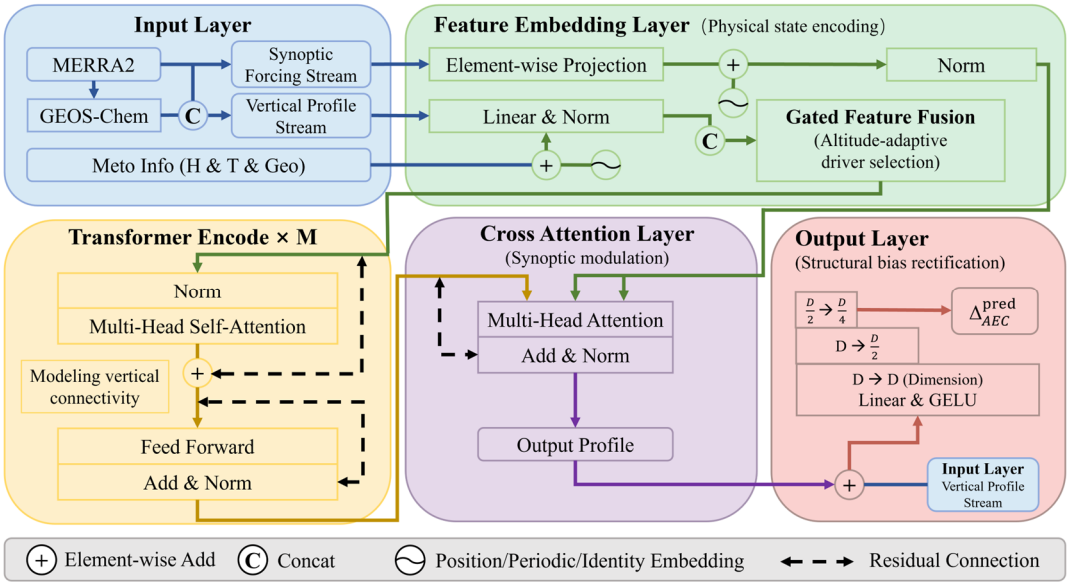

Figure 2. Architecture of the physics-informed Transformer framework. The Feature Embedding Layer integrates the VPS and SFS into unified, high-dimensional physical states. Gated Feature Fusion resolves atmospheric stratification by dynamically prioritizing altitude-dependent drivers. The Transformer Encoder stack models vertical connectivity, simulating exchange processes like turbulent mixing to ensure profile coherence. The Cross-Attention Layer retrieval synoptic constraints to modulate AEC corrections based on the meteorological background. The Output Layer performs residual bias rectification, anchoring the prediction to the initial physicochemical profiles to accurately quantify systematic simulation biases while filtering redundant noise.

