https://doi.org/10.5194/essd-2025-275
Preprint. Discussion started: 11 September 2025

Open Access Earth System Science Data Discussions

MATCHA, a novel regional hydroclimate-chemical reanalysis: System description and evaluation

Chayan Roychoudhury[1], Rajesh Kumar[2], Cenlin He[2], William Y. Y. Cheng[2], Kirpa Ram[3], Naoki Mizukami[2], and Avelino F. Arellano[1]

[1]Department of Hydrology and Atmospheric Sciences, University of Arizona, Tucson, USA
[2]Research Applications Laboratory, NSF National Center for Atmospheric Research, Boulder, USA
[3]Department of Chemistry, Indian Institute of Technology, Varanasi, India

Correspondence to: Chayan Roychoudhury (croychoudhury@arizona.edu)

Abstract. We present MATCHA (Model for Atmospheric Transport and Chemistry in Asia), a 17-year (2003-2019) regional hydroclimate-chemical reanalysis for Asia (58° – 140° E, 4° – 40° N) at 12 km resolution that is based on the Weather Research and Forecasting model coupled with Chemistry (WRF-Chem), Community Land Model (CLM), and SNow, ICe and Aerosol Radiative (SNICAR) model as well as satellite data assimilation to explicitly represent interactions between key atmospheric composition and regional hydroclimate (including aerosol-snowpack interactions) across High Mountain Asia (HMA). Approximately two decades of satellite observations of aerosol optical depth (AOD) from the Moderate Resolution Imaging Spectroradiometer (MODIS) and carbon monoxide (CO) profiles from the Measurement of Pollution in the Troposphere (MOPITT) were assimilated every three hours into WRF-Chem to further constrain the representation of aerosols and chemistry. MATCHA provides comprehensive outputs across different light-absorbing aerosol species, e.g., black carbon (BC), dust, and brown carbon (BrC), trace gases, and a range of meteorological, hydrological, and land-surface variables over the region. This paper describes the MATCHA coupled modeling and data assimilation framework and evaluates 12 key variables across aerosols (fine particulate matter ($PM_{2.5}$/PM_{10}), AOD, single scattering albedo (SSA), and surface BC concentration), trace gases (surface CO), meteorology (precipitation, planetary boundary layer height (PBLH), temperature, relative humidity, and wind speed), and hydrology (snow cover fraction) against available in-situ and satellite observations across Asia. Meteorological fields (surface and vertical profiles) are consistently well-reproduced with Kling-Gupta efficiencies (KGEs) ranging from 0.65 to 1. Notable issues include persistent cold and dry bias over high-elevation regions in winter, with stronger-than-observed surface winds. Snow cover fraction seasonality is well captured with slight underestimation during snowmelt seasons across major glacier regions. Daily accumulated precipitation estimates agree with satellite observations, particularly with the best KGE (0.6) during the monsoon season, albeit underestimated over high-elevation regions. The diurnal and seasonal evolution of PBLH is well-represented, with biases reflecting shallower heights in the morning and deeper heights in the afternoon in summer, likely due to model parameterizations and resolution limitations. MATCHA also captures the spatial and seasonal variability of AOD and SSA at 550 nm, yet overestimates summer AOD over India and southeast Asia with a strong negative bias in SSA. Biases in $PM_{2.5}$/PM_{10} are also higher, which appears to be particularly related to high biases in wind speeds, causing overestimation of natural emissions of aerosols and overestimation

https://doi.org/10.5194/essd-2025-275
Preprint. Discussion started: 11 September 2025

Open Access Earth System Science Data Discussions

of anthropogenic emissions. Comparisons with site-specific aerosol chemical composition derived from air samples at Kanpur confirm the positive bias in sea salt concentrations and lower carbonaceous aerosols during high pollution events. MATCHA captures the seasonal cycle of surface CO, but underestimates the observations, which can be attributed to the assimilation of CO profiles from MOPITT. A unique feature of MATCHA is its tagged-tracers of BC for sectoral and regional source attribution analysis. These tracers show anthropogenic BC peaking in winter, primarily from Chinese sources in the eastern and northern part of HMA, and Indian sources in western and central HMA. Biomass burning BC dominates during March–April along with substantial trans-boundary inflow throughout the year. BC emissions from Pakistan and Nepal also contribute significantly to the anthropogenic column burden of BC in parts of HMA. MATCHA is the first-of-its-kind high-resolution reanalysis to fully couple aerosols, radiation, and snow processes over HMA, offering a valuable dataset for investigating aerosol–cryosphere feedbacks and informing emission mitigation strategies in Asia. The dataset consists of hourly surface and column-integrated products and 3-hourly three-dimensional fields, and is publicly available at DOI: 10.5067/CG4OT8DJX2Z7.

1 Introduction

High-mountain Asia (HMA), often referred to as the Third Pole, is the largest freshwater source in the Earth's cryosphere after the polar regions (Immerzeel et al., 2010). It also acts as the "water tower" for major Asian river systems that supply freshwater to approximately 1.5 billion people (Immerzeel et al., 2010; Yao et al., 2022). These river systems are fed by runoff from direct precipitation and glacial melt in the region, marked by a complex terrain that acts as an elevated heat source for sustaining the Asian summer monsoon (Hahn and Manabe, 1975). Hydrological changes in HMA's glaciers and the dynamic processes affecting them are especially sensitive to climate change, with significant socio-economic consequences due to the presence of rapidly developing and highly populated economies in the vicinity of these glaciers (Pepin et al., 2022). Recent advancements in satellite observations, in situ measurements, and model simulations have characterized multiple changes in HMA, including high elevation warming, increased precipitation, heterogeneous retreating of glaciers, declining snow cover, and enhanced greening and vegetation growth (Hasson et al., 2016; Maina et al., 2022b, a, c; Notarnicola, 2022). These changes impact downstream regions by affecting water supply, regional ecology, land-use practices, and even the risk of natural hazards like landslides and glacial lake outburst floods (Kirschbaum et al., 2020).

The changes observed in HMA are driven by a combination of factors, including precipitation and temperature patterns, greenhouse gas (GHG) dynamics, and the emissions of short-lived climate forcing agents called light-absorbing particles (LAPs), including black carbon (BC), dust, and brown carbon (BrC). LAPs have been shown to cause significant reductions in snow albedo and accelerate the melting of snow and glaciers, with an efficacy comparable to greenhouse gases (Flanner et al., 2007; Qian et al., 2015; Ramachandran et al., 2023; Yasunari et al., 2013). Deposition of LAPs on the glaciers of HMA has been found to contribute to at least 10% of snow albedo reduction, with an associated radiative forcing of up to more than

https://doi.org/10.5194/essd-2025-275
Preprint. Discussion started: 11 September 2025

100 W/m^2 locally (Sarangi et al., 2019). This forcing can account for approximately 15% of total glacier melt in the southeastern Tibetan Plateau (TP) and up to a 6.3% increase in the summer melting rate of Pamir glaciers in the western TP (Schmale et al., 2017; Zhang et al., 2018). Snow albedo feedback is a significant contributor to warming in the cryosphere, which is further exacerbated by LAPs due to interactions of absorbing aerosols with the surface snow and near-surface meteorology (Flanner et al., 2007). Despite growing evidence linking deposition of LAPs to rapid warming and enhanced snowmelt in High Mountain Asia, our understanding of their precise physical impacts and the relative contribution of different types of LAPs remains limited. Observational challenges, model uncertainties, and complex interactions between aerosols and snow properties contribute to this gap. Significant gaps in our knowledge stem from uncertainties in aerosol emission estimates, incomplete understanding of aerosol radiative properties and chemical composition, and the lack of accurate representations of LAP-snowpack-meteorology interactions and BrC aerosols in coupled climate-chemistry models (Collins et al., 2017; Liu et al., 2020; Qian et al., 2015; Roychoudhury et al., 2022,2025; Skiles et al., 2018; Xu et al., 2021). Current global and regional chemical transport models tend to underestimate BC and dust concentrations, leading to underestimations of LAPs in source pollution regions, surface snow, and columnar aerosol absorption (Goto et al., 2011; He et al., 2014). These discrepancies in LAPs concentrations and LAPs-related aerosol properties result in significant errors in estimates of aerosol-climate interactions (such as radiative forcing and aerosol-induced snow albedo feedback) (Liu et al., 2020; Tuccella et al., 2021; Xu et al., 2021).

Despite the critical role of HMA as a major Asian freshwater source, monitoring the changes in this region is a challenging task because of the complex terrain of HMA along with the heterogeneous nature of glaciers, which makes it very difficult to set up and maintain in situ network measurements, and leads to relatively large uncertainties in satellite retrievals. Polar orbiting satellite retrievals such as those from the Moderate Resolution Imaging Spectroradiometer (MODIS) and Measurement of Pollution in the Troposphere (MOPITT) have a good spatial coverage but limited temporal (1-2 times per day) coverage. Dynamical models serve as another option to create datasets to study long-term changes in HMA. However, such model simulations have their own biases due to spatial resolution, simplified representation of different physics and chemistry processes via parameterizations, and inaccuracies in model input datasets (Tarek et al., 2021). This issue can be partially addressed by global and regional reanalyses that are developed by assimilating available in situ and satellite observations into models to keep the model state as close to observations as possible within the constraints of model and observation errors. The reanalyses represent spatio-temporally continuous and dynamically consistent datasets and can be used to assess trends and emerging patterns in climate, snow properties, precipitation, temperature, and aerosols. Several global meteorological and chemical reanalysis datasets are available and widely used, such as ERA5, CAMS-EAC4, and MERRA-2. However, coarser spatial resolution in global reanalyses leads to biases in precipitation, snow properties, temperature, and atmospheric chemistry estimates over HMA due to inaccurate representation of the complex topography and the inability of the models to resolve the spatial heterogeneity in such regions. Additionally, existing reanalyses that include atmospheric chemistry neither consider the coupling between aerosols and meteorology to its fullest extent nor provide concentrations for species like BrC, which also significantly contributes to LAP-induced radiative forcing. While there are ongoing efforts to

https://doi.org/10.5194/essd-2025-275
Preprint. Discussion started: 11 September 2025

enhance the representation of aerosol-meteorology-snow interactions in advanced models (Fast and others, 2006; Grell et al., 2005; He, 2022; He et al., 2014, 2018; Kumar et al., 2014) and to improve Asian emission inventories (Govardhan et al., 2016; Sadavarte and Venkataraman, 2014), recent model developments suggest that assimilation of satellite retrievals of quantities like carbon monoxide (CO) and aerosol optical depth (AOD) can significantly improve simulations of aerosol species (Arellano Jr. et al., 2010; Kumar et al., 2019; Liu et al., 2011; Saide et al., 2013; Werner et al., 2019). Regional reanalyses attempt to address this issue by employing higher spatial resolution and smaller time steps and better resolving mesoscale processes to constrain these uncertainties. Several regional reanalyses have been developed over HMA, such as the 18-year HMA Snow Reanalysis (HMASR) dataset, the High Asia Refined analysis (HAR) that focuses on snow properties and meteorological variables, as well as a multi-decadal land reanalysis focusing on hydrological budget over the region (Liu et al., 2021; Maina et al., 2024; Wang et al., 2021). However, none have been developed so far that provide a long-term (more than a decade) record across meteorological, snow and atmospheric composition quantities over HMA, especially a chemical reanalysis that constrains atmospheric composition using satellite observations and simultaneously accounts for interactions across atmospheric composition, meteorology, land surface, and snow.

Given the issues related to the representation of LAPs, their feedbacks and impact on climate over HMA, as well as the potential of data assimilation of relevant satellite retrievals to improve atmospheric composition estimates, we present a first attempt towards developing a new regional hydroclimate-chemical reanalysis called MATCHA (Model for Atmospheric Transport and Chemistry in Asia) and evaluate its ability to simulate meteorological, hydrology and air quality parameters (Kumar et al., 2024). This reanalysis encompasses around 17 years of simulations from 1st January 2003 to 31st August 2019. It is based on the Weather Research and Forecasting model coupled with Chemistry (WRF-Chem) that uses the Community Land Model (CLM) coupled with SNICAR (SNow, ICe, and Aerosol Radiative) model for land surface processes to capture aerosol-snowpack interaction. The MATCHA reanalysis provides hourly (for two-dimensional quantities) to three-hourly (for three-dimensional quantities) estimates at a spatial grid spacing of 12 km over the Asian domain (58° – 140°E, 4° – 40°N) and 35 sigma levels extending from the surface to 50 hPa. The novelty of this regional reanalysis includes:

1. coupling aerosol species with radiation and snowpack using CLM-SNICAR coupled with WRF-Chem.
2. assimilating nearly two decades of aerosol optical depth (AOD) data from MODIS and carbon monoxide profiles from MOPITT.
3. simulating the lifecycle of BrC aerosols, including its deposition on snowpack.
4. providing BC, dust, and BrC abundances and quantifying BC deposition (both wet and dry) in HMA from 10 Asian source regions using a tagged-tracer approach for source attribution of anthropogenic pollution and its impact on HMA's cryosphere.

This study 1) presents the model description behind MATCHA and the assimilation methods involved, and 2) evaluates a variety of key meteorological and chemical quantities from MATCHA with observations (ground-based and satellite

https://doi.org/10.5194/essd-2025-275
Preprint. Discussion started: 11 September 2025

measurements) to understand the model biases in simulating these quantities. The primary objective of this paper is to provide a comprehensive validation of the MATCHA dataset and document error characteristics for different parameters so that future users have a better understanding of when and where this dataset can be applied. The remainder of the paper is structured as follows: Section 2 describes the model configuration, the tagged-tracer approach, the daily assimilation setup, and the various observations that are used for evaluations. Section 3 describes the evaluation results across different meteorological and chemical variables, with a discussion on the model biases and the factors causing them. Section 4 discusses the key findings related to the evaluation and the applicability of the reanalysis for future studies.

2 Data and Methods

2.1 Model Description and Setup

We employed version 3.9.1 of the WRF-Chem model coupled with CLMv4-SNICAR to simulate the three-dimensional distributions of meteorological parameters, and chemical constituents including LAPs, their deposition and evolution in the atmosphere and snow-covered areas over HMA. Our fully-fledged WRF-Chem-CLM-SNICAR (WC-CS) modeling system simulates the interactions between atmospheric composition (trace gases and aerosols), radiation, clouds, snowpack, and land-surface processes. The model domain covers the majority of Asia on a Lambert conformal grid centered at 32°N, 97°E with a horizontal grid spacing of 12 km by 12 km and 35 vertical levels stretching from the surface up to 50 hPa. Meteorological fields from the ERA-Interim (Dee et al., 2011) were used to nudge the temperature, specific humidity, and wind variables in the model every six hours, except for the lowest ten model levels (below the planetary boundary layer), to constrain the large-scale dynamics (Bowden et al., 2012; Seaman et al., 1995). We adopt key model configurations for meteorological, chemical, and land surface processes from Kumar et al., 2013, 2015a, building on our past efforts to simulate key meteorological and chemical characteristics of south Asia (see Table 1 for the schemes and configurations). For simulation of aerosols, the Model for Simulating Aerosol Interactions and Chemistry (MOSAIC) 4-bin sectional aerosol scheme was utilized. MOSAIC simulates both the mass and size distribution of aerosols, thereby enabling interactions of aerosols with cloud microphysics and radiation as demonstrated by previous studies over Asia (Kumar et al., 2015b; Matsui, 2016; Sarangi et al., 2019).

Figure 2 highlights key environmental characteristics of the MATCHA domain, including elevation, population density, land cover, and local climate zones, to provide context for the region's physical and human landscape. The complex topography of the region can be seen through the sparsely populated high elevation regions of the Himalayas and the Tibetan Plateau (> 2.3 km) in Fig. 2a, in contrast with the low elevation regions of the Indo-Gangetic Plain and eastern China (< 1.5 km), which accommodates approximately 10% of the world's population, including some of the most polluted cities globally (Mogno et al., 2021). The clustered population hotspots over the Indo-Gangetic Plain and eastern China can be seen in Fig. 2b, depicting the total population count in the model grid (12 km) averaged between 2003-2019 from the GlobPOP dataset (Liu et al., 2024). The land cover classification averaged (mode) across the domain from the MODIS International Geosphere Biosphere

https://doi.org/10.5194/essd-2025-275
Preprint. Discussion started: 11 September 2025

Programme (IGBP) global vegetation classification scheme in Fig. 2c shows extensive croplands in areas of high population, and natural forests (mixed forests and grasslands) in areas of high elevation. Finally, in Fig. 2d, the local climate zones, mapped by Demuzere et al., 2022), highlight the level of urbanization and rapid development across Asia with pockets of major metropolitan centers (in red) interspersed with natural vegetation (green) and low-density developments (light orange). Previous studies have shown how air quality and weather patterns in Asia are impacted by the complex interactions between the complex topography, population density, urbanization effects, and land use patterns (Ganzeveld et al., 2010; Stewart et al., 2013; Tian et al., 2021).

2.1.1 Tagged-Tracer Approach for Source Contribution of BC

A tracer approach implemented in WRF-Chem tracks BC particles and their deposition fluxes (both wet and dry) from 10 different emission source regions and different sources across Asia, following the method described in our previous study (Kumar et al., 2015a). The BC tracers are independent variables added to the model simulation that experience the same atmospheric processes as standard BC particles (emissions, transport, aging, and deposition) but do not interfere with the model simulation/processes (e.g., radiation, clouds, atmospheric chemistry) or other aerosol particles. The BC tracers account for all sources of BC in the model by tracking BC emitted from anthropogenic (BC-ANT) and biomass burning (BC-BB) sources within the domain, as well as BC inflow from the lateral domain boundaries resulting from all emission sources located outside the model domain (BC-BDY) to provide insights into the background levels of BC in HMA. In addition to these emission sources, ten regional tracers are added to track BC emitted from different Asian countries or source regions, as shown in Fig. 1a. The anthropogenic emissions of BC from outside of these 10 regions are tracked separately as a tracer and defined as “Rest of Asia”. The tracers for the source regions do not strictly follow administrative (country) boundaries, some of which are defined based on elevation. For instance, grid cells above 1500 km are defined as part of the Tibetan Plateau and Nepal tracers rather than India. The ability to quantify the BC abundance from these regions individually can thus help attribute the deposition of BC over snowpacks in HMA. This tagged-tracer approach has been applied to previous air quality studies for both CO and BC to assess sectoral, regional, and local contributions in Asia and Africa (Ghude et al., 2020; Kumar et al., 2013, 2015b, 2022).

2.1.2 Implementation of BrC aerosols in WC-CS

We implement a BrC lifecycle scheme in WRF-Chem that explicitly track both primary and secondary BrC species through key processes: direct emissions, secondary formation, direct interaction with radiation (absorption and scattering), interaction with clouds (as cloud condensation nuclei), wet and dry deposition, and interaction with snowpack post-deposition (reducing snow albedo). Specifically, the wavelength-dependent BrC refractive indices follow Wang et al., 2014 who derived a best-fitting line based on previous observations. The refractive indices are used to compute BrC optical properties (extinction and absorption cross-sections, single-scattering albedo, asymmetry factor) based on the Mie theory in WRF-Chem. We only

https://doi.org/10.5194/essd-2025-275
Preprint. Discussion started: 11 September 2025

consider the biofuel and biomass burning sources for primary BrC emissions, since previous studies found that fossil fuel combustion contributes very slightly to BrC emissions and is poorly characterized (Saleh et al., 2015). Primary BrC emissions are estimated by assuming that all light absorption from total organic aerosol (OA) at the time of emissions is from BrC. This is done by matching the total OA absorption at emissions, which is computed using the BC/OA emission ratio based on the observationally derived parameterization from Lu et al., 2015. The secondary BrC formation is assumed to be aromatic secondary organic aerosol (SOA) production following previous studies (Nakayama et al., 2010; Zhang et al., 2020). The BrC photobleaching process follows the parameterization by Wang et al., 2018. The BrC transported from global boundary conditions is set to 20% of total OC boundary transport fluxes based on the mean global BrC/OC burden ratio (Jo et al., 2016). The BrC transport, interaction with radiation and clouds, and deposition processes follow the default WRF-Chem treatments of OC. The BrC evolution in snowpack after deposition and impact on snow albedo reduction are represented following Flanner et al, 2009.

2.1.3 Chemical data assimilation (DA) system and set-up

We use the three-dimensional variational scheme from the community GSI (Gridpoint Statistical Interpolation) system to assimilate satellite observations: MODIS AOD and MOPITT CO, into the WRF-Chem-CLM-SNICAR model background. This assimilation is performed every three hours to produce an optimal analysis state for chemical species. Previous studies have demonstrated the improvement in $PM_{2.5}$ forecasts from MODIS AOD and CO simulations from MOPITT CO profiles (Kumar et al., 2019, 2025). The assimilation system consists of four major components:

a. **Quality control of satellite observations to be assimilated:** Level 2 retrievals of MODIS AOD Collection 6.1 and MOPITT CO version 9 are selected within the simulation period using quality assurance flags within the products, following previous studies (Kumar et al., 2025).
b. **Background error covariance (BEC):** The BEC represents the error in the model background. The National Meteorological Center (NMC) method of the community Generalized Background Error (GEN_BE) system is used for calculating winter and summer representative BEC (Parrish and Derber, 1992). The NMC method uses two different WRF-Chem forecasts valid at the same time to calculate the statistical parameters in BEC. It calculates every 3 hours to be consistent with the assimilation cycles. Anthropogenic and biomass burning emission uncertainties are considered in the design of the BEC following a 100% uncertainty assumption in both types of emission sources (Kumar et al., 2020).
c. **Observation error covariance:** The observation error covariance for AOD is specified following Remer et al., 2005. For CO retrievals, the uncertainty reported in the MOPITT products is used for the observation error covariance (NASA/LARC/SD/ASDC, 2000).
d. **Observation operator:** The observation/forward operator converts model-simulated chemical species into corresponding satellite-observed quantities (AOD and CO here). The default forward operator in GSI is modified for

https://doi.org/10.5194/essd-2025-275
Preprint. Discussion started: 11 September 2025

transforming the MOSAIC simulated aerosols in WRF-Chem into AOD based on the parameterization from Malm and Hand, 2007 as follows,

$$WCCS_{AOD} = \sum_{i=1}^{N} \beta_{ext_i} dz_i \quad (1)$$

$$\beta_{ext_i} = 0.003 f(RH)\{[ammonium\ nitrate] + [ammonium\ sulfate]\} + 0.004[organic\ mass] + 0.010[Light - absorbing\ carbon] + 0.001[fine\ soil] + 0.0006[coarse\ soil] + 0.00137 f(RH)_{SS}\{[sea\ salt]\} \quad (2)$$

where $i = 1,2 \ldots N$ represents the vertical layers in the WC-CS model, β_{ext_i} represents the extinction coefficient for layer i, values in the square brackets represent the mass concentrations of different aerosol compounds, $f(RH)$ and $f(RH)_{ss}$ represents the relative humidity (RH) correction factor that accounts for hydroscopic growth of sulfate-nitrate-ammonium and sea salt aerosol components respectively. $f(RH)$ and $f(RH)_{ss}$ are determined from look-up tables. Extinction due to other aerosol components is assumed to be invariant with RH. The forward operator for assimilating MOPITT CO follows that of (Kumar et al., 2025). This forward operator convolves WRF-Chem CO profiles using *a priori* profiles and averaging kernels from MOPITT before calculating the innovation in CO.

The analysis and background states consist of different LAPs and CO concentrations simulated in MATCHA. Daily assimilation of MODIS AOD and MOPITT CO is performed as shown in Fig. 1c. For MODIS AOD, exploration of the satellite products from both Terra and Aqua showed that the swaths cover the HMA domain between 0130 UTC to 0930 UTC. The assimilation is performed in a series of three-hour windows until 0900 UTC daily.

2.1.4 CLM-SNICAR to simulate LAP-snow-radiation interactions

Deposition fluxes (wet and dry) of LAPs, calculated in WRF-Chem (see Table 1), are supplied online to the CLM version 4.0 land scheme (integrated within WRF-Chem) coupled with the SNICAR snow albedo model. This coupling allows for the simulation of LAPs' evolution and their impact within the snowpack. In particular, CLM-SNICAR is used to compute the albedo and radiative flux of the vertically resolved multi-layer (up to five layers) snowpack containing LAPs (Flanner et al., 2007; Flanner et al, 2009). BC and BrC particles in snow are represented as both externally and internally mixed with snow grains based on deposition mechanisms following Zhao et al., 2014 while dust in four size bins is only treated as externally mixed with snow. Snow aging processes and meltwater transport are accounted for within CLM-SNICAR, which scavenges LAPs from the top layers to the bottom layers and redistributes LAPs through the snow layers (Flanner et al., 2007). The LAP-induced snowpack water and energy changes are used to simulate corresponding changes in other land surface conditions (e.g., soil moisture and runoff) and associated feedback in the atmosphere. The CLM-SNICAR scheme coupled with WRF-Chem

https://doi.org/10.5194/essd-2025-275
Preprint. Discussion started: 11 September 2025


Open Access Earth System Science Data Discussions

has been successfully used to reproduce snow albedo and LAPs in snow over other global mid-latitude mountain regions (e.g., Huang et al., 2022).

2.2 Observations

This section provides a brief overview of the observational datasets used to evaluate the MATCHA reanalysis. We draw on available chemical, meteorological, and hydrological measurements from global networks and satellite missions, with a focus on key variables that influence aerosol–meteorology interactions and snow processes. We evaluate across 12 key variables that represent meteorology (temperature, relative humidity, and wind speeds; both surface and profiles), land-atmosphere interactions and hydrology (daily precipitation, boundary layer height, and snow cover), and atmospheric composition (aerosol optical depth, single scattering albedo, and surface measurements of particulate matter and CO).

2.2.1 OpenAQ

OpenAQ is an open-source platform that compiles air quality data from multiple sources (public sources, environmental agencies, and international organizations) across the globe (Hasenkopf et al., 2016; Velasco et al., 2024), and can be accessed at https://openaq.org/. We obtained fine particulate matter ($PM_{2.5}$), coarse particulate matter (PM_{10}), and CO measurements during 2003 – 2019 across Asia within the model domain to compare with surface layer simulations from MATCHA. The OpenAQ measurements across all species were all post-2015 at an hourly resolution. In total, we had obtained approximately 10 million hourly measurements across 1705 (for $PM_{2.5}$), 1588 (for PM_{10}), and 1705 (for CO) sites. Considering that OpenAQ data do not have any quality assurance standard within the retrieved and compiled data, we used a modified version of the probabilistic outlier detection system from Wu et al., 2018 to detect three types of outliers from the retrieved OpenAQ data for surface $PM_{2.5}$, PM_{10}, and CO. The outliers were based on large variances and periodic discrepancies for each site and each of these species, as well as spatio-temporal outliers based on neighboring sites. Detecting spatio-temporal outliers requires a characteristic localization length that defines the radius within which neighboring stations influence a given site. An exploratory analysis of how inter-site correlations decay with distance indicated two such suitable lengths: 342 km for the dense networks across China and 117 km for stations elsewhere in the domain**.** We show the percentage of outliers removed across four seasons for each of the three species ($PM_{2.5}$, PM_{10}, and CO) in Fig. A1. The number of outliers removed for $PM_{2.5}$ is highest, particularly in summer (2%, ~19600 observations removed), while CO-related outliers were removed the least (less than 0.05%). Overall, the outlier detection system shows the strongest impact in summer for $PM_{2.5}$, suggesting greater variability in the observations across the 1705 sites over the domain.

2.2.2 Aerosol Robotic Network (AERONET)

The AERONET network provides radiometer measurements of aerosol optical properties like AOD and single scattering albedo (SSA) at different wavelengths across ground-based sites globally (Sinyuk et al., 2020, Holben et al., 1998). The AOD measurements have an estimated uncertainty of 0.01 at visible wavelengths and 0.02 at near-ultraviolet wavelengths (Dubovik

https://doi.org/10.5194/essd-2025-275
Preprint. Discussion started: 11 September 2025

et al., 2000). The uncertainty due to instrumental calibration for SSA is within 0.03 for AOD at 440 nm greater than 0.4 (Dubovik et al., 2000; Giles et al., 2019). We used AERONET Level 2 (cloud-screened and quality-controlled) data version 3 for AOD at 550 nm across 185 sites that lie within the MATCHA domain, and the inversion products were used to calculate SSA at 550 nm across 153 sites. AERONET AOD is available at 340, 380, 440, 500, 675, 870, and 1020 nm, sparingly across all sites. For this study, AOD at 550 nm (denoted as AOD550 hereafter) is estimated by using AOD values at nearby available wavelengths (500-675 nm or 440-675 nm) to calculate the Angström exponent using the higher and lower wavelengths, which was further used to estimate AOD at 550 nm.

SSA at 550 nm (denoted as SSA550 hereafter), on the other hand, is calculated using the following relation,

$$SSA_{550nm} = \frac{AOD_{550nm} - AAOD_{550nm}}{AOD_{550nm}} \quad (3)$$

Where AOD and AAOD (absorption AOD) at 550 nm are calculated using the following relation,

$$\tau_{550nm} = \tau_{870nm} \frac{440^{\alpha}}{870} \quad (4)$$

Where τ is the AOD/AAOD at the corresponding wavelength, and α is the Angström exponent for 440 to 870 nm, available in the Level 2 inversion products. We note that AERONET datasets have been widely used in evaluating model aerosol simulations, reanalyses, and satellite retrievals globally (Bright and Gueymard, 2019; Gueymard and Yang, 2020; Xian et al., 2024).

2.2.3 Integrated Multi-satellitE Retrievals for GPM (IMERG)

The Integrated Multi-satellitE Retrievals for GPM (IMERG) from NASA is a global surface precipitation product at 0.1° resolution generated from the Global Precipitation Measurement (GPM) satellite constellation as a NASA-JAXA collaborative effort (Huffman et al., 2014). In this study, we use the daily accumulated precipitation from the final run version of IMERG (IMERG-F), which calibrates the satellite measurements through monthly rain gauge analysis from the Global Precipitation Climatology Centre (GPCP). We considered the calibrated daily accumulated precipitation files from 2003 through 2019 and performed a first-order conservative interpolation to the MATCHA grid for comparison. Among available satellite products, the IMERG Final Run is one of the most accurate for daily and monthly precipitation over HMA and much of Asia, although, like other precipitation datasets, it exhibits systematic biases over complex terrain and during extreme rainfall events (Dollan et al., 2024; Lee et al., 2019; Yu et al., 2021).

2.2.4 Integrated Global Radiosonde Archive (IGRA)

The Integrated Global Radiosonde Archive (IGRA) version 2 from NCEI-NOAA (National Oceanic and Atmospheric Administration's National Centers for Environmental Information) hosts radiosonde and weather balloon observations across the globe from multiple sources with observations of temperature and relative humidity profiles at 00 and 12 UTC (Durre et al., 2018). While the records are quality controlled for climatological and temperature-related outliers (in addition to other issues), there exist some known limitations, particularly unreliable humidity measurements above 250 hPa, inhomogeneities

https://doi.org/10.5194/essd-2025-275
Preprint. Discussion started: 11 September 2025

in temperature and wind profiles due to changes in instruments, station locations, and lack of uncertainty estimates in the records (Durre et al., 2006, 2018; Madonna et al., 2022). The observations are standardized at the following pressure levels: 1000, 925, 850, 700, 500, 400, 300, 250, 200, 150, 100, 70, 50, 30, 20, and 10 hPa. We obtained radiosonde profiles of temperature, relative humidity, and wind speed from 129 sites over the MATCHA domain and only considered layers up to 100 hPa for comparison, considering the model top pressure of 50 hPa in the MATCHA model configuration (see Table 1). In addition, we obtained planetary boundary layer (PBLH) height values derived within IGRA v2 (based on the parcel method in Seidel et al., 2010) at 00 and 12 UTC for 125 sites within the model domain and simulation period, for comparison with MATCHA. It is important to note here that the actual launch time of the radiosondes and their derived PBLH may differ from the standard reporting times (00/12 UTC); therefore, we used the nearest time from the hourly simulations of MATCHA to obtain the spatiotemporally collocated model PBLH. Most observations were launched before the standard reporting time, with a median lead time of approximately 30 minutes with a similar variability. Although the derived PBLH values are reported to have an uncertainty of a few 100 m (Seidel et al., 2010), sparse sampling in the vertical within IGRA can result in large uncertainties in deriving the PBLH, compared to high-resolution soundings (Guo et al., 2021; Liu and Liang, 2010). The parcel method used to derive PBLH within the IGRA archive has also been shown to yield systematically lower values than other methods, with greater diurnal and seasonal variability (Seidel et al., 2010). Uncertainties in humidity sensors onboard radiosondes also contribute to the overall uncertainty in PBLH estimates. Additionally, radiosonde observations from IGRA do not include near-surface wind measurements, which can introduce discrepancies in the derived PBLH reported within IGRA (Madonna et al., 2021). Considering the absence of quality assurance of these derived PBLH within the archive, we implemented a three-step outlier detection process for our evaluation of PBLH as follows,

1. We applied a rolling window-based method to both observed and modeled PBLH time series for each site, using a window of seven consecutive data points, corresponding to a week of measurements at 00 and 12 UTC, analyzed separately. We then calculated the local median and median absolute deviation (MAD) across this window (Leys et al., 2013) and scaled the MAD by multiplying it by 1.4826 to make it comparable to the standard deviation for Gaussian data. Modeled and observed PBLH values were then considered outliers if their absolute deviation from the median exceeded twice the scaled MAD. A multiplier of 2 was chosen to make the outlier detection conservative based on empirical testing of different values.
2. After identifying the outliers in both the modeled and observed PBLH, we computed their differences and applied the same window-based method with a multiplier of 2 to identify periods of exceptionally high disagreements between the observations and the model. These discrepancies may be attributed to measurement errors, model biases associated with strong daytime convective mixing, very stable nocturnal boundary layers, complex topography, and spatio-temporal representativeness issues.

https://doi.org/10.5194/essd-2025-275
Preprint. Discussion started: 11 September 2025

3. We finally applied a simple MAD-based filter (using a multiplier of 2 again) across the full set of observed and modeled PBLH values. This step helps identify additional extreme values that may not have been flagged by the rolling window approach.

This method allows robust filtering across datasets of spurious values that might arise from both observational and model uncertainties. The window-based approach allows us to filter out periodic outliers, while the last step allows us to filter out highly variable values from the entire dataset. Figure A2 shows histograms of both observed and modeled PBLH at 00 UTC and 12 UTC across seasons, comparing the distributions with and without outlier removal. In the top row (the original dataset with outliers), there are pronounced tails extending to higher PBLH values, particularly during summer at 12 UTC, suggesting that a small fraction of extreme values dominate the upper range of PBLH values. In the bottom row (after outlier removal), these tails are substantially reduced, resulting in more plausible distributions of PBLH.

2.2.5 NOAA's Integrated Surface Database (ISD)

The Integrated Surface Database (ISD) from NOAA consists of sub-daily measurements of several meteorological parameters pertaining to surface climate across tens of thousands of ground-based stations globally (Smith et al., 2011). We obtained six-hourly surface measurements of temperature at 2 m, dewpoint at 2 m, and wind speed at 10 m, after applying quality flags ((Lott, 2004)) from 1114 sites within the model domain and simulation period, and compared them with MATCHA. We calculated relative humidity at 2 m from temperature and dewpoint at 2 m from the ISD values based on (Hyland and Wexler, 1983) and compared them with relative humidity at 2 m in MATCHA (which was estimated based on Hyland and Wexler, 1983, using temperature, specific humidity at 2 m, and surface pressure from MATCHA). The ISD database has been used in a couple of previous studies as a validation dataset for meteorology over Asia (Faber et al., 2024; Kumar et al., 2015b; Xi, 2021; Yin et al., 2022).

2.2.6 BC Observations

Daily surface BC measurements between 2015 – 2019 were obtained from the Atmospheric Pollution and Cryospheric Changes (APCC) program, a monitoring network established over the Third Pole (Tibetan Plateau) and its surroundings to study the impact of atmospheric pollutants on the cryospheric changes over the region (Kang et al., 2022). We obtained daily BC measurements from 19 sites in the APCC program across the Tibetan Plateau and the Himalayas and compared them with surface BC abundance from MATCHA. In addition, monthly atmospheric surface BC measurements from 13 sites (across India and the Tibetan Plateau) for the year 2006 and seasonal averages of surface BC for 24 sites over India were obtained from He et al., 2014 and Kumar et al., 2015b.

2.2.7 MODIS Snow Cover Fraction

https://doi.org/10.5194/essd-2025-275
Preprint. Discussion started: 11 September 2025

Given that MATCHA incorporates CLM coupled with SNICAR, we specifically evaluated snow cover fraction (SCF in %) between MATCHA and satellite-based MODIS retrievals. We use daily SCF at a spatial resolution of 0.05° based on the Normalized Difference Snow Index (NDSI) (Hall and Riggs, 2007). Specifically, we use the MODIS (Terra and Aqua) Daily Level 3 (L3) Global 0.05 Deg Climate Modeling Grid Version 6 product with pixels having only recommended quality flags of 0 (best). We resampled the MODIS pixels to MATCHA's 12 km resolution by aggregating and averaging the finer pixels. For our comparison, our results are particularly focused on the aggregated first-order glacier regions over HMA based on the Randolph Glacier Inventory (RGI v6) (Pfeffer et al., 2014). These products have been used in previous studies over the domain, where they have reported promising results and high accuracy over HMA (Immerzeel et al., 2009; Li et al., 2018; Pu et al., 2007).

2.3 Evaluation metrics

We use four metrics: mean bias difference (MB), root mean square error (RMSE), Spearman's rank correlation (R), and a non-parametric version of the Kling-Gupta efficiency (KGE) statistic (Gupta et al., 2009; Pool et al., 2018) to evaluate the performance of MATCHA. MB represents the deviation of model (MATCHA) values from observations in both magnitude and sign, ranging from $-\infty$ to ∞ with a perfect value of 0. RMSE represents the positive deviation of the model from observations, ranging from $-\infty$ to ∞, with a perfect value of 0. R represents the degree of monotonic relationship between models and observations and is less sensitive to outliers compared to the commonly used Pearson's correlation (Wilcox, 2016). KGE quantifies the overall accuracy of the model by integrating correlation, bias, and variability. It ranges from $-\infty$ to 1, with a perfect value of 1. Here we use the non-parametric version of the original KGE metric based on Pool et al., 2018. The definitions of these metrics are as follows,

$$MB = \overline{M_i - O_i} \tag{5}$$

$$RMSE = \sqrt{\overline{(M_i - O_i)^2}} \tag{6}$$

$$R = \frac{cov(r_M, r_O)}{\sigma_{R_M}\,\sigma_{R_O}} \tag{7}$$

$$KGE_{NP} = 1 - \sqrt{(R-1)^2 + (\beta-1)^2 + (\alpha-1)^2} \tag{8}$$

where M and O are the model and observed values respectively, i refers to the seasonal average per site or grid cell, r refers to the ranks of the model and observed values, σ refers to the standard deviation of the ranks, cov refers to the covariance of the ranks r, β refers to the bias ratio, and α refers to the variability ratio. β and α are defined as follows,

$$\beta = \frac{\overline{M_i}}{\overline{O_i}} \tag{9}$$

https://doi.org/10.5194/essd-2025-275
Preprint. Discussion started: 11 September 2025

$$\alpha = 1 - \frac{1}{2}\sum_{i=1}^{N}\left|\widetilde{F_{M_i}} - \widetilde{F_{O_i}}\right| \quad (10)$$

where F_{M_i} and F_{O_i} are the model and observed values, sorted and normalized by the product of their mean and sample size, as follows,

$$\widetilde{F_i} = \frac{x_i}{\mu_x N} \quad (11)$$

where x_i is the i-th largest value of the model or observed set of values, μ_x is the model or observed mean, and N is the number of values. The non-parametric variant of KGE uses the Spearman correlation (R), instead of Pearson's correlation in the original KGE, and the variability ratio in the non-parametric version is based on quantifying the similarity in the empirical distributions of the model and observed values, rather than the ratio of their means as in the original KGE.

These metrics are based on the seasonal averages at each site or grid cell (depending on the scale of the observations used) per variable and report them per season in Tables 2,3, and 4. For our evaluation, we show a spatial representation of the seasonal average of the variable from the observation source and MATCHA, as well as their model bias. We define four seasons as follows: DJF (winter), MAM (spring), JJA (summer), and SON (fall).

3 Results

In this section, we present and discuss the evaluation of MATCHA across three key groupings: (1) meteorology, (2) land-atmosphere interactions and hydrology, and (3) atmospheric composition, each reflecting core aspects of the model's coupled atmospheric and surface processes.

3.1 Meteorology

3.1.1 Temperature

Surface temperature at 2 m from both observations and MATCHA (Fig. 3) highlights the zonal temperature gradient and captures spatial variations very well when compared to surface observations from ISD. The largest biases, exceeding 2°C, occur during winter over Asia, particularly in mainland India and southern China. Over the Indo-Gangetic Plain, MATCHA exhibits a consistent cold bias greater than 1 °C across all seasons, whereas eastern China shows a pronounced cold bias (< 2 ºC) in winter, absent in other seasons. Overall, MATCHA tends to exhibit a cold bias across most seasons, except during summer, where a slight warm bias is noted (MB of 0.07°C; Table 2). The highest cold bias appears in winter (MB of -0.38°C). RMSEs are lowest in fall (2.5°C) and peak in winter (4°C). Correlations across the domain are very strong (R > 0.95) in all seasons except for the slight decline in summer (R of 0.85), indicating that MATCHA captures surface temperature patterns effectively across all seasons. KGE values are highest during transition seasons (0.95), suggesting best model performance, which worsens a bit during winter (KGE of 0.82). The IMDAA (Indian Monsoon Data Assimilation and Analysis) regional

https://doi.org/10.5194/essd-2025-275
Preprint. Discussion started: 11 September 2025

reanalysis over the Indian subcontinent, with a resolution comparable to MATCHA, has also shown near-surface temperature biases exceeding 2°C over India, which is consistent with the findings of our study (Ashrit et al., 2020).

The temperature profiles from MATCHA align well with radiosonde profiles from IGRA in Fig. 4. These profiles reveal a consistent cold bias across the domain for most seasons, especially in the lower troposphere (below 500 hPa), and a slight warm bias in the upper troposphere. The median biases across all sites are at a maximum of 2ºC. The cold bias is most pronounced in spring (-1.3°C) and smallest in fall (-0.95°C), with RMSEs following a similar seasonal pattern (4 ºC versus 3.5ºC). Correlations across all sites and atmospheric layers remain high (R > 0.98) across all seasons, suggesting that MATCHA reasonably captures the vertical temperature structure despite the systematic cold bias. This is also confirmed by the KGE values across the seasons (0.98) suggests almost "perfect" model agreement. KGE values for the lower troposphere (below 500 hPa) are slightly higher (KGE ≥ 0.99) than the upper troposphere (KGE between 0.95 - 0.97), suggesting better agreement in the lower troposphere.

3.1.2 Relative Humidity

Relative humidity (RH) at 2 m across ISD sites and corresponding MATCHA values are shown in Fig. 5. MATCHA captures the seasonal cycle very well, with a strong moisture gradient peaking from spring and declining till winter. Correlation across the domain is strong (R > 0.64; Table 2), which varies from 0.85 in summer to 0.64 in winter. However, MATCHA generally simulates drier conditions (negative model bias) compared to observed surface RH measurements. This dry bias is pronounced in northern India, northern and eastern China, and the Himalayas, particularly during winter and spring. Conversely, a prominent moist bias is observed in central and northern India, Pakistan, Afghanistan, and southeast Asia during spring and summer. The wet bias in Pakistan and Afghanistan peaks in summer (> 15%) and becomes rather wetter during winter (> -5%). Central Asia exhibits a consistent dry bias (> -10%), which is strongest in summer (> -15%). Across seasons, spring and summer exhibit a general wet bias, while winter and fall are characterized by a general dry bias. The domain-averaged biases are highest in winter (MB = -2%) and lowest in summer (MB of 1%). RMSE values, however, are much larger (> 9%) and peak in winter (RMSE of 10%), while being lowest in fall (9%). KGE values lie between 0.6 to 0.85, suggesting good model performance, with the best agreement across the domain in summer that slightly worsens in winter. Overall, MATCHA most accurately reproduces surface RH values across all seasons.

RH profiles from MATCHA in Fig. 4 display a slight moist bias (< 5%) in the lower troposphere, transitioning to a drier bias (> 5%) in the upper troposphere, which increases near the tropopause. The median biases across all sites vary at most by 20%, with the strongest differences occurring mostly for winter across all layers. Across the domain, the moist bias below 500 hPa is relatively low (<1%) compared to the upper troposphere values, where the biases are drier and higher (-1 to -5%). This upper tropospheric dry bias is most prominent in spring and summer. Across the domain, the dry bias (see Table 2) across all tropospheric levels is strongest in winter (MB = -3.3%) and weakest in summer (-0.6%). RMSE values are comparatively high (> 10%) across all seasons, peaking at 15% in winter. Correlations are strong (> 0.66) and highest in summer (0.87). KGE values for all layers lie between 0.64-0.85, suggesting good model performance, with the best (worst) agreement in summer

https://doi.org/10.5194/essd-2025-275
Preprint. Discussion started: 11 September 2025

(winter). The model agreement, however, worsens for the upper troposphere with a strong contrast between the KGE values for winter versus summer (0.2 versus 0.7).

3.1.3 Wind Speed

The simulated and observed wind speed at 10 m across ISD sites are compared in Fig. 6. MATCHA shows moderate skills in capturing the seasonal cycle of 10 m wind speed with moderately strong correlation coefficient values of 0.4 to 0.7 across the domain. MATCHA also simulates higher wind speeds compared to observations, with a strong positive bias exceeding 2 m/s at most sites. This positive bias is particularly prominent over south Asia, especially during summer, and decreases gradually through fall and winter. In northern India, underestimations (negative bias) are observed at some sites in winter. Domain-wide seasonal biases (Table 2) indicate a consistent positive bias around 1.5-1.7 m/s across all seasons. RMSE values are also relatively stable across seasons, ranging between 2.3 m/s in winter to 1.8 m/s in fall. KGE values lie between 0.1-0.2, suggesting slightly poor model performance, with the best agreement during transition seasons (spring and fall). WRF is known to overestimate 10 m wind speeds for low to moderate wind speeds using all available PBL schemes, particularly over complex terrain (Cheng and Steenburgh, 2005; Jiménez and Dudhia, 2012; Mass and Ovens, 2010). Previous studies show a similar windier bias within the MERRA-2 reanalysis (Faber et al., 2024), particularly at inland stations (Carvalho, 2019).