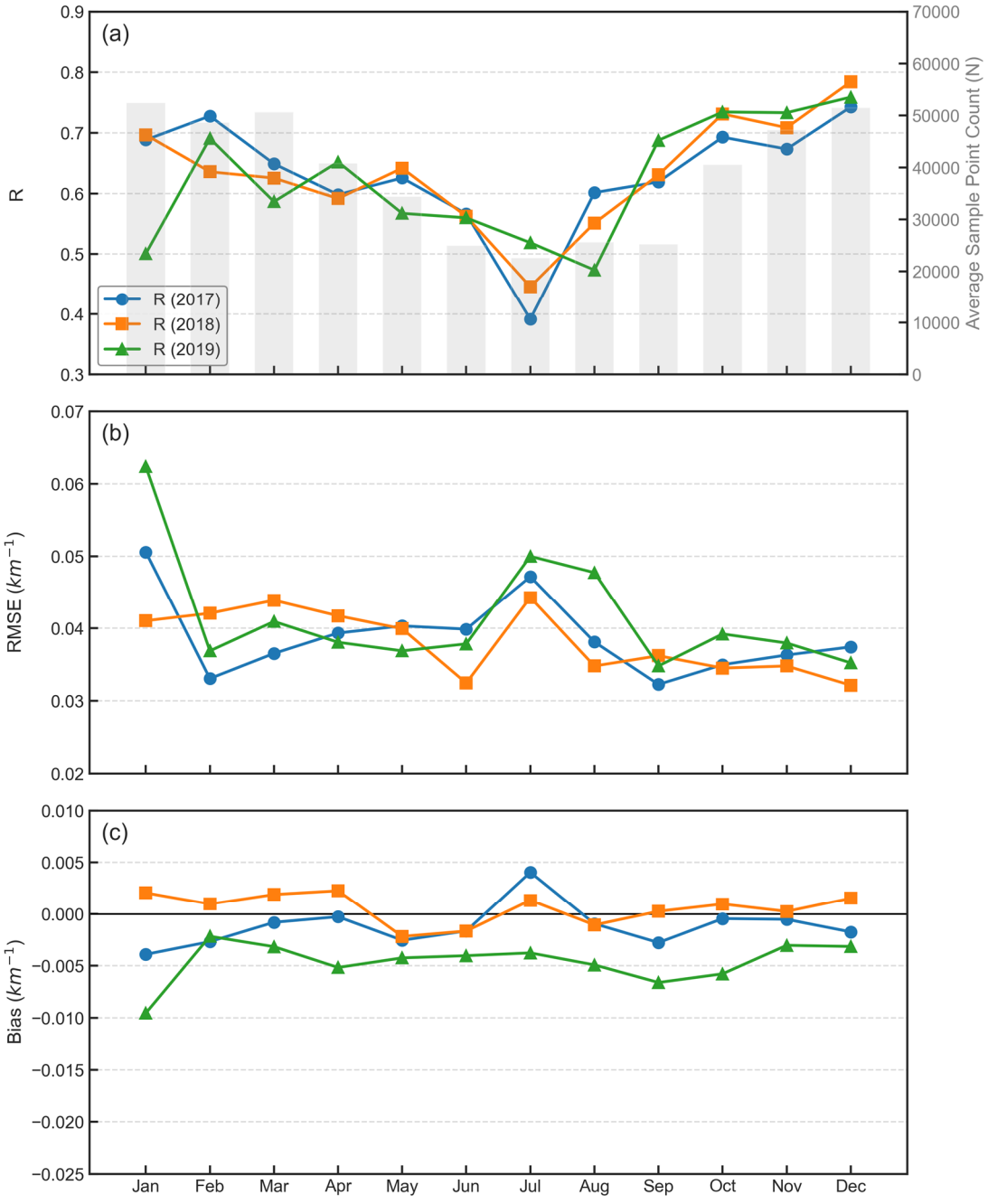

Figure 3. Monthly variations of statistical metrics evaluating the physics-informed Transformer model's predictive performance over EA across the three test years (2017, 2018, and 2019). The panels display the time series of R between the AEC bias predicted by the model and that simulated by GEOS-Chem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

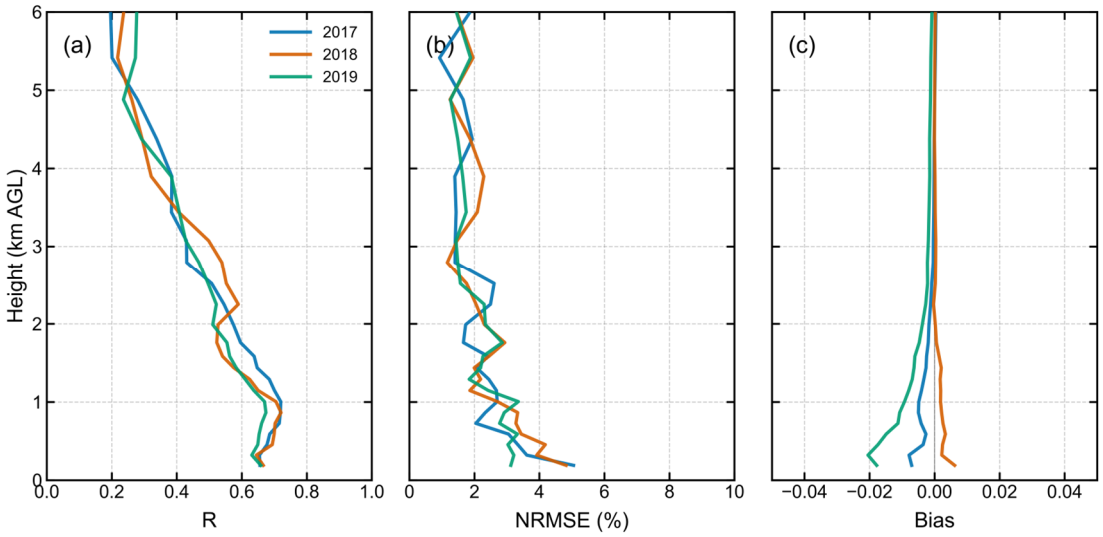

Figure 4. Vertical profiles of the physics-informed Transformer model's predictive performance for AEC simulation bias over EA across three independent test years (2017, 2018, and 2019). The panels show the height-resolved R (a), NRMSE (b), and mean bias (c). The performance is evaluated against CALIOP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