Vertical wind profiles, as shown in Fig. 4, however, demonstrate good agreement between MATCHA and IGRA observations across all seasons, as seen from the relatively high KGE values (> 0.95), suggesting good model performance across profiles. The average bias across all pressure levels is small and rather negative in contrast to the positive bias at the surface, ranging from -0.1 m/s in winter to -0.5 m/s in fall. RMSE values are higher at upper levels, peaking at 3 m/s in winter, suggesting increased variability in performance with altitude. The biases and RMSEs are lower for the lower troposphere and higher in the upper troposphere. KGE values, however, are slightly better in the upper troposphere (> 0.95) than in the lower troposphere (~ 0.9). However, correlations averaged across all pressure levels are high (> 0.9), indicating that MATCHA effectively captures overall wind profiles but exhibits biases near the surface and in the upper troposphere.

3.2 Land-Atmosphere Interactions and Hydrology

3.2.1 Precipitation

Figure 7 presents a comparison of seasonal average daily accumulated precipitation between IMERG-F and MATCHA across four seasons. The precipitation patterns show a strong south-to-north gradient, particularly in summer and fall, depicting the onset and retreat of the Asian summer monsoon, with maximum intensity observed in the Indian subcontinent, especially in summer. During winter and spring, precipitation remains low except for coastal regions and parts of south Asia across both datasets. The high precipitation intensities are clustered in high elevation regions, with MATCHA successfully reproducing the broad spatial patterns of precipitation across Asia. MATCHA also captures localized features along the mountain foothills and coastal regions, particularly during summer/monsoon, which are smoothed over in IMERG-F. Similar localized patterns

https://doi.org/10.5194/essd-2025-275
Preprint. Discussion started: 11 September 2025

are seen in the IMDAA regional reanalysis when compared with a spatially smoother precipitation from the ERA-Interim reanalysis (Ashrit et al., 2020; Dee, 2011). Average daily accumulated precipitation from IMERG-F is highest (lowest) during the summer/monsoon (winter), with a gradual increase in spring followed by a decline in fall. MATCHA captures this seasonality over the domain but shows strong biases in southeast Asia (wet bias), northern India (dry bias), southern and eastern China (dry bias), and the Tibetan Plateau (dry bias). Over India and eastern China, a dry bias progressively develops in MATCHA from winter to summer (> -1 mm), except for a prominent wet bias in summer over the Himalayan foothills and the Bay of Bengal (> 1 mm). In summer, MATCHA captures high bands of precipitation above 6 mm, especially over the Bay of Bengal, albeit drier by 4 mm or more. A consistent wet bias (~ 1 mm or more) is seen in the Tibetan Plateau across all seasons, suggesting limitations in MATCHA to represent convective processes and orographic precipitation over complex terrain and high elevation areas, which has been noted by previous studies across models (Cannon et al., 2017; Sugimoto et al., 2024). Overestimation of precipitation in higher elevation regions of HMA has also been reported across reanalyses in studies, particularly in summer (Dollan et al., 2024). A prominent dry bias is seen in southern China and southeast Asia during fall (> 0.1 mm), suggesting issues in the model to simulate the seasonal retreat of the summer monsoon. Across the domain, the evaluation statistics as discussed in Sect. 2.3 is shown in Table 3. An overall dry bias is observed, which is highest (lowest) in summer (winter) (MB of -1.8 mm versus -0.4 mm). The RMSE also shows similar seasonality (5.2 mm in summer versus 1.2 mm in winter). The correlation across the domain is relatively strong, and highest (lowest) in summer (winter) (R of 0.8 versus 0.7). The higher correlation during summer can be seen due to general patterns of high precipitation bands associated with the Indian summer monsoon, as captured by MATCHA. The KGE values vary between 0.5 to 0.6 across seasons, with the relatively best (worst) agreement in summer (winter), indicating moderate skill in reproducing the daily precipitation climatology. Xie et al., 2022 compared daily precipitation products from both IMERG and ERA5-Land (with a similar resolution of 9 km) with rain gauges over China and reported median KGE values in the 0.5 range, as in our study. Contrary to the overall biases, MATCHA captures several important features of the daily precipitation cycle well, especially in the Indian subcontinent, while the performance worsens during winter in high-elevation and desert regions of Central Asia. The regional biases observed, particularly the wet bias in high elevation regions, highlight areas where further improvements in convection, land-atmosphere coupling schemes, as well as representation of orographic precipitation across complex topography, might be necessary (Barros and Arulraj, 2020; Mishra et al., 2021).

3.2.2 Planetary Boundary Layer Height (PBLH)

PBLH measurements at 00 and 12 UTC from IGRA (following the outlier approach in Sect. 2.2.4) across four seasons are shown in Figs. 8 and 9, respectively. Within the study domain, PBLH at 0 UTC represents the dawn to morning PBLH (from 0330 hours to 0930 local hours) on average, whereas at 12 UTC, the values correspond to the late afternoon to night PBLH (from 1530 hours to 2130 local hours). Figures 8 and 9 show the approximate shadow zones (nighttime conditions) at 00 UTC and 12 UTC, based on sun elevation angle and local time. MATCHA captures the diurnal variability (increase from morning to afternoon, 00 to 12 UTC) except in winter when the 12 UTC PBLH is significantly underestimated, particularly over the

https://doi.org/10.5194/essd-2025-275
Preprint. Discussion started: 11 September 2025

Indian subcontinent. MATCHA generally reproduces the seasonal variability from a shallower PBLH in winter to a deeper PBLH in summer, especially for 12 UTC. For 00 UTC, the biases are generally negative, with the largest underestimation observed over India and near the coasts during all seasons. An exception occurs in summer, where overestimations exceeding 200 m are prominent in eastern China (Fig. 8). For 12 UTC, positive biases dominate across the domain, particularly during summer (> 350 m) (Fig. 9). In winter, however, MATCHA sometimes underestimates PBLH by more than -200 m, particularly over the Indian subcontinent. Biases over eastern China remain predominantly positive across seasons, except for winter at certain sites. Over India, positive biases (above 330 m) are observed exclusively during summer.

Domain-wise metrics (in Table 3) at 00 UTC reveal low biases across all seasons, which is highest in summer (MB of 123 m), except for a mild underestimation in spring (MB of -2 m) and winter (MB of -14 m). RMSE values range from 167 m in winter to 210 m in summer, indicating larger variability in summer. At 00 UTC, the mean bias remains within ±125 m in every season, indicating that the WC-CS model reproduces the nocturnal PBLH (pre-sunrise) reasonably well. At 12 UTC, however, the bias becomes positive and highly variable across seasons, rising from 6 m in winter to 1400 m in summer, indicating a significant seasonal variability. There are strong overestimations at 12 UTC across the domain, particularly in summer, where the errors are the highest (MB of 1399 m, RMSE of 1552 m), compared to winter, where the errors are quite low (MB = 6 m, RMSE = 259 m). Models often struggle to capture the diurnal cycle of PBLH at the transition periods during morning and evening, where the decay during the transition from the daytime convective boundary layer to the nocturnal stable layer (and vice versa, a rapid rise from stable BL to convective BL during morning) can lead to the significant PBLH biases around 12 or 00 UTC (Cuchiara and Rappenglück, 2017; Hong, 2010; Taylor et al., 2014). For our domain, most of the observations belong to this this transition period because radiosonde launches at and 00 (12) UTC in most of Asia span from (pre-dawn) afternoon to daytime (night) locally, and we can see that (western) eastern part of the domain remains at nighttime conditions in Figs. 8 and 9 across all seasons. In Fig. A3a, we see that around 42% of the launches are at the transition zone or night times conditions (referring to D(LS), CT(M), CT(E), and N; low sun during day, civil twilight in the morning, evening, and night respectively) across all seasons, which coincide with the largest PBLH biases (MATCHA-IGRA, -500 m to 2200 m) (in Fig. A3b) as well. We group the local launch times of all the observations across five solar regimes, based on the solar elevation angle: N/night (≤ –6°), CT/civil twilight (–6° to 0°, morning (M) or evening (E) by the solar azimuth), D (LS)/Day Low Sun (0° to 10°), and D(HS)/Day High Sun (> 10°). The model biases are particularly high for D(LS) and CT(E) solar regimes, particularly during summer, as seen in Table 3. At 00 UTC, when most of the transition-related launches (~19% of total launches) fall into the N, CT(M) and D(LS) regimes, MATCHA shows negligible biases (< 125 m) and lower RMSEs (150 - 210 m, Table 3), confirmed by the smaller spread in Fig. A3b. In contrast, at 12 UTC, the large variability in errors is seen by the large spread in Fig. A3b, where the transition-related launches (~23%) fall into the N, CT(E), and D(LS) regimes. In Fig. A3e-f, we also show the distribution of shortwave downward radiation (SWDOWN) and surface sensible heat flux (sSHF) that drives convective BL, collocated at the observations. We can see that MATCHA still simulates high radiances and fluxes during the transition regimes, particularly at D(LS), CT(E), and even N, well after the Sun has crossed the horizon locally (within 10° elevation). This can indicate that the sustained surface forcing leads to increased buoyant turbulence, where the model PBLH continues

https://doi.org/10.5194/essd-2025-275
Preprint. Discussion started: 11 September 2025

to deepen, when it should already collapse. For instance, in summer, we see that the median sSHF in CT(E) exceeds 150 W $m^{-2,}$ leading to PBLH biases up to 2000 m, while during CT(M) or N, the sSHF falls near/below 0, and the biases collapse to 200 m. It is also important to note that we are limited to the nearest hour by the collocated MATCHA samples, when the actual launch times might be varied by minutes. A preliminary analysis showed that during the transition hours, around half (52%) of the radiosonde launches are released within ±15 minutes of the model hour, with an inter-quartile range of model biases up to 1000 m across all seasons. These findings indicate that residual heating due to a mismatch in radiation timing within the model can drive these high biases within MATCHA, particularly in summer, due to the delayed collapse in PBL in the transition hours.

For 00 UTC, correlation coefficients are weak across seasons, except weak correlations (0.1-0.2) for transition seasons (spring and fall), reflecting weak agreement between MATCHA simulations and IGRA observations. Correlations, however, are improved for 12 UTC compared to morning/daytime PBLH values, with the strongest correlation during summer (R of 0.5). KGE values, however, are slightly better for 00 UTC than 12 UTC, with the best model performance during spring for both 00 UTC and 12 UTC. The overestimation during summer suggests enhanced vertical mixing within the model due to strong convection over tropical regions. Previous studies have reported an overestimation of nighttime PBLH by factors of three or at least by 1400 m over regions in Asia from models (Lee et al., 2023; Zhang et al., 2022). A global analysis of daytime PBLH across reanalysis products like ERA5 and MERRA-2 has shown an average bias of up to 640 m in Asia across seasons (Guo et al., 2021). These discrepancies, characterized by high RMSE and low correlation, likely stem from the coarser model resolution in the horizontal and vertical (only 35 vertical levels), parameterizations to represent sub-grid scale convection and surface energy budgets, particularly in complex and high terrain, as well as the diurnal phase mismatch during the evening transition, as discussed in previous studies (Meng et al., 2023; Mues et al., 2018; Shin and Dudhia, 2016). Note that this validation is not representative of the model's ability to capture the diurnal cycle of PBL but is only an attempt to validate the model performance at two specific times (00 and 12 UTC) that belong to morning and evening transition zones in daily PBLH evolution. Nevertheless, our results point to the need for a deeper investigation into the uncertainties in both the observed and modeled PBLH, particularly in such regions with complex terrain (Chen et al., 2023).

3.2.3 Snow Cover

In Fig. 10, we compare SCF (in %) from MODIS and MATCHA across four seasons over the six aggregated major first-order glacier regions of HMA, based on the Randolph Glacier Inventory (RGI v6). The six HMA glacier regions from RGI v6 are defined as follows: INT: Inner Tibet, S and E Tibet; HTQ: Hengduan Shan, Qilian Shan; TNS: W and E Tien Shan; KNL: W and E Kun Lun; HIM: W, C, and E Himalayas; HKPH: Hindu Kush, Karakoram, Pamir, and Hissar Alay. MATCHA in general captures the seasonal as well as spatial distribution of SCF observed by MODIS, with the highest SCF in winter and lowest in summer. Both MODIS and MATCHA show widespread high SCF in winter across most regions, except for strong overestimation in INT (>20%) and strong underestimation in TNS (>20%). As snow begins to ablate in spring, MATCHA seems to suggest more snowmelt as seen by the negative biases across most regions, while some regions of INT, eastern HIM,

https://doi.org/10.5194/essd-2025-275
Preprint. Discussion started: 11 September 2025

and southern HTQ show positive biases, indicating lagged snow ablation. In summer, the highest SCF is mostly limited to the higher altitudes but is much lower in MATCHA compared to MODIS, particularly in HIM and HKPH. In fall, as snow begins to accumulate, MATCHA continues to exhibit negative biases over regions the high elevation areas, with some sporadic overestimations in the HMA. Across the glacier regions, MATCHA shows reasonable agreement with MODIS in winter over HIM, except for the larger biases during snowmelt. Highly rugged regions like HKPH also show strong biases likely due to limitations in resolving sub-grid processes. Areas like INT and HTQ show the strongest biases during winter. Domain-wide metrics in Table 3 show the lowest bias in winter (MB of -2%) and the strongest bias in spring (MB of -9%). RMSE values peak in winter (24%) and are lowest in summer (12%), suggesting greater variability in SCF in winter despite the lowest bias. Correlation values are strong across all seasons ($R > 0.7$), while the KGE values show the best agreement during winter, followed by spring (0.7 and 0.6). These results suggest that overall, MATCHA seems to struggle mostly during snowmelt periods. A recent study comparing SCF across various reanalyses over the TP found that the TP-averaged SCF was around 12-14%, which closely aligns with our estimates of SCF from MATCHA (11%) and MODIS (12%) over the INT region (Yan et al., 2024).

3.3 Atmospheric Composition

3.3.1 AOD and SSA at 550 nm

We compare mean daily AOD550 across seasons as shown in Fig. 11. MATCHA successfully captures the overall spatial patterns observed in AERONET, including consistently low values over the Tibetan Plateau, the Gobi and Taklamakan deserts, and the higher values over the Indian subcontinent due to pollution hotspots across all seasons except in summer, where the positive biases are the highest (>0.5). The seasonal cycle is also captured well, with AOD550 peaking in summer and declining during fall. A consistent negative bias is seen in spring, transitioning to a strong positive bias in southeast Asia and over India during summer. Over eastern China, the bias remains consistently positive during winter and spring and turns generally negative post-spring. Domain-wide metrics (Table 4) show a positive bias in all seasons, peaking in winter (MB of 0.2) with a slight negative bias in spring (MB of -0.01). RMSE values lie between 0.2 to 0.6, with the highest RMSE in winter (0.4) and the lowest in spring (0.2). Correlation values are highest in fall (0.7) and lowest in winter (0.5), with KGE values ranging between 0.7 in spring to 0.2 in summer, suggesting good agreement between MATCHA and AERONET values, with the best (worst) performance in spring (summer). Previous studies comparing AOD550 from AERONET across several reanalysis products have highlighted similar model biases over Asia, with bias values around 0.1-0.2 and RMSE values exceeding 0.2 for polluted regions, which is in line with our evaluation (Ansari and Ramachandran, 2024; Gueymard and Yang, 2020; Singh et al., 2017; Xian et al., 2024). These studies have also highlighted the fact that AOD550 biases over Asia are higher than in other regions of the globe. In addition to the magnitude of AOD550, we also look at the monthly climatology of AOD550 as seen from AERONET and MATCHA in Fig. A7, where we visualize the median AOD550 across months for sites grouped by tagged-tracer regions within MATCHA. While the seasonality and variability of AOD550 are captured at most sites, especially

https://doi.org/10.5194/essd-2025-275
Preprint. Discussion started: 11 September 2025

in China, we see major discrepancies during summer months among sites over India and Nepal. Similar discrepancies can be seen for Bangladesh, Myanmar, and southeast Asia, where AOD550 is much higher in the model during summer months than in AERONET. This highlights issues in the model regarding wet scavenging and deposition during the Asian monsoon, as well as convective transport to high-elevation regions, which leads to the higher bias during the summer months.

We also evaluated SSA550 from AERONET and MATCHA simulations in Fig. 12 to assess the relative contributions of scattering and absorbing aerosols represented in MATCHA. MATCHA generally underestimates SSA550 across the domain, indicating a higher fraction of absorbing aerosols in its simulations compared to AERONET. The negative bias in SSA550 peaks at -0.1 during winter and fall over most regions except in summer, where there is a slight positive bias (0.05) over India and southeast Asia, suggesting the influence of marine aerosols (which are scattering in nature) during monsoon. The seasonality of SSA550 is also captured in MATCHA as seen by the peak SSA550 (> 0.93) in summer, with the lowest SSA550 in winter (~0.65). Across the domain, a negative bias is seen across all seasons that peaks in winter (-0.2) and is lowest in summer (-0.02). RMSE values are highest in winter (0.2) and lowest in summer (0.04) as well. Correlation and KGE values are relatively low, except in fall (R of 0.6), suggesting relatively strong agreement. These results suggest that while MATCHA captures the optical properties post-spring, it struggles particularly in winter. The negative SSA550 bias across the domain indicates more absorbing aerosols simulated within MATCHA, which, when deposited on snow via the SNICAR module, darken the surface and accelerate snowmelt and can help drive the negative SCF bias observed in MATCHA (Sect. 3.2.3). Uncertainties in SSA550 in previous studies across WRF-Chem and other reanalyses have been reported to be up to 0.05, in contrast with the biases in MATCHA particularly for winter, suggesting differences in atmospheric composition within the model that can arise due to the emission inputs, and model parameterization related to deposition schemes and transport within the model. Biases in SSA within models often arise from physical and chemical assumptions related to aerosol composition and treatment. Internal mixing of absorbing particles (e.g., BC with sulphate), as employed in the MOSAIC aerosol model, tends to increase absorption, thereby lowering SSA values. Studies have shown that accounting for proper mixing state, refractive index, and including brown carbon and realistic dust size distributions can reduce SSA biases. Others report SSA underestimations in polluted regions, largely driven by emissions that overestimate absorbing aerosols inventories or the underestimation of scattering from hygroscopic growth (Dong et al., 2023).

3.3.2 $PM_{2.5}$, PM_{10}, and CO surface measurements

Fig. 13 shows the spatial distribution of seasonally averaged surface $PM_{2.5}$ mass concentrations from OpenAQ observations and MATCHA simulations. While MATCHA reproduces $PM_{2.5}$ seasonality in general, it diverges in magnitude, with pronounced regional differences. OpenAQ observations show the highest $PM_{2.5}$ concentrations during winter, followed by a decline in spring, the lowest levels in summer, and a rebound in fall. MATCHA captures this seasonal pattern over China but deviates over India, where it successfully simulates the fall-to-winter increase and the winter-to-spring decrease but inaccurately predicts an increase from spring to summer, contrary to the observed decline. MATCHA exhibits large positive biases over eastern China across all seasons, peaking in winter (> 107 $\mu g/m^3$) and reaching a minimum in summer (< 62 $\mu g/m^3$).

https://doi.org/10.5194/essd-2025-275
Preprint. Discussion started: 11 September 2025

In contrast, MATCHA consistently underestimates $PM_{2.5}$ over western China and HMA, with negative biases exceeding -6 µg/m³ throughout the year. Over India, biases are predominantly positive across all seasons, with the smallest deviations observed in spring and fall (within ±62 µg/m³). These discrepancies point to uncertainties in emission inventories and limitations in MATCHA's representation of transport, transformation, and deposition processes, particularly in regions with complex topography.

Across the domain (see Table 2), the positive bias is highest in winter (188 µg/m³), while the lowest is in spring (59 µg/m³). RMSE values are also substantial, peaking in winter (212 µg/m³) and reaching their lowest in both spring and summer (77 µg/m³). The significant RMSE values indicate marked seasonal variability, particularly in spring and summer, where similar mean biases (~61 µg/m³) and RMSE values are seen (117 µg/m³ in spring versus 94 µg/m³ in summer). Correlation values across the domain are moderate (~ 0.5-0.6) except in spring (R of 0.2), while KGE values are negative, reflecting limited agreement between MATCHA simulations and observed $PM_{2.5}$ concentrations. The KGE values suggest that MATCHA performs best during the transition seasons and struggles particularly in winter.

To investigate the processes contributing to the positive biases in surface $PM_{2.5}$, we analyzed the contributions of various aerosol species to surface $PM_{2.5}$ mass concentrations across the model domain. Additional analyses using the MATCHA dataset revealed several potential causes for these strong positive biases:

1. Elevated near-surface wind speeds along the Arabian Sea coastline, as seen in MATCHA, contribute to higher concentrations of surface sea salt and dust aerosols. An overestimation of near-surface winds (see Sect. 3.1.3) impacts the sea salt parameterization scheme within MATCHA based on (Gong et al., 1997), which strongly depends on 10 m wind speeds. Previous studies have demonstrated the overestimation of sea salt emissions from models based on this parameterization (Chen et al., 2016; Neumann et al., 2016; Saide et al., 2012). Sect. 3.1.3 highlights a positive bias in 10 m wind speeds, which aligns with this observation. Figure A4 shows the spatial distribution of sodium and chloride (components of sea salt) across four seasons, alongside average 10 m wind speeds. During summer, average wind speeds exceed 4 m/s along the western coast of the Indian subcontinent, transporting marine aerosols from the Arabian Sea to inland regions. Consequently, high concentrations of sodium and chloride are simulated in summer, exceeding 51 µg/m³.
2. Lower precipitation intensity simulated during summer, compared to observations, reduces wet scavenging efficiency, allowing aerosols to persist in the atmosphere. Precipitation totals in Fig. 7 highlight a dry bias over the Arabian Sea during summer, further supporting this underestimation. This reduced precipitation intensity limits aerosol removal, contributing to higher concentrations of marine aerosols in the region.
3. Errors in the partitioning of gas-phase NO_2 into particle-phase NO_3 aerosols lead to an overestimation of nitrate aerosols over India. The MOSAIC aerosol scheme used in MATCHA relies on several factors for gas-to-particle partitioning: 1) availability of hydroxyl (OH) radicals for oxidizing NO_2 to HNO_3, 2) high aerosol concentrations promoting HNO_3 condensation, and 3) favorable conditions such as lower ambient temperatures and higher RH for HNO_3-to-NO_3 partitioning. Several studies have mentioned the overestimation of nitrates in models (Sha et al., 2022;

https://doi.org/10.5194/essd-2025-275
Preprint. Discussion started: 11 September 2025

Zakoura and Pandis, 2018). Figure S5 compares daily tropospheric column NO_2 concentrations from MATCHA (for 2014) to satellite-based observations from the Ozone Monitoring Instrument (OMI) (Boersma et al., 2017). In summer, a distinct negative bias in tropospheric column NO_2 (> 1.4 molecules cm^{-2}) is seen over India, suggesting excessive conversion of gas-phase NO_2 to nitrate aerosols, which increases $PM_{2.5}$ concentrations in MATCHA.

4. The high positive biases in $PM_{2.5}$ over eastern China can be attributed to discrepancies in the emission inventory. Specifically, the CAMS-ANT-GLOB inventory used in MATCHA fails to accurately capture the observed decreasing trend in emissions over China (Soulie et al., 2023). This discrepancy likely results in an overestimation of $PM_{2.5}$ in MATCHA for this region.