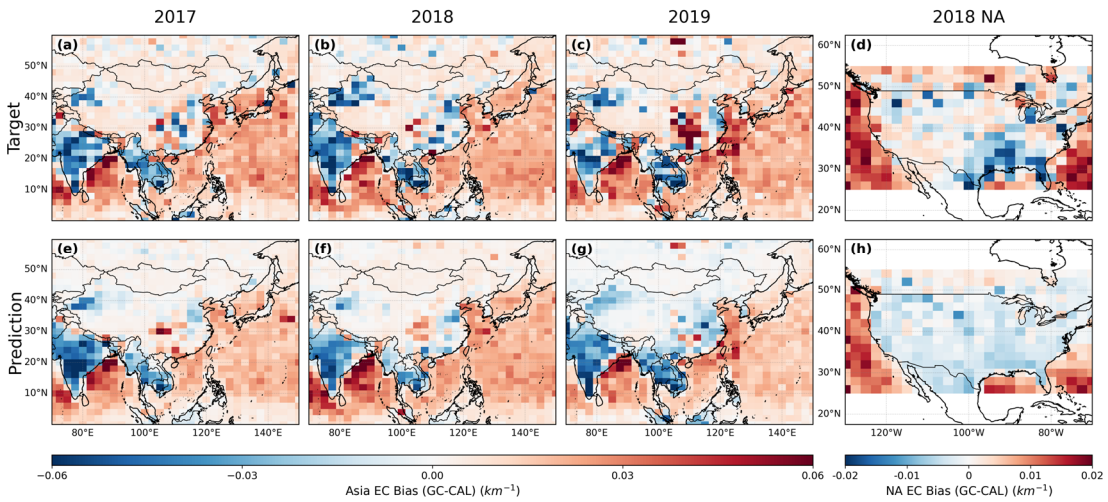

Figure 5. Spatial comparison of the vertically averaged (0–6 km AGL) AEC simulation bias over EA and NA. The top row (a–d) represents the Target systematic bias. The bottom row (e–h) displays the corresponding bias Predicted by the physics-informed Transformer model. Columns 1–3 show results for the 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.

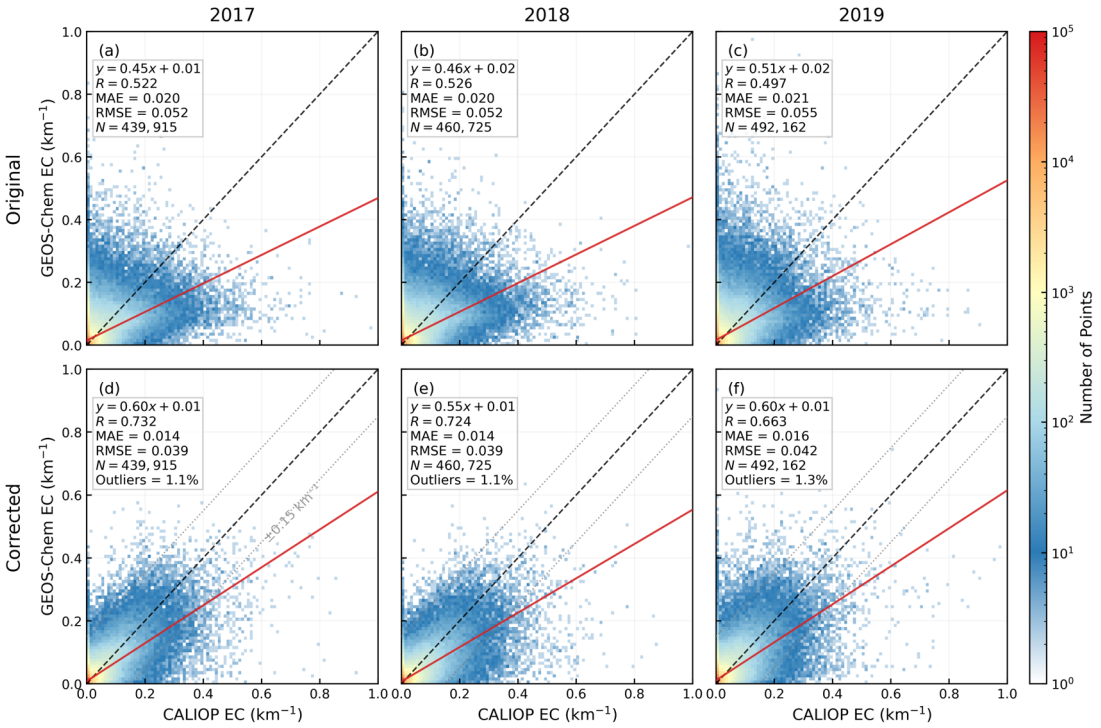

Figure 6. Density scatter plots comparing the simulated AEC against CALIOP observations over EA for the three test years: 2017 (a, d), 2018 (b, e), and 2019 (c, f). The top row (a, b, c) displays the validation results for the original GEOS-Chem simulation, while the bottom row (d, e, f) shows the results after correction by the physics-informed Transformer model. The dashed gray lines in the bottom panels (d–f) delineate the ± 0.15 km^{-1} error envelope, with the corresponding percentage of outliers (points falling outside this envelope) indicated in the statistical boxes.

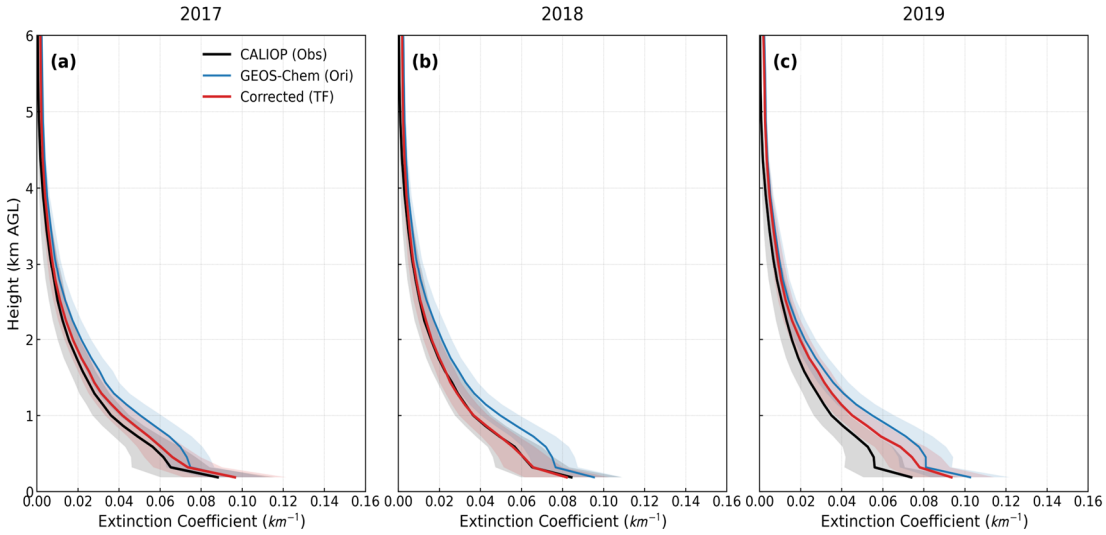

Figure 7. Annual mean vertical profiles of AEC for the three test years (2017–2019). The profiles are averaged over the entire study domain (EA). The shaded areas indicate the standard deviation ($\pm 0.5\sigma$) of the vertical distribution, representing the spatial and temporal variability.

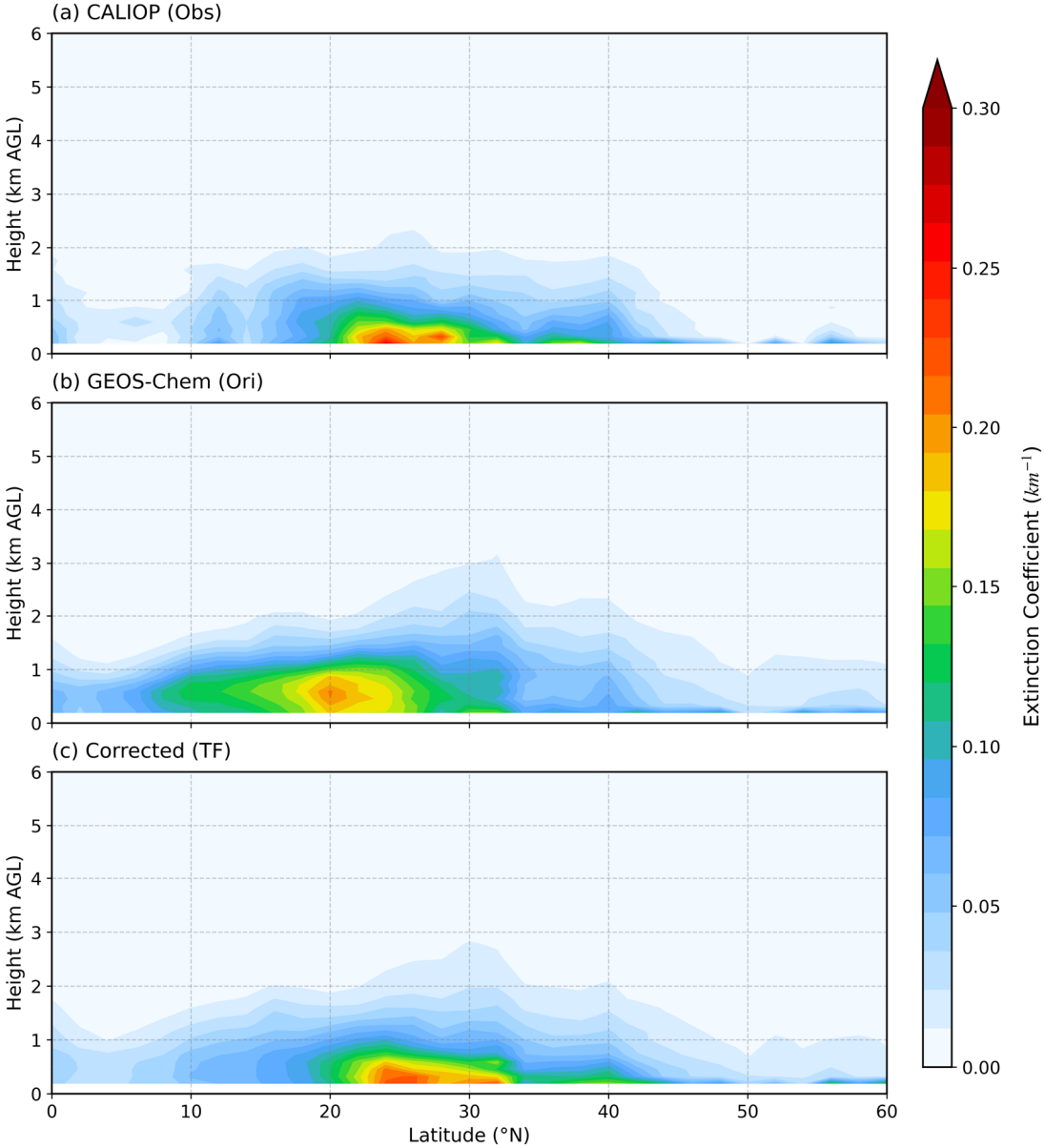

Figure 8. Zonal mean vertical distributions of AEC averaged over the longitude range (70°E–150°E) for the test year 2019. The panels display the latitude-altitude cross-sections for CALIOP observations (a), original GEOS-Chem simulations (b), and the corrected GC-TF results (c).