These findings highlight key limitations in the representation of marine aerosols, wet scavenging processes, gas-particle partitioning, and emission trends within the MATCHA model, each of which contributes to the positive biases in surface $PM_{2.5}$.

Similar to $PM_{2.5}$, MATCHA exhibits significant biases in PM_{10}, particularly during the summer season (Fig. 14). To understand if errors in simulations of natural aerosols are the primary contributors to biases in PM_{10}, we recalculated the mean bias in PM_{10} by removing the contributions of these species from MATCHA simulations, as shown in Fig. 14 and the domain-wide statistics in Table 4. As in $PM_{2.5}$, the seasonality of PM_{10} concentrations is consistent between OpenAQ observations and MATCHA. After removing sea salt and dust that are overestimated by MATCHA (Table 4), the mean bias in winter decreases substantially from 203 to 62 μg/m³, and in spring, the bias reduces from 39 μg/m³ to just 9 μg/m³. Overall, removing sea salt and dust improves the mean bias, with the largest reduction observed during winter. Despite this adjustment, MATCHA still shows a consistent positive bias, particularly in winter, indicating an important contribution of error in anthropogenic emissions to the overestimation of PM_{10} by MATCHA. Compared to $PM_{2.5}$, the strong positive wintertime bias in PM_{10} is more localized, primarily observed over eastern China, while $PM_{2.5}$ biases are prominent over both India and China. For the adjusted surface PM_{10} concentrations (excluding sea salt and dust), the highest domain-wide bias occurs in winter (MB of 62 μg/m³) and the lowest in fall (MB of 6 μg/m³). RMSE values are highest in winter (100 μg/m³) and lowest in summer (42 μg/m³), highlighting greater variability in winter. Correlation values remain moderate overall, peaking in summer (R of 0.5) except in spring (R of -0.01). KGE values suggest the best model agreement of the adjusted PM_{10} values in summer and fall. If sea salt and dust are not removed, the positive biases increase substantially, particularly in winter (from 62 μg/m³ to 203 μg/m³) and fall (from 6 μg/m³ to 73 μg/m³), further emphasizing the contribution of these species to the observed overestimation in MATCHA.

Surface CO, in contrast, is consistently underestimated by MATCHA across all seasons, although the biases are smaller compared to $PM_{2.5}$ and PM_{10}. In winter, high surface CO concentrations are observed over the Indo-Gangetic Plain and eastern China in both OpenAQ and MATCHA (Fig. 15). MATCHA exhibits strong seasonality in surface CO, similar to OpenAQ, albeit with a negative bias persistent across the domain and all seasons. The persistent negative bias across all seasons is likely due to constraints imposed by the daily data assimilation (DA) system used in MATCHA. The domain-wide smallest negative biases occur in summer (MB of -197 ppbv), while the largest is seen in fall (MB = -246 ppbv). Correlation values for CO are

https://doi.org/10.5194/essd-2025-275
Preprint. Discussion started: 11 September 2025

consistently moderate across seasons (R of 0.3-0.4), with the highest correlation in winter (R of 0.35). KGE values suggest similar moderate agreement across the seasons (0.2-0.3), with the best performance during winter (KGE of 0.3). These biases are likely driven by regional inaccuracies in surface CO emissions (Gaubert et al., 2020; Ojha et al., 2016), particularly over India and China, despite the constraints from daily MOPITT CO profile assimilation, which may primarily influence CO in the free troposphere rather than its surface concentrations (Deeter et al., 2022; Hooghiemstra et al., 2012; Jiang et al., 2015).

3.3.3 Surface BC

Atmospheric black carbon (BC) abundances near the surface are evaluated in Fig. 16. MATCHA shows a strong overestimate of surface BC concentrations over India during winter, while underestimates are observed in most other seasons. Both observations and model simulations exhibit strong seasonality in BC abundances. Higher-elevation regions, such as the Tibetan Plateau and the Himalayas, consistently show underestimation by the model across all seasons, whereas urban areas at lower elevations exhibit weak positive biases. This spatial pattern of the biases and reproduction of the BC seasonal cycle by MATCHA suggests that the elevated PM burden previously noted in MATCHA is primarily driven by natural aerosol emissions (e.g., sea salt and dust) rather than anthropogenic sources such as BC. Across the domain, the biases in MATCHA are negative across all seasons except in winter, where the bias is positive and highest (MB of 3 μg/m³). Winter also exhibits the highest RMSE (8 μg/m³), reflecting significant variability in BC concentrations. This is also suggested by the positive bias in PM_{10} observed in Fig. 14 (Sect. 3.3.2) after removing the contribution of natural aerosols to the total PM_{10} burden. The largest negative bias occurs in spring (MB = -2 μg/m³), while summer shows the lowest bias (-1 μg/m³) and RMSE (2 μg/m³). Correlations between MATCHA and observations are relatively strong across all seasons (0.6-0.7), with the highest correlation observed in spring and winter (R of 0.75), while the lowest is in summer (R of 0.6). KGE values indicate consistent performance across all seasons, with the best agreement in fall (0.5) followed by winter (0.4). These results indicate that while MATCHA captures the general spatial and seasonal trends of BC, it struggles to accurately represent BC abundances in some regions, particularly over India during winter and in high-altitude regions throughout the year.

3.3.4 Additional surface aerosol species

To gain insights into whether anthropogenic emission sources might be contributing to errors in MATCHA simulated PM, we compare surface observations of PM_{10} as well as eight other aerosol for an urban site named Kanpur (26.5°N, 80.3°E) in the northwestern India with MATCHA based on daily samples collected between January of 2007 to March of 2008 (Ram et al., 2010). The observations are daily averages (between 0600 and 1800 local time) collected using high-volume samplers. Samples for the monsoon months (July to September) were not collected and hence not shown in Fig. 17. We see reasonable agreement for bulk PM_{10} and several species (as seen by KGE ≥ 0.2 in Fig. 17) in all months except for Ca^{2+}, Na^{+}, and Cl^{-}, where we see significant differences (as seen by the negative KGEs). Na^{+} shows significant overestimations across all months, Ca^{2+} is almost absent in most months, while Cl^{-} concentrations are significantly high during spring. These biases can stem from the parameterizations related to natural aerosols and dust speciation, (as discussed in Sect. 3.3.2). In Fig. A6, we show the monthly

https://doi.org/10.5194/essd-2025-275
Preprint. Discussion started: 11 September 2025

average percentage contributions of these eight species to total PM, from the observations and MATCHA (based on collocated observations and model simulations for the site at Kanpur, as well as based on the average 17-year contribution over the site from MATCHA). Observations reveal organic carbon (OC) as the dominant contributor, peaking in winter months (November–February), with consistently higher contributions compared to MATCHA. Contribution of SO_4^{2-} is overestimated in MATCHA during summer (May–June), particularly in the 17-year average, which shows a more uniform seasonal contribution. NO_3^- contribution peaks during winter and is consistently underestimated by MATCHA. Thus, observations indicate stronger seasonal variability in components like OC, NO_3^-, and NH_4^+, while MATCHA (both 1-year and 17-year averages) smooths out these fluctuations. MATCHA overpredicts SO_4^{2-} and NH_4^+ m during non-winter months while underestimating components like OC and elemental carbon (EC), particularly in winter. These results point towards more accurate emission inventories and improved parameterizations of sea salt and dust speciation for accurate simulation of PM within models.

3.3.5 Source attribution of BC

We used the tagged-tracers of anthropogenic BC across ten Asian regions within MATCHA, along with biomass burning and boundary inflow, to quantify the relative contribution of black carbon (BC) from anthropogenic emissions originating in ten different countries or regions, as well as sectors to the BC burden in HMA. Figure 18a–c shows the monthly averaged total column burden BC (tBC in $\mu g/m^2$) for these tracers across 17 years (2003-2019), spatially averaged over HMA (based on aggregated first-order regions from the Randolph Glacier Inventory (RGI) v6; see Sect. 3.2.3). Winter exhibits the highest tBC for total BC (total refers to the sum of anthropogenic, BB, and transboundary sources of BC) and the highest variability, while anthropogenic tBC shows the highest variability across all months. Biomass burning tBC peaks in March–April, coinciding with seasonal agricultural burning in south, east, and southeast Asia (Wiedinmyer et al., 2023). Anthropogenic tBC is primarily dominated by China and India (Fig. 18b), with China showing the largest seasonal variability, particularly in winter. Pakistan and the Tibetan Plateau (Fig. 18c) show peak concentrations and variability during the summer and monsoon months, followed by the remaining regions of the model domain (Rest of Asia; Fig. 18c). In Fig. 18d, we show the average ~17-year contribution of 12 tags (ten tagged regions + biomass burning + boundary inflow) to tBC over the six major glacier regions of HMA. India and China are the dominant sources across all seasons, with India contributing more to the western and central regions of HMA (e.g., HKH, HIM) and China to the northern and eastern regions (e.g., HTQ, TNS). Contributions to anthropogenic tBC from other regions, like southeast Asia and Pakistan, increase during spring and summer, especially in the southern areas of HMA. Contributions from boundary inflow also dominate for most of the glacier regions across all seasons, which indicates the significance of long-range transport of BC at high-altitude areas. The tagged-tracers within MATCHA can thus assist in inferring the contribution of major Asian regions to anthropogenic BC, as well as separating the influence of biomass burning and boundary inflow over the domain. This source attribution capacity can help determine which emission sources affect the BC loading over HMA during different seasons. These estimates can provide critical input for process-level studies in

https://doi.org/10.5194/essd-2025-275
Preprint. Discussion started: 11 September 2025

investigating aerosol-meteorology-snow interactions over HMA, where snowmelt impacts downstream freshwater availability (Roychoudhury et al., 2022,2025).

4 Summary and Conclusions

This paper presents a novel hydroclimate-chemical regional reanalysis, MATCHA (Model for Atmospheric Transport and Chemistry in Asia), to support research on light-absorbing particles and their impacts on the cryosphere over High Mountain Asia (HMA). Compared to existing global reanalyses, MATCHA offers strong coupling between atmospheric chemistry, land surface processes, and aerosol–snowpack interactions, thus making it the only regionally focused, fully coupled chemical reanalysis currently available for HMA. Our recent work analyzed the degree of coupling incorporated in the model framework across MATCHA, ERA5, and MERRA-2 over HMA and found that MATCHA reflects the highest degree of coupling that contributes to a more accurate representation of snow cover during the snow ablation season over HMA (Roychoudhury et al., 2025). MATCHA assimilates nearly two decades (2003–2019) of MODIS AOD and MOPITT CO satellite retrievals into the WRF-Chem–CLM–SNICAR modeling framework, producing a 12 km resolution dataset with hourly to three-hourly temporal outputs for ~17 years. Key features of this dataset include the simulation of BrC aerosol processes, source-tagged-tracers of BC, and coupling of aerosols with radiation and snowpack processes. We evaluate the dataset against multiple in situ and satellite observations, and the results provide detailed insights into model biases and uncertainties across meteorological, land, and atmospheric composition variables.

We summarize the evaluation results as follows,

1. Surface temperature, relative humidity, wind speed, and their vertical profiles show the strongest overall skill (KGE of 0.65–1.0) across all evaluated variables. Minor issues include a cold, dry bias in winter at high elevations and a domain-wide 10 m wind overestimate; in the upper troposphere, temperature/RH are slightly low, whereas winds are high.
2. Daily accumulated precipitation compared with satellite measurements from IMERG Final Runs is realistically represented, with the monsoon rainband and localized features over the Indian sub-continent reproduced the best (domain-mean KGE ≈ 0.6 in JJA). Systematic underestimation persists over high terrain, peaking in summer, reflecting the difficulty most regional models face with orographic convection.
3. Diurnal and seasonal cycles of PBLH are effectively captured, but MATCHA underestimates early-morning PBLH (00 UTC) over India/coastal zones and overestimates afternoon PBLH (12 UTC) across most sites in summer, leading to biases that trace back to the YSU PBL scheme, limited vertical resolution, complex terrain, as well as transition of daytime to nocturnal PBLH. The skill is highest in spring, and the magnitude of the biases is within the ranges reported in the literature.
4. Over HMA, MATCHA recreates the observed snow cover fraction seasonality (R of 0.7–0.9; winter KGE ≈ 0.7), with only minor negative biases during the melt period. This can be attributed to negative SSA550 biases observed

https://doi.org/10.5194/essd-2025-275
Preprint. Discussion started: 11 September 2025

across the domain, suggesting more absorbing aerosols, which lead to increased snowmelt, with the CLM-SNICAR coupling within MATCHA.

5. Spatial and seasonal patterns of AOD550 and SSA550 are well reproduced, capturing major pollution hot spots. MATCHA overestimates summer AOD550 over the Indian sub-continent and SE Asia and exhibits a larger-than-typical negative SSA550 bias, pointing to shortcomings in emission inventories and model parameterizations of natural emissions of aerosols.
6. Surface particulate matter (PM) measurements, based on a network of over 1500 sites from OpenAQ, exhibit seasonality and spatial patterns consistent with Asian pollution hotspots identified in MATCHA. However, MATCHA significantly overestimates surface $PM_{2.5}$ and PM_{10} concentrations, particularly during winter. Preliminary analysis attributes this bias to four interlinked factors: (i) overstated Arabian Sea winds inflating sea salt and dust emissions, (ii) underestimated monsoon rainfall reducing wet scavenging, (iii) excessive NO_2-NO_3^- partitioning enhancing nitrate mass, and (iv) the absence of China's recent downward PM trends in the CAMS-GLOB-ANT emission inventory used as input for MATCHA.
7. Furthermore, based on aerosol chemical composition measurements of different chemical species from an urban site in Kanpur, India, MATCHA captures the bulk PM_{10} load and the seasonality of most constituents. However, it overestimates marine and soil tracers (Na^+, Cl^-, Ca^{2+}) while failing to reproduce the strong winter peaks in carbonaceous (OC, EC) and nitrate (NO_3^-) aerosols seen in observations. Specifically, the reanalysis underestimates OC, EC, and NO_3^- during November–February and overestimates SO_4^{2-} and NH_4^+ during the pre-monsoon and post-monsoon periods. These discrepancies suggest overly efficient secondary inorganic aerosol production and underrepresentation of primary combustion sources in MATCHA. Overall, the species-specific biases are consistent with the broader regional PM discrepancies and underscore the need for refined emission inventories and improved aerosol chemistry parameterizations in future MATCHA updates.
8. The tagged-tracers from MATCHA are a unique feature within the MATCHA reanalysis that offer insights into the sources (both regional and sectoral) of BC in the domain across two decades. Initial analysis of these tracers reveals that the total column burden of BC peaks in winter, with China and India being the dominant anthropogenic sources of the total column burden of anthropogenic BC. China shows the greatest seasonal variability in total column burden of anthropogenic BC, particularly over eastern and northern China, while India contributes more to the total column burden over western and central HMA. Biomass burning BC peaks in March to April, while trans-boundary inflow (representing the influence of emissions sources located outside our domain) is significant across all seasons. Regional contributions vary, with additional influences from Pakistan, the Tibetan Plateau, and southeast Asia during spring and summer.

Overall, MATCHA effectively captures the spatiotemporal distributions and general trends of key meteorological and chemical variables. Its performance in representing meteorological fields is particularly strong, with KGE values greater than 0.65. In

https://doi.org/10.5194/essd-2025-275
Preprint. Discussion started: 11 September 2025

contrast, variables associated with land–atmosphere interactions, such as precipitation, planetary boundary layer height (PBLH), and snow cover fraction (SCF), exhibit moderate skill, with maximum KGE values around 0.7. For atmospheric composition, MATCHA accurately reproduces the spatiotemporal patterns of anthropogenic aerosols, although biases in natural aerosol species contribute to discrepancies in surface $PM_{2.5}$ and PM_{10} concentrations. Surface CO measurements, however, are more tightly constrained, likely due to the implementation of a daily assimilation workflow within MATCHA. The identified biases underscore the need for improved parameterizations of surface-atmosphere interactions, cloud microphysics, and precipitation processes, as well as the improvements required in emission inventories to enhance the simulation of chemical species. While this study presents an initial assessment of 12 variables spanning meteorology, land–atmosphere coupling, and atmospheric composition at a seasonal scale, further analyses are necessary. Specifically, extending evaluations to capture regional differences and long-term trends over the past 17 years, across finer spatiotemporal scales, will be crucial to fully understand the value added by the high-resolution datasets provided by MATCHA for a region as complex as Asia. Moreover, rigorous intercomparisons with other reanalysis products are needed, alongside investigations into additional essential climate variables offered by MATCHA.

Despite its limitations, MATCHA represents a state-of-the-art, high-resolution, long-term regional chemical reanalysis for Asia, specifically designed to address major gaps in representing LAPs and their feedbacks in climate-vulnerable regions like HMA. MATCHA's most distinguishing feature is its explicit coupling of aerosols, radiation, and snowpack processes, allowing it to capture complex non-linear interactions among atmospheric composition, meteorology, and the cryosphere (e.g., snow hydrology). In contrast to global reanalyses such as ERA5 and MERRA-2, which offer longer periods exceeding 17 years but operate at coarser resolutions (25-50 km) and lack aerosol-snow coupling, MATCHA provides higher-resolution outputs (12 km) and includes parameterizations that allow for more accurate representation of the cryosphere-atmosphere feedbacks critical to regional hydroclimate. Among current reanalyses, MERRA-2 includes aerosol–radiation interactions but not aerosol–snow feedbacks, limiting its utility for evaluating LAPs' impacts on snowpack and glacier melt. MATCHA further distinguishes itself by assimilating nearly two decades of MODIS AOD and MOPITT carbon monoxide retrievals and simulating the full cycle of black carbon (BC), dust, and brown carbon (BrC), including their deposition on snow surfaces. Beyond retrospective analysis across the last two decades, MATCHA offers high-resolution datasets that facilitate assessments of the sensitivity of regional hydroclimate and air quality to nonlinear aerosol–climate interactions. Additionally, tagging of regional sources of anthropogenic BC with MATCHA enable more precise source attribution studies, providing critical insights to inform emission-driven policy interventions, particularly for hydroclimate-vulnerable regions such as the HMA.

https://doi.org/10.5194/essd-2025-275
Preprint. Discussion started: 11 September 2025

Open Access Earth System Science Data Discussions

5 Appendix

A1 Additional Figures

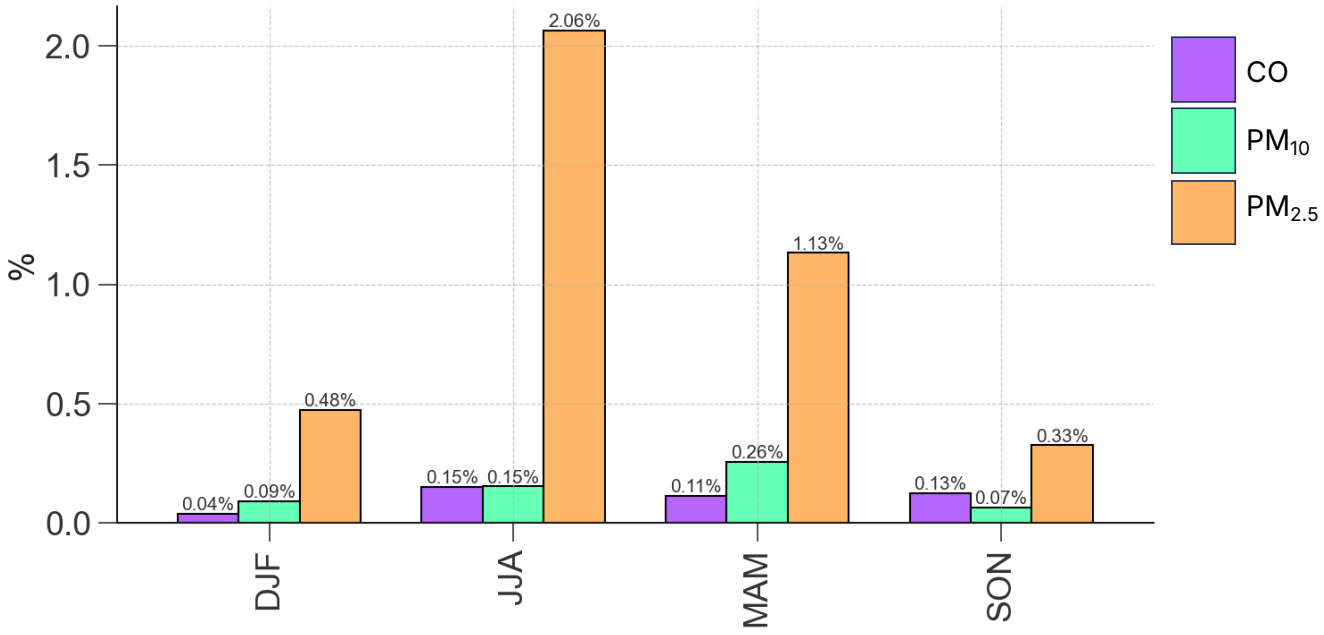


Figure A1: Percentage of outliers removed from retrieved OpenAQ data of surface CO, PM_{10}, and $PM_{2.5}$ (purple, green and orange, respectively) using the outlier detection system in Wu et al., 2018 (see main text).

https://doi.org/10.5194/essd-2025-275
Preprint. Discussion started: 11 September 2025

Open Access Earth System Science Data Discussions

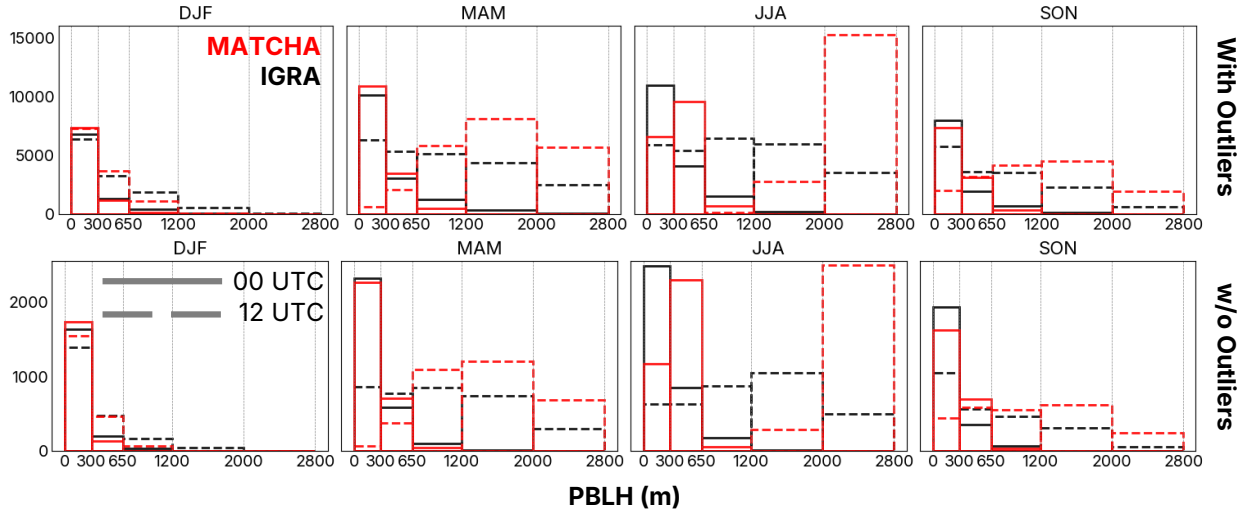


Figure A2: Histograms of both model (red) and observed (black) PBLH values compared across seasons and standard radiosonde launch times (solid for 00 UTC and dashed for 12 UTC) before and after applying the outlier-detection approach described in Section 2.2.4 of the main text.