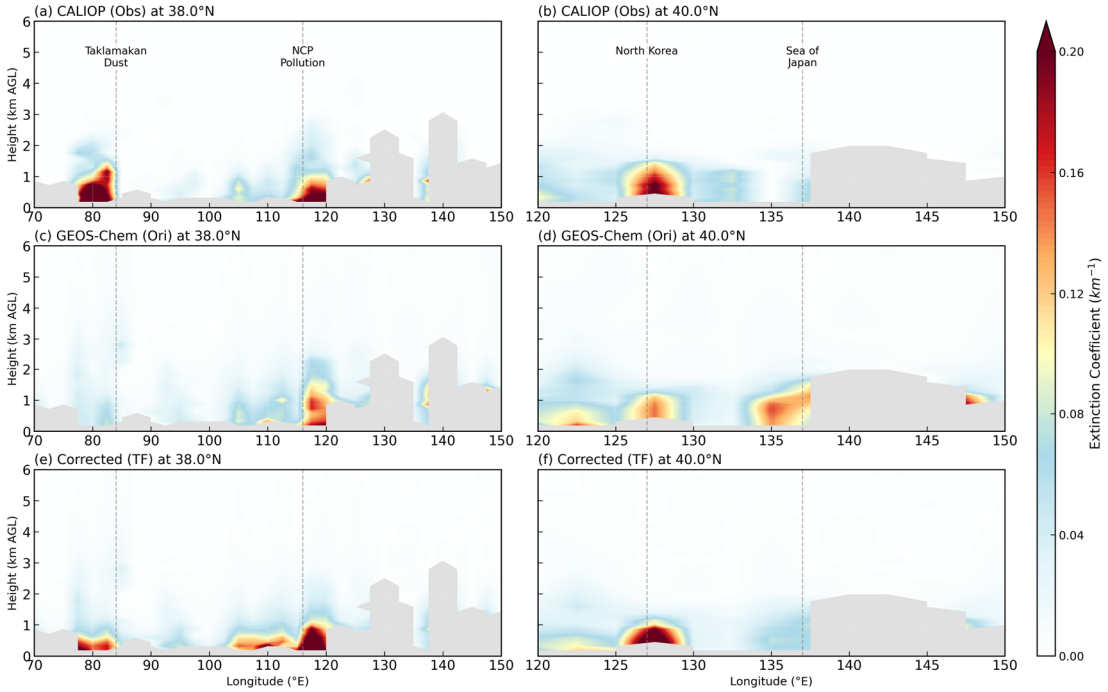

Figure 9. Longitudinal vertical cross-sections of annual mean AEC along two selected latitudes for the test year 2019. The left column displays the profiles at 38.0°N, highlighting the Taklamakan Desert and the NCP. The right column displays the profiles at 40.0°N, capturing the aerosol transport over 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.