https://doi.org/10.5194/essd-2025-275
Preprint. Discussion started: 11 September 2025

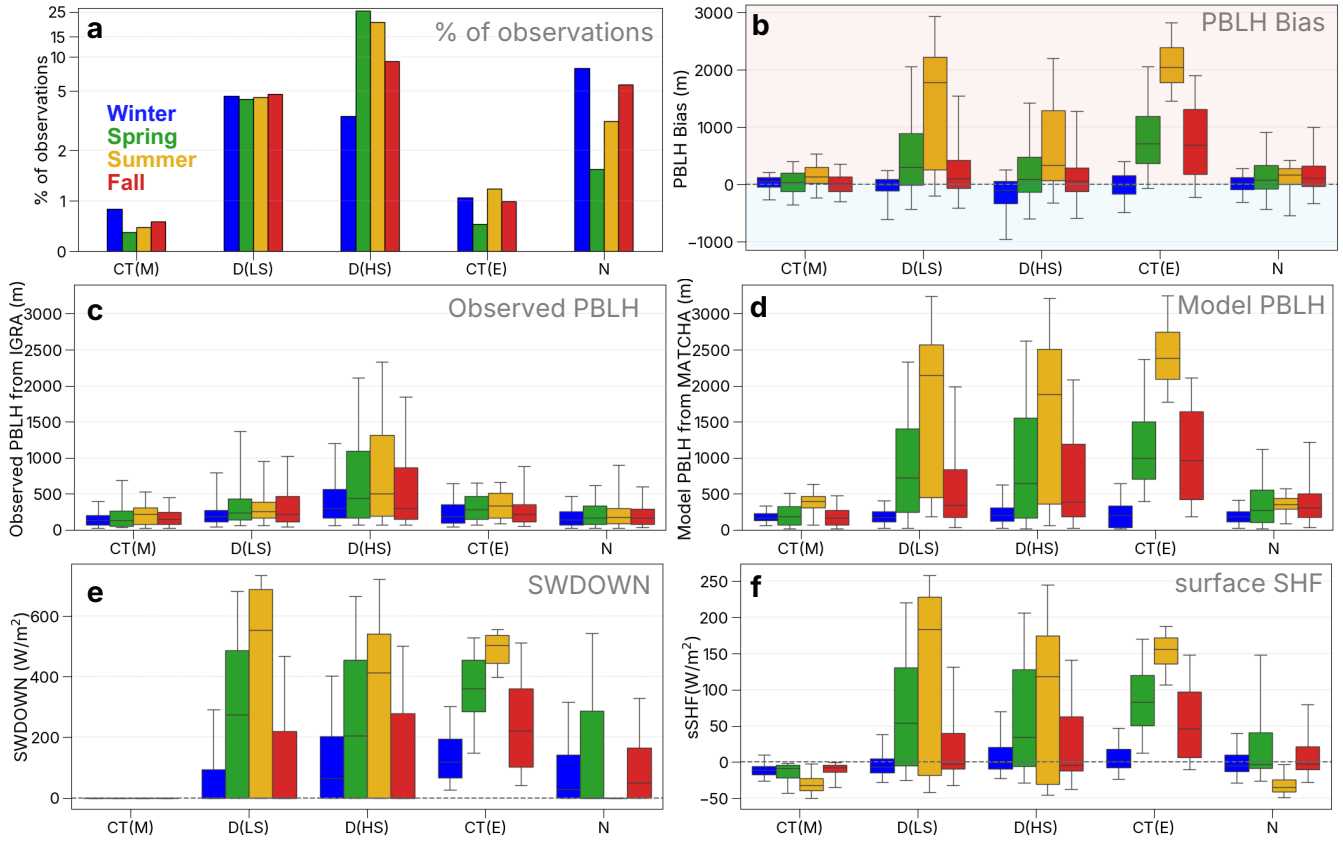
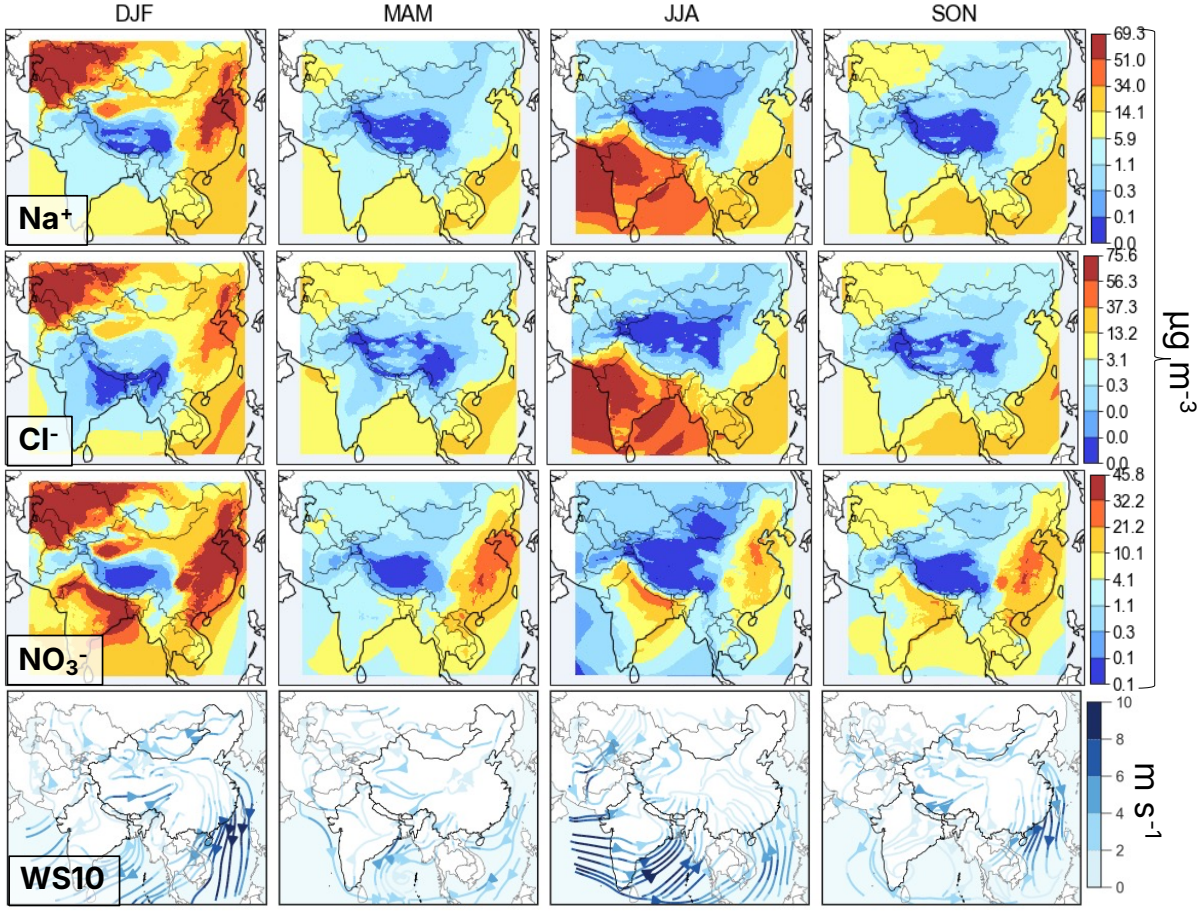


Figure A3: (a) Fraction (in %) of all 00 UTC and 12 UTC radiosonde launches in the domain across different solar regimes (CT(M) = Morning Civil Twilight; D(LS) = Day – Low Sun; D(HS) = Day – High Sun; CT(E) = Evening Civil Twilight; N = Night) for each season. Solar regimes were defined by computing the solar elevation angle at the exact launch time of each radiosonde observations and classifying them into five exclusive categories: N (≤ –6°), CT (–6° to 0°, split into morning or evening by solar azimuth), D (LS) (0° to 10°), and D(HS) (> 10°). (b) PBLH bias (model-observed), (c) observed PBLH, and (d) modeled PBLH from IGRA and MATCHA across the solar regimes and seasons. (e) Downward shortwave radiation flux (SWDOWN) and (f) surface sensible heat flux (sSHF) collocated at IGRA sites from MATCHA across solar regimes and seasons.

https://doi.org/10.5194/essd-2025-275
Preprint. Discussion started: 11 September 2025



940 **Figure A4:** Seasonal averages of surface sea-salt partitioned to sodium and chloride (in μg/m^3), surface nitrate aerosols (in μg/m^3), and wind speed at 10 m (in m/s) over the model domain.

https://doi.org/10.5194/essd-2025-275
Preprint. Discussion started: 11 September 2025

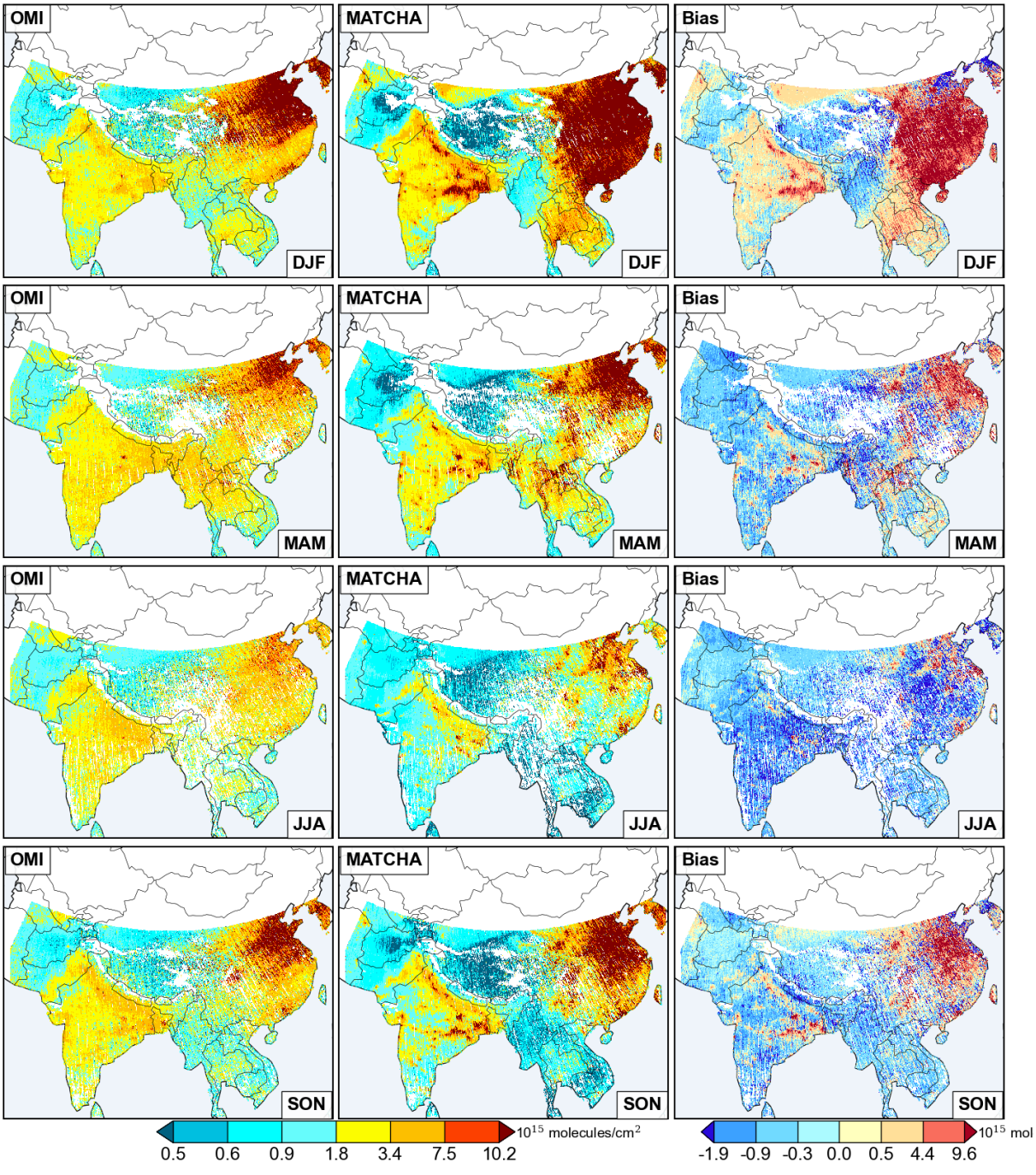


Figure A5: Seasonal averages for the year 2014 of tropospheric column NO_2 concentrations (molecules cm^{-2}) from OMI satellite observations (L2 products) regridded to MATCHA's 12 km resolution (left column), MATCHA (middle column), and the associated model bias (MATCHA-OMI, right column).

https://doi.org/10.5194/essd-2025-275
Preprint. Discussion started: 11 September 2025

Open Access Earth System Science Data Discussions

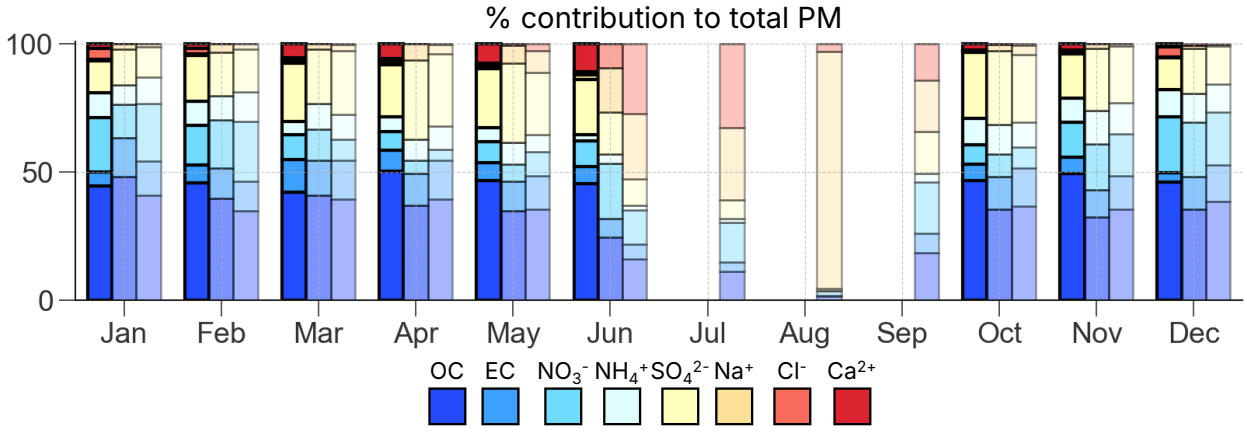


Figure A6: Monthly averaged percentage contribution of eight chemical species to total particulate matter over the Kanpur site based on ~1 year of observations from the site (first group of bars), ~1 year of model simulation over the site (second group) and 17-year contribution (third group) from the model simulation over the site.

https://doi.org/10.5194/essd-2025-275
Preprint. Discussion started: 11 September 2025

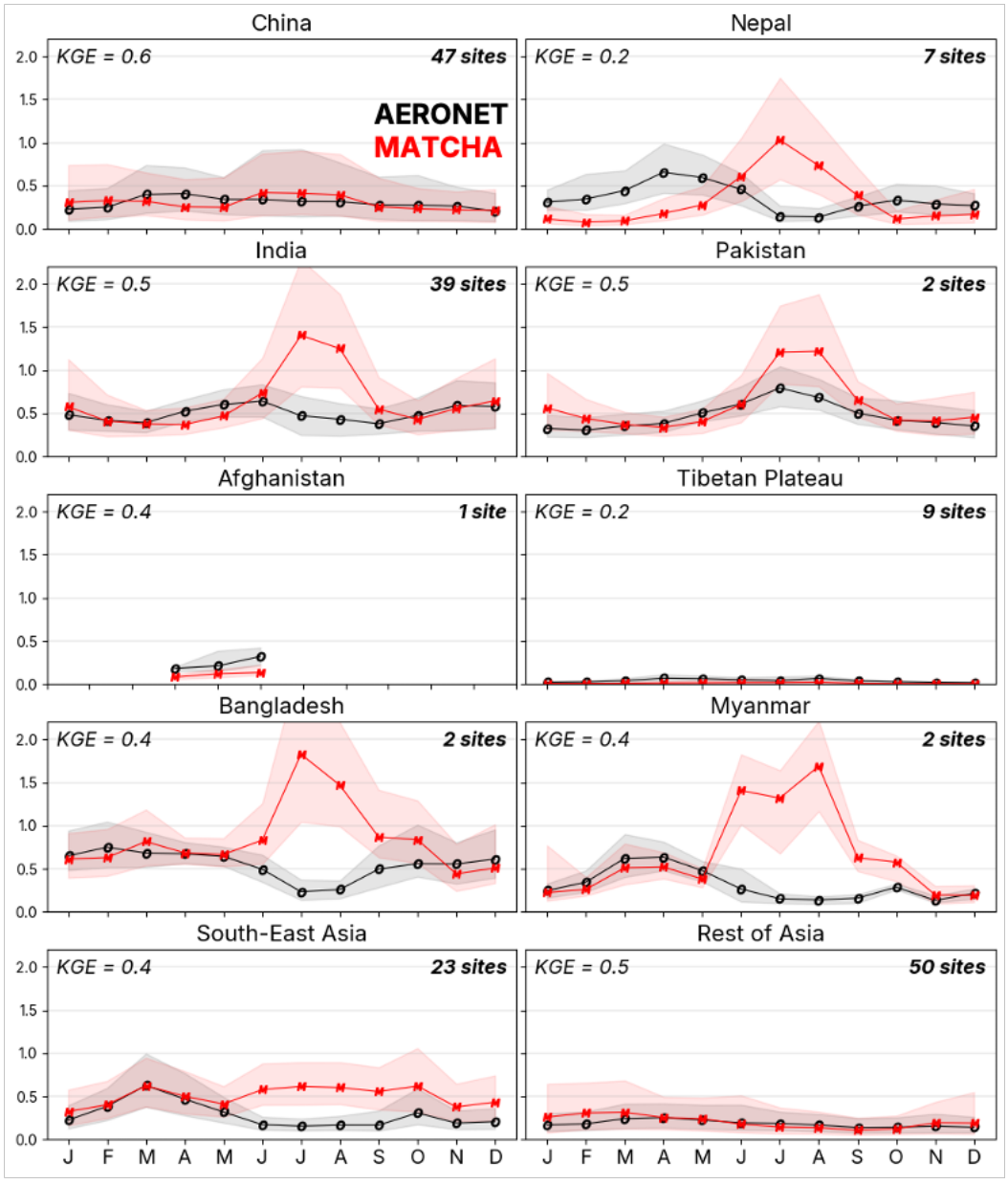


Figure A7: Monthly median AOD at 550 nm with their interquartile ranges across AERONET sites (black) compared with corresponding values from MATCHA (red) across the ten tagged tracer regions within MATCHA (defined in the main text).

https://doi.org/10.5194/essd-2025-275
Preprint. Discussion started: 11 September 2025

6 Data availability

The MATCHA reanalysis is available through the National Snow and Ice Data Center (NSIDC) (https://nsidc.org/data/hma2_matcha/versions/1; DOI: https://doi.org/10.5067/CG4OT8DJX2Z7; Kumar et al., 2024). MODIS SCF (v6.1), AOD (v6.1), and MOPITT CO (v8) were downloaded from NASA Earthdata (https://search.earthdata.nasa.gov/search). AERONET AOD and associated inversion products (v3) were downloaded from the AERONET website (https://aeronet.gsfc.nasa.gov/new_web/download_all_v3_aod.html). OpenAQ data can be accessed through the OpenAQ API (https://docs.openaq.org/about/about). IMERG Final Run daily data were downloaded from NASA GES DISC (https://disc.gsfc.nasa.gov/). IGRA sounding data and derived products were downloaded from NOAA NCEI's website (https://www.ncei.noaa.gov/products/weather-balloon/integrated-global-radiosonde-archive). ISD products for sub-daily surface meteorological measurements across sites can be obtained from NOAA NCEI's website (https://www.ncei.noaa.gov/products/land-based-station/integrated-surface-database).

7 Author contribution

CR contributed to the conceptualization, methodology, formal analysis, investigation, visualization, writing- original draft, reviewing, and editing. RK contributed to the conceptualization, investigation, project administration, validation, resources, software, funding acquisition, writing-review & editing. CH contributed to the resources, software, validation, project administration, funding acquisition, writing-review & editing. WC contributed to the data curation, software, validation, writing-review & editing. KR contributed to the validation, writing-review & editing. NM contributed to validation, writing-review & editing. AFAJ contributed to the conceptualization, supervision, funding acquisition, and writing-review & editing.

8 Competing Interests

The authors declare that they have no competing interests.

9 Acknowledgement

This work is supported by a NASA HiMAT2 grant (#NNH19ZDA001N-HMA). HiMAT2 is an interdisciplinary multi-investigator effort to understand the cryospheric and hydrological state of HMA. This work is in tandem with the goals of the Aerosol subgroup under HiMAT2, to quantify the deposition of aerosols over snow in HMA. We also acknowledge the National Center for Atmospheric Research (NCAR) (sponsored by the National Science Foundation (NSF)) for assisting this ongoing study through high-performance computing support from Cheyenne and Derecho HPC from NCAR's Computational and Information Systems Laboratory, as well as for model runs and data curation.

https://doi.org/10.5194/essd-2025-275
Preprint. Discussion started: 11 September 2025

References

Ansari, K. and Ramachandran, S.: Optical and physical characteristics of aerosols over Asia: AERONET, MERRA-2 and CAMS, Atmospheric Environment, 326, 120470, https://doi.org/10.1016/j.atmosenv.2024.120470, 2024.

Arellano Jr., A. F., Hess, P. G., Edwards, D. P., and Baumgardner, D.: Constraints on black carbon aerosol distribution from Measurement of Pollution in the Troposphere (MOPITT) CO, Geophysical Research Letters, 37, https://doi.org/10.1029/2010GL044416, 2010.

Ashrit, R., Indira Rani, S., Kumar, S., Karunasagar, S., Arulalan, T., Francis, T., Routray, A., Laskar, S. I., Mahmood, S., Jermey, P., Maycock, A., Renshaw, R., George, J. P., and Rajagopal, E. N.: IMDAA Regional Reanalysis: Performance Evaluation During Indian Summer Monsoon Season, Journal of Geophysical Research: Atmospheres, 125, e2019JD030973, https://doi.org/10.1029/2019JD030973, 2020.

Barros, A. P. and Arulraj, M.: Remote Sensing of Orographic Precipitation, in: Satellite Precipitation Measurement: Volume 2, edited by: Levizzani, V., Kidd, C., Kirschbaum, D. B., Kummerow, C. D., Nakamura, K., and Turk, F. J., Springer International Publishing, Cham, 559–582, https://doi.org/10.1007/978-3-030-35798-6_6, 2020.

Boersma, K. F., Eskes, H., Richter, A., De Smedt, I., Lorente, A., Beirle, S., Van Geffen, J., Peters, E., Van Roozendael, M., and Wagner, T.: QA4ECV NO2 tropospheric and stratospheric vertical column data from OMI (Version 1.1) [Data set], , https://doi.org/10.21944/qa4ecv-no2-omi-v1.1, 2017.