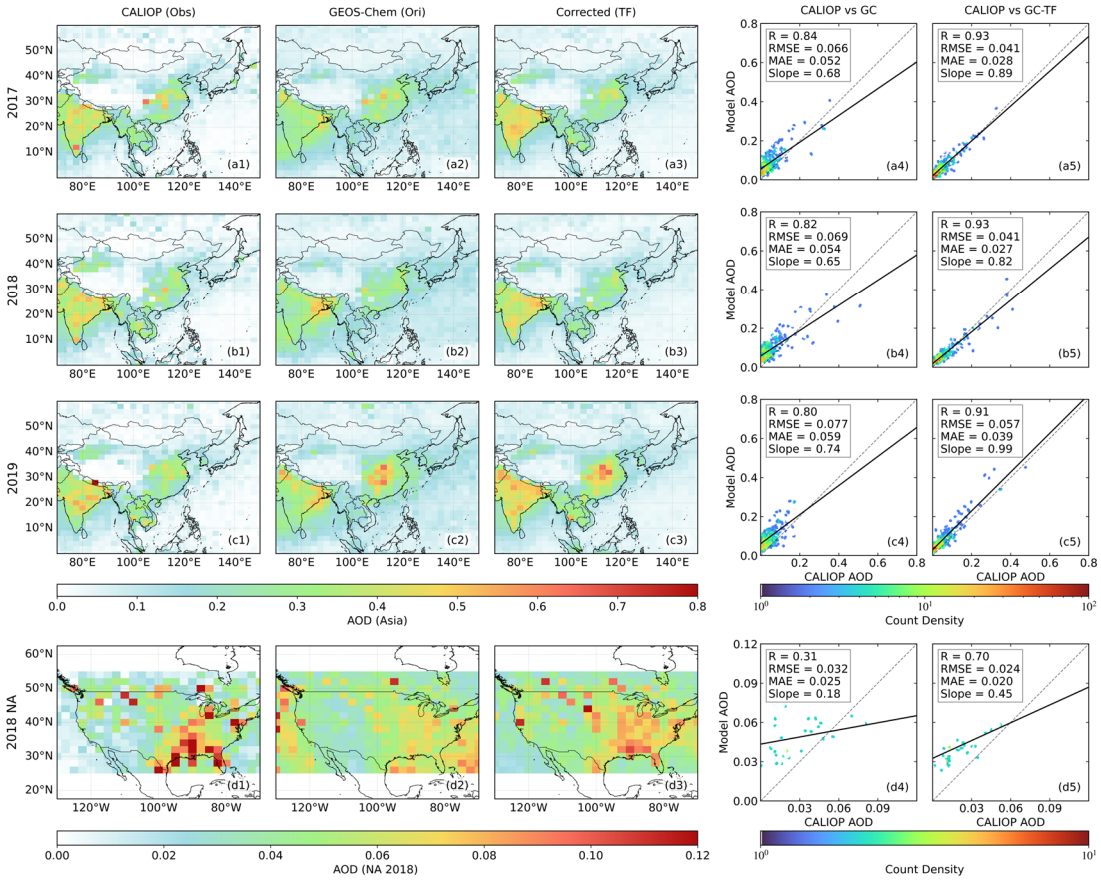

Figure 10. Spatial distributions and statistical evaluations of the annual mean AOD derived from CALIOP observations, GEOS-Chem simulations, and the GC-TF model corrections. Results for the EA domain during the test years 2017 (a), 2018 (b), and 2019 (c), respectively. Columns 1–3 display the spatial 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).

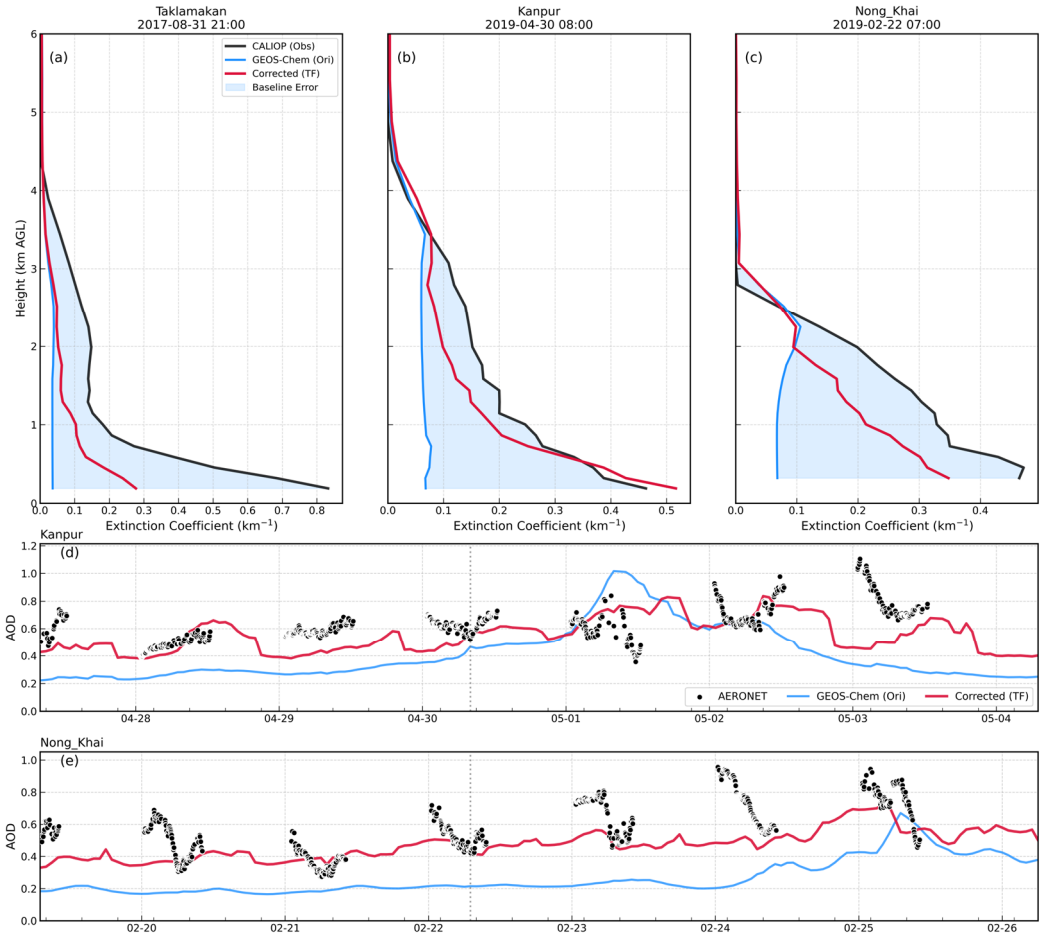

Figure 11. Composite analysis of aerosol vertical structures and temporal evolution during selected pollution episodes. Vertical profiles of AEC at three representative sites: Taklamakan (Dust, a), Kanpur (Anthropogenic Pollution, b), and Nong Khai (Biomass Burning, c). Time series of AOD at the Kanpur (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.

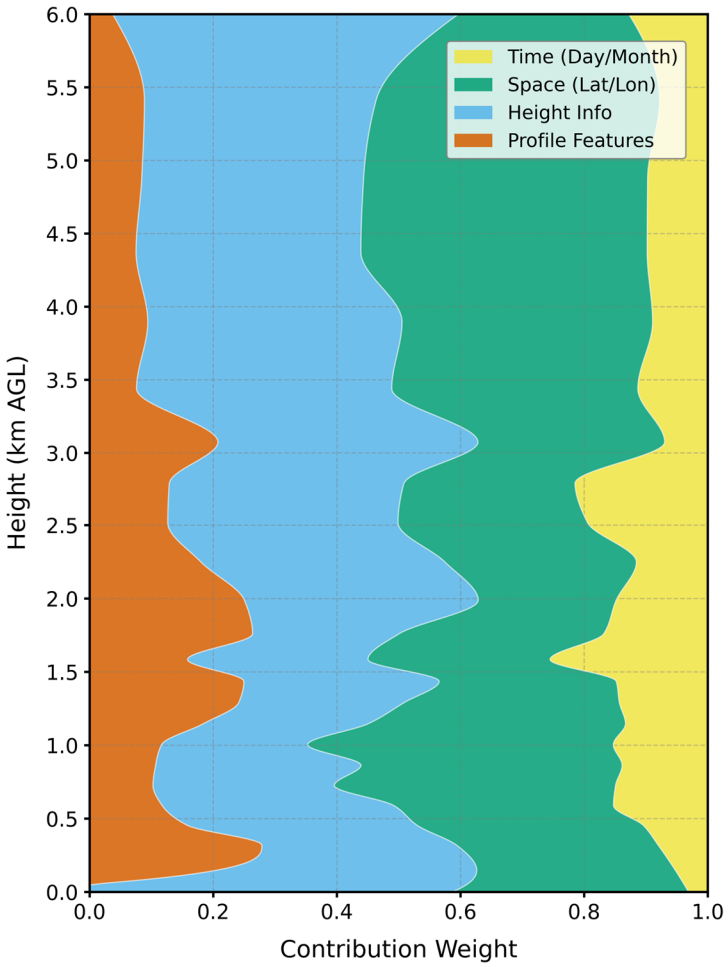

Figure 12. Altitude-dependent prioritization of information within the VPS. The stacked plot displays the learnable contribution weights for the four VPS sub-components—baseline physicochemical profiles, height information, spatial coordinates, and temporal indices—as determined by the gated feature fusion mechanism. This analysis visualizes how the model's internal decision-making adapts to different atmospheric stratifications.

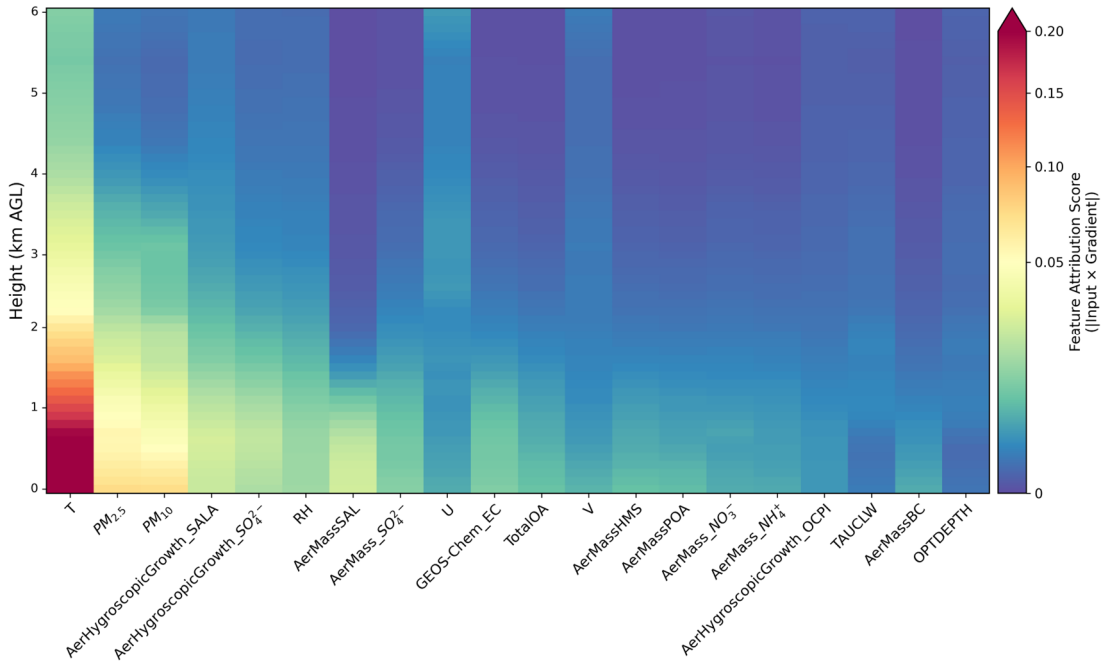

Figure 13. Vertical distribution of cross-attention weights for synoptic forcing variables. The heatmap characterizes the interaction strength between the SFS and the VPS. It reveals how specific variables within the SFS (sorted by total contribution from left to right) modulate AEC bias corrections at different altitude levels.