Bowden, J. H., Otte, T. L., Nolte, C. G., and Otte, M. J.: Examining Interior Grid Nudging Techniques Using Two-Way Nesting in the WRF Model for Regional Climate Modeling, https://doi.org/10.1175/JCLI-D-11-00167.1, 2012.

Bright, J. M. and Gueymard, C. A.: Climate-specific and global validation of MODIS Aqua and Terra aerosol optical depth at 452 AERONET stations, Solar Energy, 183, 594–605, https://doi.org/10.1016/j.solener.2019.03.043, 2019.

Buchholz, R. R., Emmons, L. K., Tilmes, S., and Team, T. C. D.: CESM2.1/CAM-chem instantaneous output for boundary conditions, 2019.

Cannon, F., Carvalho, L. M. V., Jones, C., Norris, J., Bookhagen, B., and Kiladis, G. N.: Effects of topographic smoothing on the simulation of winter precipitation in High Mountain Asia, Journal of Geophysical Research: Atmospheres, 122, 1456–1474, https://doi.org/10.1002/2016JD026038, 2017.

Carvalho, D.: An Assessment of NASA's GMAO MERRA-2 Reanalysis Surface Winds, https://doi.org/10.1175/JCLI-D-19-0199.1, 2019.

Chen, X., Yang, T., Wang, Z., Wang, F., and Wang, H.: An ensemble method for improving the estimation of planetary boundary layer height from radiosonde data, Atmospheric Measurement Techniques, 16, 4289–4302, https://doi.org/10.5194/amt-16-4289-2023, 2023.

Chen, Y., Cheng, Y., Ma, N., Wolke, R., Nordmann, S., Schüttauf, S., Ran, L., Wehner, B., Birmili, W., van der Gon, H. A. C. D., Mu, Q., Barthel, S., Spindler, G., Stieger, B., Müller, K., Zheng, G.-J., Pöschl, U., Su, H., and Wiedensohler, A.: Sea salt emission, transport and influence on size-segregated nitrate simulation: a case study in northwestern Europe by WRF-Chem, Atmospheric Chemistry and Physics, 16, 12081–12097, https://doi.org/10.5194/acp-16-12081-2016, 2016.

Cheng, W. Y. Y. and Steenburgh, W. J.: Evaluation of Surface Sensible Weather Forecasts by the WRF and the Eta Models over the Western United States, https://doi.org/10.1175/WAF885.1, 2005.

https://doi.org/10.5194/essd-2025-275
Preprint. Discussion started: 11 September 2025

Collins, W. J., Lamarque, J.-F., Schulz, M., Boucher, O., Eyring, V., Hegglin, M. I., Maycock, A., Myhre, G., Prather, M., Shindell, D., and Smith, S. J.: AerChemMIP: quantifying the effects of chemistry and aerosols in CMIP6, Geoscientific Model Development, 10, 585–607, https://doi.org/10.5194/gmd-10-585-2017, 2017.

Cuchiara, G. C. and Rappenglück, B.: Single-column model and large eddy simulation of the evening transition in the planetary boundary layer, Environmental Fluid Mechanics, 17, 777–798, https://doi.org/10.1007/s10652-017-9518-z, 2017.

Dee, D. U., S. ;. Simmons, Adrian; Berrisford, Paul; Poli, Paul; Kobayashi, Shinya; Andrae, Ulf; Balmaseda, Magdalena; Balsamo, Gianpaolo; Bauer, Peter; Bechtold, Peter; Beljaars, Anton; van de Berg, L. ;. Bidlot, Jean; Bormann, Niels; Delsol, C. ;. Dragani, Rossana; Fuentes, Manuel; Geer, Alan J. ;. Haimberger, Leopold; Healy, Sean; Hersbach, Hans; Hólm, Elías; Isaksen, Lars; Kallberg, P. ;. Köhler, Martin; Matricardi, Marco; McNally, A. P. ;. Monge-Sanz, B. M. ;. Morcrette, Jean-Jacques; Park, B. K. ;. Peubey, Carole; de Rosnay, P. ;. Tavolato, Christina; Thépaut, Jean-Noël; Vitart, Frederic: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, Quarterly Journal of the Royal Meteorological Society, 137, 553–597, https://doi.org/10.1002/qj.828, 2011.

Deeter, M. N., Francis, G., Gille, J. C., Mao, D., Martínez-Alonso, S., Worden, H. M., Ziskin, D., Drummond, J. R., Commane, R., Diskin, G. S., and McKain, K.: The MOPITT Version 9 CO product: sampling enhancements and validation, Atmospheric Measurement Techniques, 15, https://doi.org/10.5194/amt-15-2325-2022, 2022.

Demuzere, M., Kittner, J., Martilli, A., Mills, G., Moede, C., Stewart, I. D., van Vliet, J., and Bechtel, B.: A global map of local climate zones to support earth system modelling and urban-scale environmental science, Earth System Science Data, 14, 3835–3873, https://doi.org/10.5194/essd-14-3835-2022, 2022.

Dollan, I. J., Maina, F. Z., Kumar, S. V., Nikolopoulos, E. I., and Maggioni, V.: An assessment of gridded precipitation products over High Mountain Asia, Journal of Hydrology: Regional Studies, 52, 101675, https://doi.org/10.1016/j.ejrh.2024.101675, 2024.

Dong, Y., Li, J., Yan, X., Li, C., Jiang, Z., Xiong, C., Chang, L., Zhang, L., Ying, T., and Zhang, Z.: Retrieval of aerosol single scattering albedo using joint satellite and surface visibility measurements, Remote Sensing of Environment, 294, 113654, https://doi.org/10.1016/j.rse.2023.113654, 2023.

Dubovik, O., Smirnov, A., Holben, B. N., King, M. D., Kaufman, Y. J., Eck, T. F., and Slutsker, I.: Accuracy assessments of aerosol optical properties retrieved from Aerosol Robotic Network (AERONET) Sun and sky radiance measurements, Journal of Geophysical Research: Atmospheres, 105, 9791–9806, https://doi.org/10.1029/2000JD900040, 2000.

Durre, I., Vose, R. S., and Wuertz, D. B.: Overview of the Integrated Global Radiosonde Archive, https://doi.org/10.1175/JCLI3594.1, 2006.

Durre, I., Yin, X., Vose, R. S., Applequist, S., and Arnfield, J.: Enhancing the Data Coverage in the Integrated Global Radiosonde Archive, Journal of Atmospheric and Oceanic Technology, 35, 1753–1770, https://doi.org/10.1175/JTECH-D-17-0223.1, 2018.

Easter, R. C., Ghan, S. J., Zhang, Y., Saylor, R. D., Chapman, E. G., Laulainen, N. S., Abdul-Razzak, H., Leung, L. R., Bian, X., and Zaveri, R. A.: MIRAGE: Model description and evaluation of aerosols and trace gases, Journal of Geophysical Research: Atmospheres, 109, https://doi.org/10.1029/2004JD004571, 2004.

Emmons, L. K., Walters, S., Hess, P. G., Lamarque, J.-F., Pfister, G. G., Fillmore, D., Granier, C., Guenther, A., Kinnison, D., Laepple, T., Orlando, J., Tie, X., Tyndall, G., Wiedinmyer, C., Baughcum, S. L., and Kloster, S.: Description and evaluation of the Model for Ozone and Related chemical Tracers, version 4 (MOZART-4), Geoscientific Model Development, 3, 43–67, https://doi.org/10.5194/gmd-3-43-2010, 2010.

https://doi.org/10.5194/essd-2025-275
Preprint. Discussion started: 11 September 2025

Faber, E., Rocha-Lima, A., Colarco, P., and Baker, B.: Evaluation of 10-m Wind Speed From ISD Meteorological Stations and the MERRA-2 Reanalysis: Impacts on Dust Emission in the Arabian Peninsula, Journal of Geophysical Research: Atmospheres, 129, e2024JD040885, https://doi.org/10.1029/2024JD040885, 2024.

Fast, J. D. and others: Modeling regional aerosol and tracer transport using the WRF-chemistry model during the MIRAGE field experiment, Journal of Geophysical Research: Atmospheres, 111, D19303, https://doi.org/10.1029/2005JD006168, 2006.

Fast, J. D., Gustafson Jr., W. I., Easter, R. C., Zaveri, R. A., Barnard, J. C., Chapman, E. G., Grell, G. A., and Peckham, S. E.: Evolution of ozone, particulates, and aerosol direct radiative forcing in the vicinity of Houston using a fully coupled meteorology-chemistry-aerosol model, Journal of Geophysical Research: Atmospheres, 111, https://doi.org/10.1029/2005JD006721, 2006.

Flanner, M. G., Zender, C. S., Randerson, J. T., and Rasch, P. J.: Present-day climate forcing and response from black carbon in snow, Journal of Geophysical Research: Atmospheres, 112, https://doi.org/10.1029/2006JD008003, 2007.

Flanner, M. G., Liu, X., Zhou, C., Penner, J. E., and Jiao, C.: Enhanced solar energy absorption by internally-mixed black carbon in snow grains, Atmospheric Chemistry and Physics, 12, 4699–4721, https://doi.org/10.5194/acp-12-4699-2012, 2012.

Flanner, M. Z., Charles S. ;. Hess, Peter; Mahowald, Natalie M. ;. Painter, Thomas H. ;. Ramanathan, Veerabhadran; Rasch, Philip J.: Springtime warming and reduced snow cover from carbonaceous particles, Atmospheric Chemistry and Physics, 9, 2481–2497, https://doi.org/10.5194/acp-9-2481-2009, 2009.

Freitas, S. R., Longo, K. M., Chatfield, R., Latham, D., Silva Dias, M. a. F., Andreae, M. O., Prins, E., Santos, J. C., Gielow, R., and Carvalho, J. A. J.: Including the sub-grid scale plume rise of vegetation fires in low resolution atmospheric transport models, Atmospheric Chemistry and Physics, 7, 3385–3398, https://doi.org/10.5194/acp-7-3385-2007, 2007.

Ganzeveld, L., Bouwman, L., Stehfest, E., van Vuuren, D. P., Eickhout, B., and Lelieveld, J.: Impact of future land use and land cover changes on atmospheric chemistry-climate interactions, Journal of Geophysical Research: Atmospheres, 115, https://doi.org/10.1029/2010JD014041, 2010.

Gaubert, B., Emmons, L. K., Raeder, K., Tilmes, S., Miyazaki, K., Arellano Jr., A. F., Elguindi, N., Granier, C., Tang, W., Barré, J., Worden, H. M., Buchholz, R. R., Edwards, D. P., Franke, P., Anderson, J. L., Saunois, M., Schroeder, J., Woo, J.-H., Simpson, I. J., Blake, D. R., Meinardi, S., Wennberg, P. O., Crounse, J., Teng, A., Kim, M., Dickerson, R. R., He, H., Ren, X., Pusede, S. E., and Diskin, G. S.: Correcting model biases of CO in East Asia: impact on oxidant distributions during KORUS-AQ, Atmospheric Chemistry and Physics, 20, 14617–14647, https://doi.org/10.5194/acp-20-14617-2020, 2020.

Ghude, S. D., Kumar, R., Jena, C., Debnath, S., Kulkarni, R. G., Alessandrini, S., Biswas, M., Kulkrani, S., Pithani, P., Kelkar, S., Sajjan, V., Chate, D. M., Soni, V. K., Singh, S., Nanjundiah, R. S., and Rajeevan, M.: Evaluation of PM2.5 forecast using chemical data assimilation in the WRF-Chem model, Current Science, 118, 1803–1815, 2020.

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

Gong, S. L.: A parameterization of sea-salt aerosol source function for sub- and super-micron particles, Global Biogeochemical Cycles, 17, https://doi.org/10.1029/2003GB002079, 2003.

https://doi.org/10.5194/essd-2025-275
Preprint. Discussion started: 11 September 2025

Gong, S. L., Barrie, L. A., and Blanchet, J.-P.: Modeling sea-salt aerosols in the atmosphere: 1. Model development, Journal of Geophysical Research: Atmospheres, 102, 3805–3818, https://doi.org/10.1029/96JD02953, 1997.

Goto, D., Schutgens, N. a. J., Nakajima, T., and Takemura, T.: Sensitivity of aerosol to assumed optical properties over Asia using a global aerosol model and AERONET, Geophysical Research Letters, 38, https://doi.org/10.1029/2011GL048675, 2011.

Govardhan, G. R., Nanjundiah, R. S., Satheesh, S. K., Moorthy, K. K., and Takemura, T.: Inter-comparison and performance evaluation of chemistry transport models over Indian region, Atmospheric Environment, 125, 486–504, https://doi.org/10.1016/j.atmosenv.2015.10.065, 2016.

Granier, C., Darras, S., Denier van der Gon, H., Doubalova, J., Elguindi, N., Galle, B., Gauss, M., Guevara, M., Jalkanen, J.-P., Kuenen, J., Liousse, C., Quack, B., Simpson, D., and Sindelarova, K.: The Copernicus Atmosphere Monitoring Service global and regional emissions (April 2019 version), Copernicus Atmosphere Monitoring Service (CAMS), https://doi.org/10.24380/d0bn-kx16, 2019.

Grell, G. A. and Freitas, S. R.: A scale and aerosol aware stochastic convective parameterization for weather and air quality modeling, Atmospheric Chemistry and Physics, 14, 5233–5250, https://doi.org/10.5194/acp-14-5233-2014, 2014.

Grell, G. A., Peckham, S. E., Schmitz, R., McKeen, S. A., Frost, G., Skamarock, W. C., and Eder, B.: Fully coupled "online" chemistry within the WRF model, Atmospheric Environment, 39, 6957–6975, https://doi.org/10.1016/j.atmosenv.2005.04.027, 2005.

Guenther, A., Karl, T., Harley, P., Wiedinmyer, C., Palmer, P. I., and Geron, C.: Estimates of global terrestrial isoprene emissions using MEGAN (Model of Emissions of Gases and Aerosols from Nature), Atmospheric Chemistry and Physics, 6, 3181–3210, https://doi.org/10.5194/acp-6-3181-2006, 2006.

Gueymard, C. A. and Yang, D.: Worldwide validation of CAMS and MERRA-2 reanalysis aerosol optical depth products using 15 years of AERONET observations, Atmospheric Environment, 225, 117216, https://doi.org/10.1016/j.atmosenv.2019.117216, 2020.

Guo, J., Zhang, J., Yang, K., Liao, H., Zhang, S., Huang, K., Lv, Y., Shao, J., Yu, T., Tong, B., Li, J., Su, T., Yim, S. H. L., Stoffelen, A., Zhai, P., and Xu, X.: Investigation of near-global daytime boundary layer height using high-resolution radiosondes: first results and comparison with ERA5, MERRA-2, JRA-55, and NCEP-2 reanalyses, Atmospheric Chemistry and Physics, 21, 17079–17097, https://doi.org/10.5194/acp-21-17079-2021, 2021.

Gupta, H. V., Kling, H., Yilmaz, K. K., and Martinez, G. F.: Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling, Journal of Hydrology, 377, 80–91, https://doi.org/10.1016/j.jhydrol.2009.08.003, 2009.

Hahn, D. G. and Manabe, S.: The Role of Mountains in the South Asian Monsoon Circulation, Journal of the Atmospheric Sciences, 32, 1515–1541, https://doi.org/10.1175/1520-0469(1975)032<1515:TROMIT>2.0.CO;2, 1975.

Hall, D. K. and Riggs, G. A.: Accuracy assessment of the MODIS snow products, Hydrological Processes, 21, 1534–1547, https://doi.org/10.1002/hyp.6715, 2007.

Hasenkopf, C. A., Flasher, J. C., Veerman, O., Scalamogna, A., Silva, D., Salmon, M., Buuralda, D., and DeWitt, L. H.: Stories from OpenAQ, a Global and Grassroots Open Air Quality Community, AGU Fall Meeting Abstracts, ADS Bibcode: 2016AGUFMPA43A2162H, PA43A-2162, 2016.

https://doi.org/10.5194/essd-2025-275
Preprint. Discussion started: 11 September 2025

Hasson, S. ul, Gerlitz, L., Schickhoff, U., Scholten, T., and Böhner, J.: Recent Climate Change over High Asia, in: Climate Change, Glacier Response, and Vegetation Dynamics in the Himalaya: Contributions Toward Future Earth Initiatives, edited by: Singh, R., Schickhoff, U., and Mal, S., Springer International Publishing, Cham, 29–48, https://doi.org/10.1007/978-3-319-28977-9_2, 2016.

He, C.: Modelling light-absorbing particle–snow–radiation interactions and impacts on snow albedo: fundamentals, recent advances and future directions, Environ. Chem., 19, 296–311, https://doi.org/10.1071/EN22013, 2022.

He, C., Li, Q., Liou, K.-N., Takano, Y., Gu, Y., Qi, L., Mao, Y., and Leung, L. R.: Black carbon radiative forcing over the Tibetan Plateau, Geophysical Research Letters, 41, 7806–7813, https://doi.org/10.1002/2014GL062191, 2014.

He, C., Flanner, M. G., Chen, F., Barlage, M., Liou, K.-N., Kang, S., Ming, J., and Qian, Y.: Black carbon-induced snow albedo reduction over the Tibetan Plateau: uncertainties from snow grain shape and aerosol–snow mixing state based on an updated SNICAR model, Atmospheric Chemistry and Physics, 18, 11507–11527, https://doi.org/10.5194/acp-18-11507-2018, 2018.

Holben, B. N., Eck, T. F., Slutsker, I., Tanré, D., Buis, J. P., Setzer, A., Vermote, E., Reagan, J. A., Kaufman, Y. J., Nakajima, T., Lavenu, F., Jankowiak, I., and Smirnov, A.: AERONET—A Federated Instrument Network and Data Archive for Aerosol Characterization, Remote Sensing of Environment, 66, 1–16, https://doi.org/10.1016/S0034-4257(98)00031-5, 1998.

Hong, S.-Y.: A new stable boundary-layer mixing scheme and its impact on the simulated East Asian summer monsoon, Quarterly Journal of the Royal Meteorological Society, 136, 1481–1496, https://doi.org/10.1002/qj.665, 2010.

Hooghiemstra, P. B., Krol, M. C., Bergamaschi, P., de Laat, A. T. J., van der Werf, G. R., Novelli, P. C., Deeter, M. N., Aben, I., and Röckmann, T.: Comparing optimized CO emission estimates using MOPITT or NOAA surface network observations, Journal of Geophysical Research: Atmospheres, 117, https://doi.org/10.1029/2011JD017043, 2012.

Huang, H., Qian, Y., He, C., Bair, E. H., and Rittger, K.: Snow Albedo Feedbacks Enhance Snow Impurity-Induced Radiative Forcing in the Sierra Nevada, Geophysical Research Letters, 49, e2022GL098102, https://doi.org/10.1029/2022GL098102, 2022.

Huffman, G. J., Bolvin, D., Braithwaite, D., Hsu, K., Joyce, R., and Xie, P.: GPM IMERG Final Precipitation L3 1 day 0.1 degree x 0.1 degree V06, https://doi.org/10.5067/GPM/IMERGDF/DAY/06, 2014.

Hyland, R. W. and Wexler, A.: Formulations for the thermodynamic properties of the saturated phases of H2O from 173.15 K to 473.15 K, ASHRAE transactions, 89, 500–519, 1983.

Iacono, M. J., Delamere, J. S., Mlawer, E. J., Shephard, M. W., Clough, S. A., and Collins, W. D.: Radiative forcing by long-lived greenhouse gases: Calculations with the AER radiative transfer models, Journal of Geophysical Research: Atmospheres, 113, https://doi.org/10.1029/2008JD009944, 2008.

Immerzeel, W. W., Droogers, P., de Jong, S. M., and Bierkens, M. F. P.: Large-scale monitoring of snow cover and runoff simulation in Himalayan river basins using remote sensing, Remote Sensing of Environment, 113, 40–49, https://doi.org/10.1016/j.rse.2008.08.010, 2009.

Immerzeel, W. W., van Beek, L. P. H., and Bierkens, M. F. P.: Climate Change Will Affect the Asian Water Towers, Science, 328, 1382–1385, https://doi.org/10.1126/science.1183188, 2010.

Inness, A., Ades, M., Agustí-Panareda, A., Barré, J., Benedictow, A., Blechschmidt, A.-M., Dominguez, J. J., Engelen, R., Eskes, H., Flemming, J., Huijnen, V., Jones, L., Kipling, Z., Massart, S., Parrington, M., Peuch, V.-H., Razinger, M., Remy,

https://doi.org/10.5194/essd-2025-275
Preprint. Discussion started: 11 September 2025

S., Schulz, M., and Suttie, M.: The CAMS reanalysis of atmospheric composition, Atmospheric Chemistry and Physics, 19, 3515–3556, https://doi.org/10.5194/acp-19-3515-2019, 2019.

Jiang, Z., Jones, D. B. A., Worden, H. M., and Henze, D. K.: Sensitivity of top-down CO source estimates to the modeled vertical structure in atmospheric CO, Atmospheric Chemistry and Physics, 15, 1521–1537, https://doi.org/10.5194/acp-15-1521-2015, 2015.

Jiménez, P. A. and Dudhia, J.: Improving the Representation of Resolved and Unresolved Topographic Effects on Surface Wind in the WRF Model, Journal of Applied Meteorology and Climatology, 51, 300–316, https://doi.org/10.1175/JAMC-D-11-084.1, 2012.

Jo, D. S., Park, R. J., Lee, S., Kim, S.-W., and Zhang, X.: A global simulation of brown carbon: implications for photochemistry and direct radiative effect, Atmospheric Chemistry and Physics, 16, 3413–3432, https://doi.org/10.5194/acp-16-3413-2016, 2016.

Kang, S., Zhang, Y., Chen, P., Guo, J., Zhang, Q., Cong, Z., Kaspari, S., Tripathee, L., Gao, T., Niu, H., Zhong, X., Chen, X., Hu, Z., Li, X., Li, Y., Neupane, B., Yan, F., Rupakheti, D., Gul, C., Zhang, W., Wu, G., Yang, L., Wang, Z., and Li, C.: Black carbon and organic carbon dataset over the Third Pole, Earth System Science Data, 14, 683–707, https://doi.org/10.5194/essd-14-683-2022, 2022.

Kirschbaum, D., Kapnick, S. B., Stanley, T., and Pascale, S.: Changes in Extreme Precipitation and Landslides Over High Mountain Asia, Geophysical Research Letters, 47, e2019GL085347, https://doi.org/10.1029/2019GL085347, 2020.

Kumar, R., Naja, M., Pfister, G. G., Barth, M. C., and Brasseur, G. P.: Source attribution of carbon monoxide in India and surrounding regions during wintertime, Journal of Geophysical Research: Atmospheres, 118, 1981–1995, https://doi.org/10.1002/jgrd.50134, 2013.

Kumar, R., Barth, M. C., Pfister, G. G., Naja, M., and Brasseur, G. P.: WRF-Chem simulations of a typical pre-monsoon dust storm in northern India: influences on aerosol optical properties and radiation budget, Atmospheric Chemistry and Physics, 14, 2431–2446, https://doi.org/10.5194/acp-14-2431-2014, 2014.