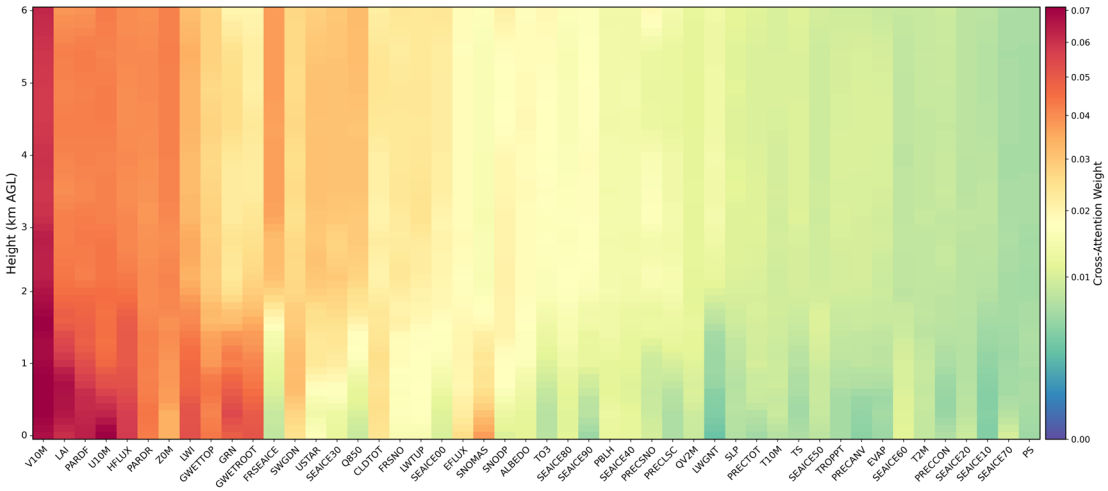

Figure 14. Vertical distribution of the attention weights for global variables learned by the Cross-Attention module. The heatmap displays the interaction strength between the VPS and the SFS. The variables are sorted by their total contribution from left to right.

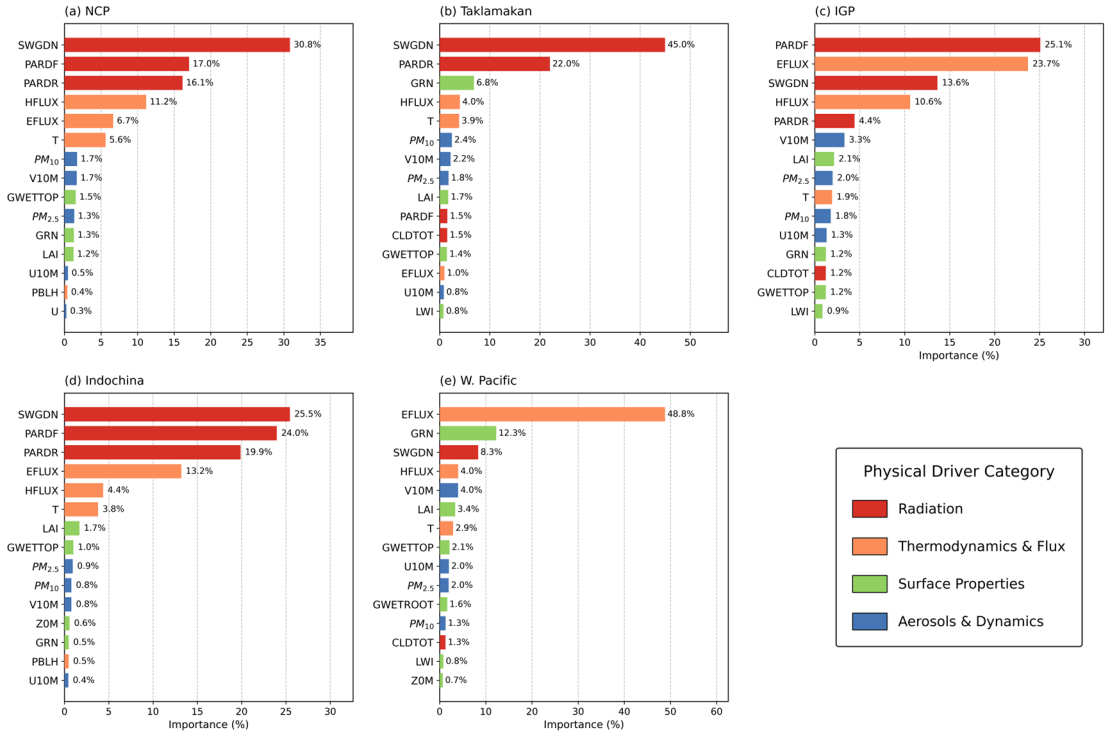

Figure 15. Regional variability in feature importance drivers identified by SHAP analysis for the test year 2019. The panels display the top 15 most influential features for predicting AEC simulation biases in five representative regions: NCP (a), Taklamakan Desert (b), IGP (c), Indochina (d), and Western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