Kumar, R., Barth, M. C., Nair, V. S., Pfister, G. G., Suresh Babu, S., Satheesh, S. K., Krishna Moorthy, K., Carmichael, G. R., Lu, Z., and Streets, D. G.: Sources of black carbon aerosols in South Asia and surrounding regions during the Integrated Campaign for Aerosols, Gases and Radiation Budget (ICARB), Atmospheric Chemistry and Physics, 15, 5415–5428, https://doi.org/10.5194/acp-15-5415-2015, 2015a.

Kumar, R., Barth, M. C., Pfister, G. G., Nair, V. S., Ghude, S. D., and Ojha, N.: What controls the seasonal cycle of black carbon aerosols in India?, Journal of Geophysical Research: Atmospheres, 120, 7788–7812, https://doi.org/10.1002/2015JD023298, 2015b.

Kumar, R., Delle Monache, L., Bresch, J., Saide, P. E., Tang, Y., Liu, Z., da Silva, A. M., Alessandrini, S., Pfister, G., Edwards, D., Lee, P., and Djalalova, I.: Toward Improving Short-Term Predictions of Fine Particulate Matter Over the United States Via Assimilation of Satellite Aerosol Optical Depth Retrievals, Journal of Geophysical Research: Atmospheres, 124, 2753–2773, https://doi.org/10.1029/2018JD029009, 2019.

Kumar, R., Ghude, S. D., Biswas, M., Jena, C., Alessandrini, S., Debnath, S., Kulkarni, S., Sperati, S., Soni, V. K., Nanjundiah, R. S., and Rajeevan, M.: Enhancing Accuracy of Air Quality and Temperature Forecasts During Paddy Crop Residue Burning Season in Delhi Via Chemical Data Assimilation, Journal of Geophysical Research: Atmospheres, 125, e2020JD033019, https://doi.org/10.1029/2020JD033019, 2020.

Kumar, R., He, C., Bhardwaj, P., Lacey, F., Buchholz, R. R., Brasseur, G. P., Joubert, W., Labuschagne, C., Kozlova, E., and Mkololo, T.: Assessment of regional carbon monoxide simulations over Africa and insights into source attribution and regional transport, Atmospheric Environment, 277, 119075, https://doi.org/10.1016/j.atmosenv.2022.119075, 2022.

Kumar, R., He, C., Roychoudhury, C., Cheng, W., Mizukami, N. and Arellano A.F.: High Mountain Asia 12 Km Modeled Estimates Of Aerosol Transport, Chemistry, And Deposition Reanalysis, 2003-2019, Version 1, 2024. https://doi.org/10.5067/CG4OT8DJX2Z7

Kumar, R., Bhardwaj, P., He, C., Boehnert, J., Lacey, F., Alessandrini, S., Sampson, K., Casali, M., Swerdlin, S., Wilhelmi, O., Pfister, G. G., Gaubert, B., and Worden, H.: A long-term high-resolution air quality reanalysis with a public-facing air quality dashboard over the Contiguous United States (CONUS), Earth System Science Data, 17, 1807–1834, https://doi.org/10.5194/essd-17-1807-2025, 2025.

Lawrence, D. M., Oleson, K. W., Flanner, M. G., Thornton, P. E., Swenson, S. C., Lawrence, P. J., Zeng, X., Yang, Z.-L., Levis, S., Sakaguchi, K., Bonan, G. B., and Slater, A. G.: Parameterization improvements and functional and structural advances in Version 4 of the Community Land Model, Journal of Advances in Modeling Earth Systems, 3, https://doi.org/10.1029/2011MS00045, 2011.

Lee, H.-J., Jo, H.-Y., Kim, J.-M., Bak, J., Park, M.-S., Kim, J.-K., Jo, Y.-J., and Kim, C.-H.: Nocturnal Boundary Layer Height Uncertainty in Particulate Matter Simulations during the KORUS-AQ Campaign, Remote Sensing, 15, 300, https://doi.org/10.3390/rs15020300, 2023.

Lee, J., Lee, E.-H., and Seol, K.-H.: Validation of Integrated MultisatellitE Retrievals for GPM (IMERG) by using gauge-based analysis products of daily precipitation over East Asia, Theor Appl Climatol, 137, 2497–2512, https://doi.org/10.1007/s00704-018-2749-1, 2019.

Leys, C., Ley, C., Klein, O., Bernard, P., and Licata, L.: Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median, Journal of Experimental Social Psychology, 49, 764–766, https://doi.org/10.1016/j.jesp.2013.03.013, 2013.

Li, C., Su, F., Yang, D., Tong, K., Meng, F., and Kan, B.: Spatiotemporal variation of snow cover over the Tibetan Plateau based on MODIS snow product, 2001–2014, International Journal of Climatology, 38, 708–728, https://doi.org/10.1002/joc.5204, 2018.

Liu, Y., Y. F., and Margulis., S. A.: High mountain asia UCLA daily snow reanalysis, version 1, , https://doi.org/10.5067/HNAUGJQXSCVU, 2021.

Liu, D., He, C., Schwarz, J. P., and Wang, X.: Lifecycle of light-absorbing carbonaceous aerosols in the atmosphere, npj Clim Atmos Sci, 3, 1–18, https://doi.org/10.1038/s41612-020-00145-8, 2020.

Liu, L., Cao, X., Li, S., and Jie, N.: A 31-year (1990–2020) global gridded population dataset generated by cluster analysis and statistical learning, Sci Data, 11, 124, https://doi.org/10.1038/s41597-024-02913-0, 2024.

Liu, S. and Liang, X.-Z.: Observed Diurnal Cycle Climatology of Planetary Boundary Layer Height, Journal of Climate, 23, 5790–5809, https://doi.org/10.1175/2010JCLI3552.1, 2010.

Liu, Z., Liu, Q., Lin, H.-C., Schwartz, C. S., Lee, Y.-H., and Wang, T.: Three-dimensional variational assimilation of MODIS aerosol optical depth: Implementation and application to a dust storm over East Asia, Journal of Geophysical Research: Atmospheres, 116, https://doi.org/10.1029/2011JD016159, 2011.

Lott, J. N.: The quality control of the Integrated Surface Hourly database, in: 14th conference on applied climatology, Citation Key: lott2004quality, 2004.

Lu, Z., Streets, D. G., Winijkul, E., Yan, F., Chen, Y., Bond, T. C., Feng, Y., Dubey, M. K., Liu, S., Pinto, J. P., and Carmichael, G. R.: Light Absorption Properties and Radiative Effects of Primary Organic Aerosol Emissions, Environ. Sci. Technol., 49, 4868–4877, https://doi.org/10.1021/acs.est.5b00211, 2015.

Madonna, F., Summa, D., Di Girolamo, P., Marra, F., Wang, Y., and Rosoldi, M.: Assessment of Trends and Uncertainties in the Atmospheric Boundary Layer Height Estimated Using Radiosounding Observations over Europe, Atmosphere, 12, 301, https://doi.org/10.3390/atmos12030301, 2021.

Madonna, F., Tramutola, E., Sy, S., Serva, F., Proto, M., Rosoldi, M., Gagliardi, S., Amato, F., Marra, F., Fassò, A., Gardiner, T., and Thorne, P. W.: The New Radiosounding HARMonization (RHARM) Data Set of Homogenized Radiosounding Temperature, Humidity, and Wind Profiles With Uncertainties, Journal of Geophysical Research: Atmospheres, 127, e2021JD035220, https://doi.org/10.1029/2021JD035220, 2022.

Maina, F. Z., Kumar, S. V., Dollan, I. J., and Maggioni, V.: Development and Evaluation of Ensemble Consensus Precipitation Estimates over High Mountain Asia, Journal of Hydrometeorology, 23, 1469–1486, https://doi.org/10.1175/JHM-D-21-0196.1, 2022a.

Maina, F. Z., Kumar, S. V., and Gangodagamage, C.: Irrigation and warming drive the decreases in surface albedo over High Mountain Asia, Scientific Reports, 12, 16163, https://doi.org/10.1038/s41598-022-20564-2, 2022b.

Maina, F. Z., Kumar, S. V., Albergel, C., and Mahanama, S. P.: Warming, increase in precipitation, and irrigation enhance greening in High Mountain Asia, Commun Earth Environ, 3, 1–8, https://doi.org/10.1038/s43247-022-00374-0, 2022c.

Maina, F. Z., Xue, Y., Kumar, S. V., Getirana, A., McLarty, S., Appana, R., Forman, B., Zaitchik, B., Loomis, B., Maggioni, V., and Zhou, Y.: Development of a multidecadal land reanalysis over High Mountain Asia, Sci Data, 11, 827, https://doi.org/10.1038/s41597-024-03643-z, 2024.

Malm, W. C. and Hand, J. L.: An examination of the physical and optical properties of aerosols collected in the IMPROVE program, Atmospheric Environment, 41, 3407–3427, https://doi.org/10.1016/j.atmosenv.2006.12.012, 2007.

Mass, C. and Ovens, D.: WRF model physics: problems, solutions and a new paradigm for progress, Preprints, 2010 WRF Users' Workshop, NCAR, Boulder, Colo., 2010.

Matsui, H.: Black carbon simulations using a size- and mixing-state-resolved three-dimensional model: 1. Radiative effects and their uncertainties, Journal of Geophysical Research: Atmospheres, 121, https://doi.org/10.1002/2015jd023998, 2016.

Meng, J., Gao, Y., Wang, Y., Sheng, L., and Zhang, S.: Evaluating the climate projections for China through the lens of the simulations of planetary boundary layer height, Atmospheric Research, 294, 106975, https://doi.org/10.1016/j.atmosres.2023.106975, 2023.

Mishra, S. K., Rupper, S., Kapnick, S., Casey, K., Chan, H. G., Ciraci', E., Haritashya, U., Hayse, J., Kargel, J. S., Kayastha, R. B., Krakauer, N. Y., Kumar, S. V., Lammers, R. B., Maggioni, V., Margulis, S. A., Olson, M., Osmanoglu, B., Qian, Y., McLarty, S., Rittger, K., Rounce, D. R., Shean, D., Velicogna, I., Veselka, T. D., and Arendt, A.: Grand Challenges of Hydrologic Modeling for Food-Energy-Water Nexus Security in High Mountain Asia, Front. Water, 3, https://doi.org/10.3389/frwa.2021.728156, 2021.

https://doi.org/10.5194/essd-2025-275
Preprint. Discussion started: 11 September 2025

Mogno, C., Palmer, P. I., Knote, C., Yao, F., and Wallington, T. J.: Seasonal distribution and drivers of surface fine particulate matter and organic aerosol over the Indo-Gangetic Plain, Atmospheric Chemistry and Physics, 21, 10881–10909, https://doi.org/10.5194/acp-21-10881-2021, 2021.

Morrison, H., Thompson, G., and Tatarskii, V.: Impact of Cloud Microphysics on the Development of Trailing Stratiform Precipitation in a Simulated Squall Line: Comparison of One- and Two-Moment Schemes, https://doi.org/10.1175/2008MWR2556.1, 2009.

Mues, A., Lauer, A., Lupascu, A., Rupakheti, M., Kuik, F., and Lawrence, M. G.: WRF and WRF-Chem v3.5.1 simulations of meteorology and black carbon concentrations in the Kathmandu Valley, Geoscientific Model Development, 11, 2067–2091, https://doi.org/10.5194/gmd-11-2067-2018, 2018.

Nakayama, T., Matsumi, Y., Sato, K., Imamura, T., Yamazaki, A., and Uchiyama, A.: Laboratory studies on optical properties of secondary organic aerosols generated during the photooxidation of toluene and the ozonolysis of α-pinene, Journal of Geophysical Research: Atmospheres, 115, https://doi.org/10.1029/2010JD014387, 2010.

NASA/LARC/SD/ASDC: MOPITT CO gridded monthly means (Near and Thermal Infrared Radiances) V008, 2000.

Neu, J. L. and Prather, M. J.: Toward a more physical representation of precipitation scavenging in global chemistry models, Atmospheric Chemistry and Physics, 12, 1791–1812, https://doi.org/10.5194/acp-12-1791-2012, 2012.

Neumann, D., Matthias, V., Bieser, J., Aulinger, A., and Quante, M.: A comparison of sea salt emission parameterizations in northwestern Europe using a chemistry transport model setup, Atmospheric Chemistry and Physics, 16, 9905–9933, https://doi.org/10.5194/acp-16-9905-2016, 2016.

Notarnicola, C.: Overall negative trends for snow cover extent and duration in global mountain regions over 1982–2020, Sci Rep, 12, 13731, https://doi.org/10.1038/s41598-022-16743-w, 2022.

Ojha, N., Pozzer, A., Rauthe-Schöch, A., Baker, A. K., Yoon, J., Brenninkmeijer, C. A. M., and Lelieveld, J.: Ozone and carbon monoxide over India during the summer monsoon: regional emissions and transport, Atmospheric Chemistry and Physics, 16, 3013–3032, https://doi.org/10.5194/acp-16-3013-2016, 2016.

Parrish, D. F. and Derber, J. C.: The National Meteorological Center's Spectral Statistical-Interpolation Analysis System, Monthly Weather Review, 120, 1747–1763, https://doi.org/10.1175/1520-0493(1992)120<1747:TNMCSS>2.0.CO;2, 1992.

Pepin, N. C., Arnone, E., Gobiet, A., Haslinger, K., Kotlarski, S., Notarnicola, C., Palazzi, E., Seibert, P., Serafin, S., Schöner, W., Terzago, S., Thornton, J. M., Vuille, M., and Adler, C.: Climate Changes and Their Elevational Patterns in the Mountains of the World, Reviews of Geophysics, 60, e2020RG000730, https://doi.org/10.1029/2020RG000730, 2022.

Pfeffer, W. T., Arendt, A. A., Bliss, A., Bolch, T., Cogley, J. G., Gardner, A. S., Hagen, J.-O., Hock, R., Kaser, G., Kienholz, C., Miles, E. S., Moholdt, G., Mölg, N., Paul, F., Radić, V., Rastner, P., Raup, B. H., Rich, J., Sharp, M. J., and Consortium, T. R.: The Randolph Glacier Inventory: a globally complete inventory of glaciers, Journal of Glaciology, 60, 537–552, https://doi.org/10.3189/2014JoG13J176, 2014.

Pool, S., Vis ,Marc, and and Seibert, J.: Evaluating model performance: towards a non-parametric variant of the Kling-Gupta efficiency, Hydrological Sciences Journal, 63, 1941–1953, https://doi.org/10.1080/02626667.2018.1552002, 2018.

Pu, Z., Xu, L., and Salomonson, V. V.: MODIS/Terra observed seasonal variations of snow cover over the Tibetan Plateau, Geophysical Research Letters, 34, https://doi.org/10.1029/2007GL029262, 2007.

Qian, Y., Yasunari, T. J., Doherty, S. J., Flanner, M. G., Lau, W. K. M., Ming, J., Wang, H., Wang, M., Warren, S. G., and Zhang, R.: Light-absorbing particles in snow and ice: Measurement and modeling of climatic and hydrological impact, Adv. Atmos. Sci., 32, 64–91, https://doi.org/10.1007/s00376-014-0010-0, 2015.

Ram, K., Sarin, M. M., and Tripathi, S. N.: A 1 year record of carbonaceous aerosols from an urban site in the Indo-Gangetic Plain: Characterization, sources, and temporal variability, Journal of Geophysical Research: Atmospheres, 115, https://doi.org/10.1029/2010JD014188, 2010.

Ramachandran, S., Rupakheti, M., Cherian, R., and Lawrence, M. G.: Aerosols heat up the Himalayan climate, Science of The Total Environment, 894, 164733, https://doi.org/10.1016/j.scitotenv.2023.164733, 2023.

Remer, L. A., Kaufman, Y. J., Tanré, D., Mattoo, S., Chu, D. A., Martins, J. V., Li, R.-R., Ichoku, C., Levy, R. C., Kleidman, R. G., Eck, T. F., Vermote, E., and Holben, B. N.: The MODIS Aerosol Algorithm, Products, and Validation, Journal of the Atmospheric Sciences, 62, 947–973, https://doi.org/10.1175/JAS3385.1, 2005.

Roychoudhury, C., He, C., Kumar, R., McKinnon, J. M., and Arellano Jr., A. F.: On the Relevance of Aerosols to Snow Cover Variability Over High Mountain Asia, Geophysical Research Letters, 49, e2022GL099317, https://doi.org/10.1029/2022GL099317, 2022.

Roychoudhury, C., He, C., Kumar, R., and Arellano Jr., A. F.: Diagnosing aerosol–meteorological interactions on snow within Earth system models: a proof-of-concept study over High Mountain Asia, Earth Syst. Dynam., 16, 1237–1266, https://doi.org/10.5194/esd-16-1237-2025, 2025.

Sadavarte, P. and Venkataraman, C.: Trends in multi-pollutant emissions from a technology-linked inventory for India: I. Industry and transport sectors, Atmospheric Environment, 99, 353–364, https://doi.org/10.1016/j.atmosenv.2014.09.081, 2014.

Saide, P. E., Spak, S. N., Carmichael, G. R., Mena-Carrasco, M. A., Yang, Q., Howell, S., Leon, D. C., Snider, J. R., Bandy, A. R., Collett, J. L., Benedict, K. B., de Szoeke, S. P., Hawkins, L. N., Allen, G., Crawford, I., Crosier, J., and Springston, S. R.: Evaluating WRF-Chem aerosol indirect effects in Southeast Pacific marine stratocumulus during VOCALS-REx, Atmospheric Chemistry and Physics, 12, 3045–3064, https://doi.org/10.5194/acp-12-3045-2012, 2012.

Saide, P. E., Carmichael, G. R., Liu, Z., Schwartz, C. S., Lin, H. C., da Silva, A. M., and Hyer, E.: Aerosol optical depth assimilation for a size-resolved sectional model: impacts of observationally constrained, multi-wavelength and fine mode retrievals on regional scale analyses and forecasts, Atmospheric Chemistry and Physics, 13, 10425–10444, https://doi.org/10.5194/acp-13-10425-2013, 2013.

Saleh, R., Marks, M., Heo, J., Adams, P. J., Donahue, N. M., and Robinson, A. L.: Contribution of brown carbon and lensing to the direct radiative effect of carbonaceous aerosols from biomass and biofuel burning emissions, Journal of Geophysical Research: Atmospheres, 120, 10,285-10,296, https://doi.org/10.1002/2015JD023697, 2015.

Sarangi, C., Qian, Y., Rittger, K., Bormann, K. J., Liu, Y., Wang, H., Wan, H., Lin, G., and Painter, T. H.: Impact of light-absorbing particles on snow albedo darkening and associated radiative forcing over high-mountain Asia: high-resolution WRF-Chem modeling and new satellite observations, Atmospheric Chemistry and Physics, 19, 7105–7128, https://doi.org/10.5194/acp-19-7105-2019, 2019.

Schmale, J., Flanner, M., Kang, S., Sprenger, M., Zhang, Q., Guo, J., Li, Y., Schwikowski, M., and Farinotti, D.: Modulation of snow reflectance and snowmelt from Central Asian glaciers by anthropogenic black carbon, Sci Rep, 7, 40501, https://doi.org/10.1038/srep40501, 2017.

https://doi.org/10.5194/essd-2025-275
Preprint. Discussion started: 11 September 2025

Seaman, N. L., Stauffer, D. R., and Lario-Gibbs, A. M.: A Multiscale Four-Dimensional Data Assimilation System Applied in the San Joaquin Valley during SARMAP. Part I: Modeling Design and Basic Performance Characteristics, 1995.

Seidel, D. J., Ao, C. O., and Li, K.: Estimating climatological planetary boundary layer heights from radiosonde observations: Comparison of methods and uncertainty analysis, Journal of Geophysical Research: Atmospheres, 115, https://doi.org/10.1029/2009JD013680, 2010.

Sha, T., Ma, X., Wang, J., Tian, R., Zhao, J., Cao, F., and Zhang, Y.-L.: Improvement of inorganic aerosol component in PM2.5 by constraining aqueous-phase formation of sulfate in cloud with satellite retrievals: WRF-Chem simulations, Science of The Total Environment, 804, 150229, https://doi.org/10.1016/j.scitotenv.2021.150229, 2022.

Shin, H. H. and Dudhia, J.: Evaluation of PBL Parameterizations in WRF at Subkilometer Grid Spacings: Turbulence Statistics in the Dry Convective Boundary Layer, Monthly Weather Review, 144, 1161–1177, https://doi.org/10.1175/MWR-D-15-0208.1, 2016.

Singh, R., Singh, C., Ojha, S. P., Kumar, A. S., and Kumar, A. S. K.: Development of an improved aerosol product over the Indian subcontinent: Blending model, satellite, and ground-based estimates, Journal of Geophysical Research: Atmospheres, 122, 367–390, https://doi.org/10.1002/2016JD025335, 2017.

Sinyuk, A., Holben, B. N., Eck, T. F., Giles, D. M., Slutsker, I., Korkin, S., Schafer, J. S., Smirnov, A., Sorokin, M., and Lyapustin, A.: The AERONET Version 3 aerosol retrieval algorithm, associated uncertainties and comparisons to Version 2, Atmos. Meas. Tech., 13, 3375–3411, https://doi.org/10.5194/amt-13-3375-2020, 2020.

Skiles, S. M., Flanner, M., Cook, J. M., Dumont, M., and Painter, T. H.: Radiative forcing by light-absorbing particles in snow, Nature Clim Change, 8, 964–971, https://doi.org/10.1038/s41558-018-0296-5, 2018.

Smith, A., Lott, N., and Vose, R.: The Integrated Surface Database: Recent Developments and Partnerships, Bulletin of the American Meteorological Society, 92, 704–708, https://doi.org/10.1175/2011BAMS3015.1, 2011.

Soulie, A., Granier, C., Darras, S., Zilbermann, N., Doumbia, T., Guevara, M., Jalkanen, J.-P., Keita, S., Liousse, C., Crippa, M., Guizzardi, D., Hoesly, R., and Smith, S.: Global Anthropogenic Emissions (CAMS-GLOB-ANT) for the Copernicus Atmosphere Monitoring Service Simulations of Air Quality Forecasts and Reanalyses, ESSD – Atmosphere/Energy and Emissions, https://doi.org/10.5194/essd-2023-306, 2023.

Stewart, I. D., Oke, T. R., and Krayenhoff, E. S.: Evaluation of the 'Local Climate Zone' Scheme Using Temperature Observations and Model Simulations, International Journal of Climatology, 34, 1062–1080, https://doi.org/10.1002/joc.3746, 2013.

Sugimoto, S., Xue, Y., Sato, T., and Takahashi, H. G.: Influence of convective processes on weather research and forecasting model precipitation biases over East Asia, Clim Dyn, 62, 2859–2875, https://doi.org/10.1007/s00382-022-06587-5, 2024.

Tarek, M., Brissette, F., and Arsenault, R.: Uncertainty of gridded precipitation and temperature reference datasets in climate change impact studies, Hydrology and Earth System Sciences, 25, 3331–3350, https://doi.org/10.5194/hess-25-3331-2021, 2021.

Taylor, A. C., Beare, R. J., and Thomson, D. J.: Simulating Dispersion in the Evening-Transition Boundary Layer, Boundary-Layer Meteorol, 153, 389–407, https://doi.org/10.1007/s10546-014-9960-0, 2014.

https://doi.org/10.5194/essd-2025-275
Preprint. Discussion started: 11 September 2025

Tian, R., Ma, Y., Ma, W., Zhao, X., and Zha, D.: Longer Time-Scale Variability of Atmospheric Vertical Motion Over the Tibetan Plateau and North Pacific and the Climate in East Asia, Atmosphere, 12, 630, https://doi.org/10.3390/atmos12050630, 2021.

Tuccella, P., Pitari, G., Colaiuda, V., Raparelli, E., and Curci, G.: Present-day radiative effect from radiation-absorbing aerosols in snow, Atmospheric Chemistry and Physics, 21, 6875–6893, https://doi.org/10.5194/acp-21-6875-2021, 2021.

Velasco, E., Retama, A., and Stratoulias, D.: Ambient Air Quality Monitoring, in: Air Quality Management and Research in Southeast Asia, edited by: Velasco, E., Retama, A., and Stratoulias, D., Springer Nature Switzerland, Cham, 45–67, https://doi.org/10.1007/978-3-031-69088-4_5, 2024.

Wang, X., Heald, C. L., Ridley, D. A., Schwarz, J. P., Spackman, J. R., Perring, A. E., Coe, H., Liu, D., and Clarke, A. D.: Exploiting simultaneous observational constraints on mass and absorption to estimate the global direct radiative forcing of black carbon and brown carbon, Atmospheric Chemistry and Physics, 14, 10989–11010, https://doi.org/10.5194/acp-14-10989-2014, 2014.

Wang, X., Heald, C. L., Liu, J., Weber, R. J., Campuzano-Jost, P., Jimenez, J. L., Schwarz, J. P., and Perring, A. E.: Exploring the observational constraints on the simulation of brown carbon, Atmospheric Chemistry and Physics, 18, 635–653, https://doi.org/10.5194/acp-18-635-2018, 2018.

Wang, X., Tolksdorf, V., Otto, M., and Scherer, D.: WRF-based dynamical downscaling of ERA5 reanalysis data for High Mountain Asia: Towards a new version of the High Asia Refined analysis, International Journal of Climatology, 41, 743–762, https://doi.org/10.1002/joc.6686, 2021.

Werner, M., Kryza, M., Pagowski, M., and Guzikowski, J.: Assimilation of PM2.5 ground base observations to two chemical schemes in WRF-Chem – The results for the winter and summer period, Atmospheric Environment, 200, 178–189, https://doi.org/10.1016/j.atmosenv.2018.12.016, 2019.

Wesely, M. L.: Parameterization of surface resistances to gaseous dry deposition in regional-scale numerical models, Atmospheric Environment, 23, 1293–1304, https://doi.org/10.1016/0004-6981(89)90153-4, 1989.

Wiedinmyer, C., Kimura, Y., McDonald-Buller, E. C., Emmons, L. K., Buchholz, R. R., Tang, W., Seto, K., Joseph, M. B., Barsanti, K. C., Carlton, A. G., and Yokelson, R.: The Fire Inventory from NCAR version 2.5: an updated global fire emissions model for climate and chemistry applications, Geoscientific Model Development, 16, 3873–3891, https://doi.org/10.5194/gmd-16-3873-2023, 2023.

Wilcox, R. R.: Comparing dependent robust correlations, British Journal of Mathematical and Statistical Psychology, 69, 215–224, https://doi.org/10.1111/bmsp.12069, 2016.

Wu, H., Tang, X., Wang, Z., Wu, L., Lu, M., Wei, L., and Zhu, J.: Probabilistic Automatic Outlier Detection for Surface Air Quality Measurements from the China National Environmental Monitoring Network, Adv. Atmos. Sci., 35, 1522–1532, https://doi.org/10.1007/s00376-018-8067-9, 2018.

Xi, X.: Revisiting the Recent Dust Trends and Climate Drivers Using Horizontal Visibility and Present Weather Observations, Journal of Geophysical Research: Atmospheres, 126, e2021JD034687, https://doi.org/10.1029/2021JD034687, 2021.

Xian, P., Reid, J. S., Ades, M., Benedetti, A., Colarco, P. R., da Silva, A., Eck, T. F., Flemming, J., Hyer, E. J., Kipling, Z., Rémy, S., Sekiyama, T. T., Tanaka, T., Yumimoto, K., and Zhang, J.: Intercomparison of aerosol optical depths from four reanalyses and their multi-reanalysis consensus, Atmospheric Chemistry and Physics, 24, 6385–6411, https://doi.org/10.5194/acp-24-6385-2024, 2024.

Xie, W., Yi, S., Leng, C., Xia, D., Li, M., Zhong, Z., and Ye, J.: The evaluation of IMERG and ERA5-Land daily precipitation over China with considering the influence of gauge data bias, Sci Rep, 12, 8085, https://doi.org/10.1038/s41598-022-12307-0, 2022.

Xu, L., Peng, Y., Ram, K., Zhang, Y., Bao, M., and Wei, J.: Investigation of the Uncertainties of Simulated Optical Properties of Brown Carbon at Two Asian Sites Using a Modified Bulk Aerosol Optical Scheme of the Community Atmospheric Model Version 5.3, Journal of Geophysical Research: Atmospheres, 126, e2020JD033942, https://doi.org/10.1029/2020JD033942, 2021.

Yan, S., Chen, Y., Hou, Y., Liu, K., Li, X., Xing, Y., Wu, D., Cui, J., Zhou, Y., Pu, W., and Wang, X.: Which global reanalysis dataset has better representativeness in snow cover on the Tibetan Plateau?, The Cryosphere, 18, 4089–4109, https://doi.org/10.5194/tc-18-4089-2024, 2024.

Yao, T., Bolch, T., Chen, D., Gao, J., Immerzeel, W., Piao, S., Su, F., Thompson, L., Wada, Y., Wang, L., Wang, T., Wu, G., Xu, B., Yang, W., Zhang, G., and Zhao, P.: The imbalance of the Asian water tower, Nat Rev Earth Environ, 3, 618–632, https://doi.org/10.1038/s43017-022-00299-4, 2022.

Yasunari, T. J., Tan, Q., Lau, K.-M., Bonasoni, P., Marinoni, A., Laj, P., Ménégoz, M., Takemura, T., and Chin, M.: Estimated range of black carbon dry deposition and the related snow albedo reduction over Himalayan glaciers during dry pre-monsoon periods, Atmospheric Environment, 78, 259–267, https://doi.org/10.1016/j.atmosenv.2012.03.031, 2013.

Yin, G., Baik ,Jongjin, and and Park, J.: Comprehensive analysis of GEO-KOMPSAT-2A and FengYun satellite-based precipitation estimates across Northeast Asia, GIScience & Remote Sensing, 59, 782–800, https://doi.org/10.1080/15481603.2022.2067970, 2022.

Yu, L., Leng, G., Python, A., and Peng, J.: A Comprehensive Evaluation of Latest GPM IMERG V06 Early, Late and Final Precipitation Products across China, Remote Sensing, 13, 1208, https://doi.org/10.3390/rs13061208, 2021.

Zakoura, M. and Pandis, S. N.: Overprediction of aerosol nitrate by chemical transport models: The role of grid resolution, Atmospheric Environment, 187, 390–400, https://doi.org/10.1016/j.atmosenv.2018.05.066, 2018.

Zaveri, R. A., Easter, R. C., Fast, J. D., and Peters, L. K.: Model for Simulating Aerosol Interactions and Chemistry (MOSAIC), Journal of Geophysical Research: Atmospheres, 113, https://doi.org/10.1029/2007JD008782, 2008.

Zhang, A., Wang, Y., Zhang, Y., Weber, R. J., Song, Y., Ke, Z., and Zou, Y.: Modeling the global radiative effect of brown carbon: a potentially larger heating source in the tropical free troposphere than black carbon, Atmospheric Chemistry and Physics, 20, 1901–1920, https://doi.org/10.5194/acp-20-1901-2020, 2020.

Zhang, D. and Anthes, R. A.: A High-Resolution Model of the Planetary Boundary Layer—Sensitivity Tests and Comparisons with SESAME-79 Data, 1982.

Zhang, Y., Kang, S., Sprenger, M., Cong, Z., Gao, T., Li, C., Tao, S., Li, X., Zhong, X., Xu, M., Meng, W., Neupane, B., Qin, X., and Sillanpää, M.: Black carbon and mineral dust in snow cover on the Tibetan Plateau, The Cryosphere, 12, 413–431, https://doi.org/10.5194/tc-12-413-2018, 2018.

Zhang, Z., Mu, L., and Li, C.: Comparison of Planetary Boundary Layer Height Derived from Lidar in AD-Net and ECMWFs Reanalysis Data over East Asia, Atmosphere, 13, 1976, https://doi.org/10.3390/atmos13121976, 2022.

Zhao, C., Hu, Z., Qian, Y., Ruby Leung, L., Huang, J., Huang, M., Jin, J., Flanner, M. G., Zhang, R., Wang, H., Yan, H., Lu, Z., and Streets, D. G.: Simulating black carbon and dust and their radiative forcing in seasonal snow: a case study over North

https://doi.org/10.5194/essd-2025-275
Preprint. Discussion started: 11 September 2025

China with field campaign measurements, Atmospheric Chemistry and Physics, 14, 11475–11491, https://doi.org/10.5194/acp-14-11475-2014, 2014.

https://doi.org/10.5194/essd-2025-275
Preprint. Discussion started: 11 September 2025

Open Access Earth System Science Data Discussions

Figures

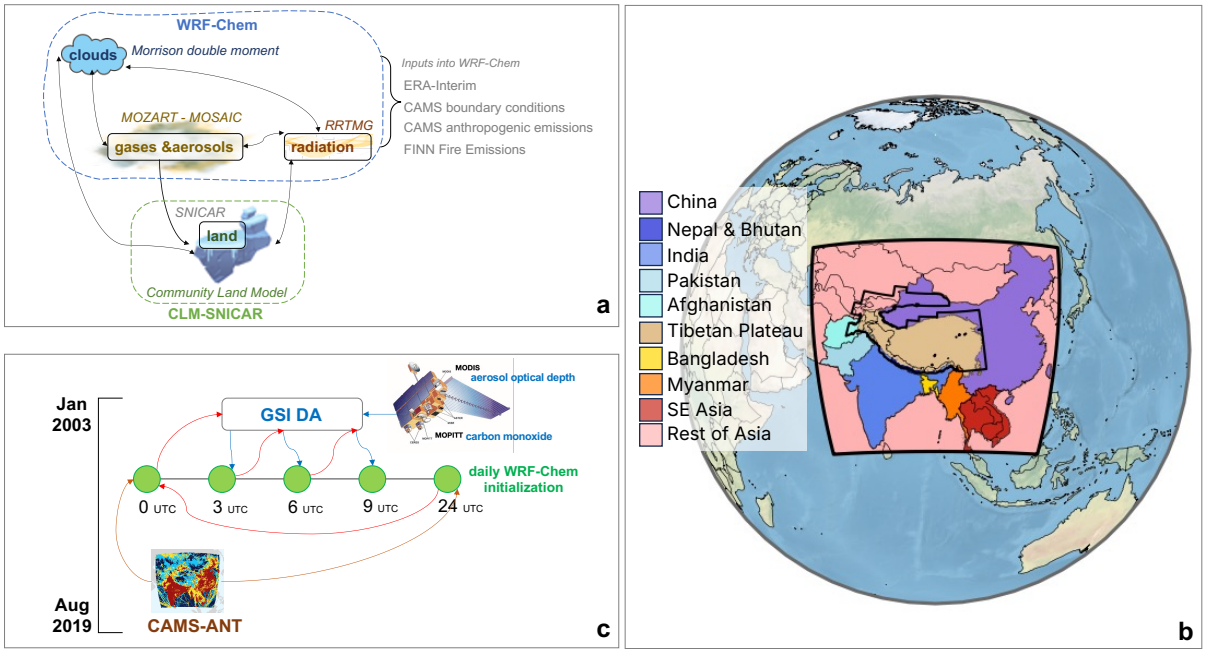


Figure 1: MATCHA framework. (a) schematic representation of the model setup in MATCHA. (b) the 10 tagged source regions for anthropogenic BC over the entire domain. (c) the assimilation workflow used in MATCHA, where MODIS AOD and MOPITT CO profiles were assimilated every 3 hours into the WRF-Chem-CLM-SNICAR model for the ~17 years of simulation period.

https://doi.org/10.5194/essd-2025-275
Preprint. Discussion started: 11 September 2025

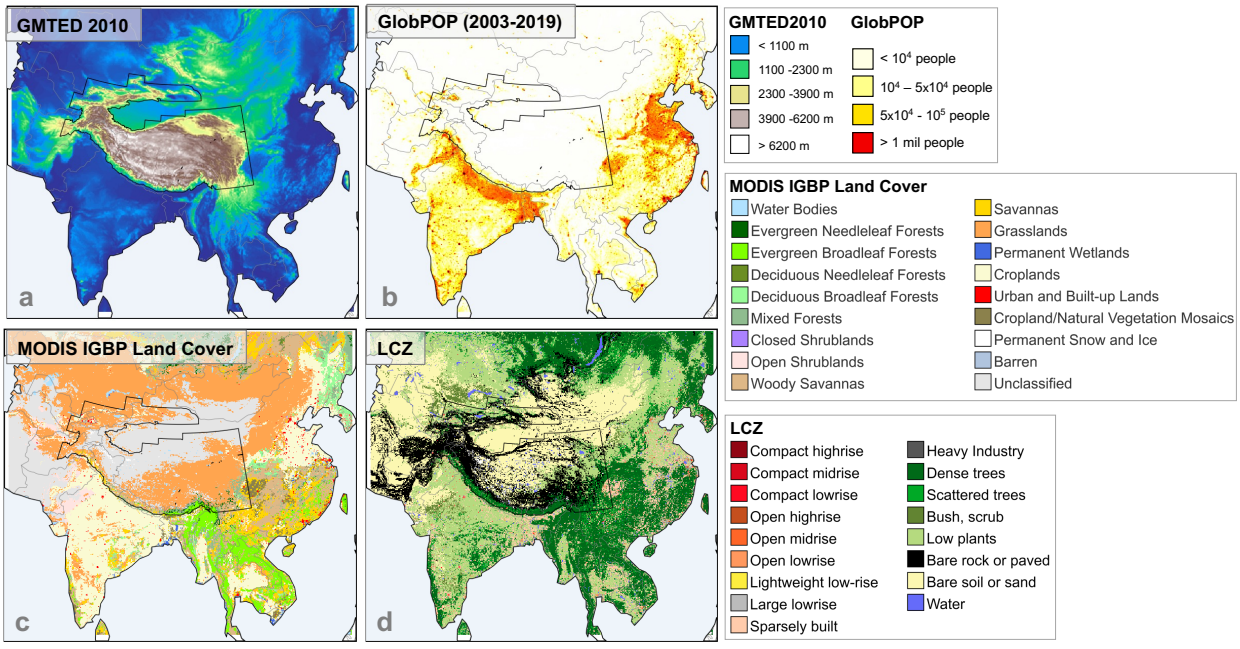


Figure 2: General characteristics over the model domain in MATCHA, spanning across topography (elevation from GMTED 2010) in (a), population count from the GlobPOP dataset (b), land cover from MODIS IGBP averaged from 2003-2019 (c), and local climate zones (see Sect. 2) mapped over the domain (d).

https://doi.org/10.5194/essd-2025-275
Preprint. Discussion started: 11 September 2025

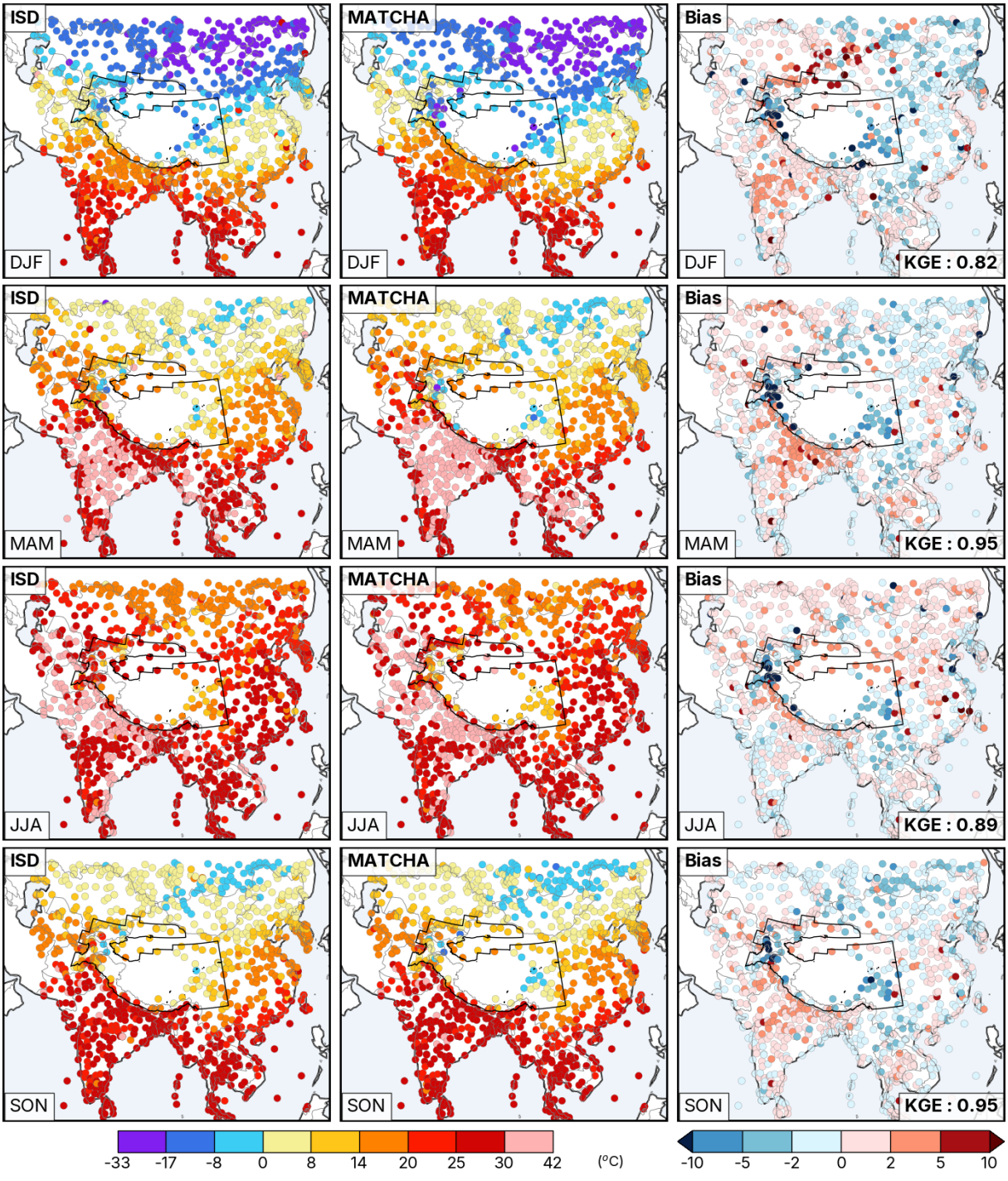


Figure 3: Seasonal averages of temperature at 2 m (in ºC) from the ISD archive (left column), MATCHA (middle column), and the associated model bias (MATCHA - ISD, right column).

505

https://doi.org/10.5194/essd-2025-275
Preprint. Discussion started: 11 September 2025

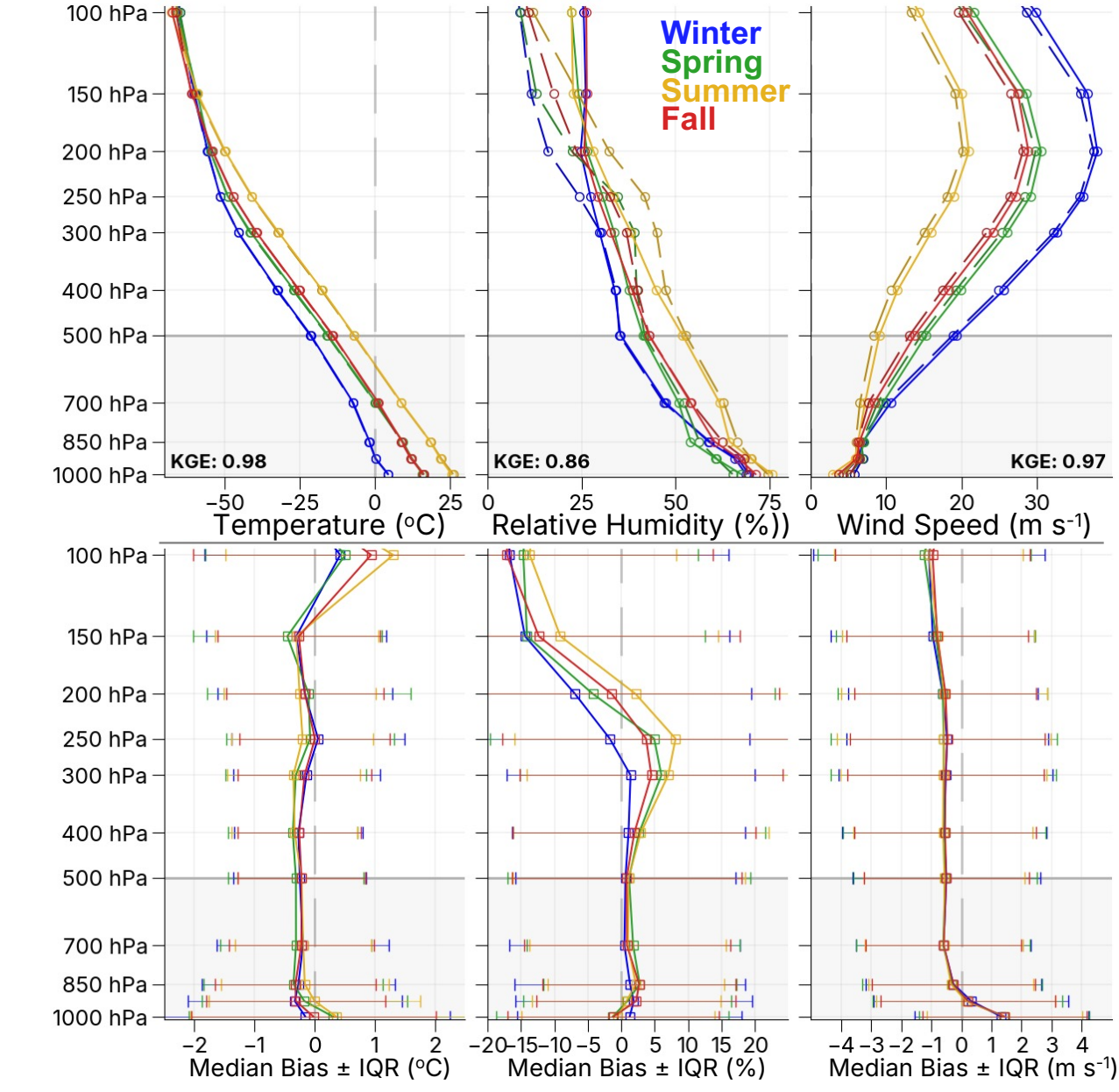


Figure 4: Seasonal averages of temperature, relative humidity, and wind speed profiles across the domain (top panel) from the IGRA radiosonde archive, and their relative biases (MATCHA-IGRA) in the bottom panel. The dashed lines in the top panel refer to profiles from MATCHA, while the solid lines refer to profiles from IGRA. Horizontal error bars in the bottom panel correspond to the interquartile range of the relative bias across all sites.

https://doi.org/10.5194/essd-2025-275
Preprint. Discussion started: 11 September 2025

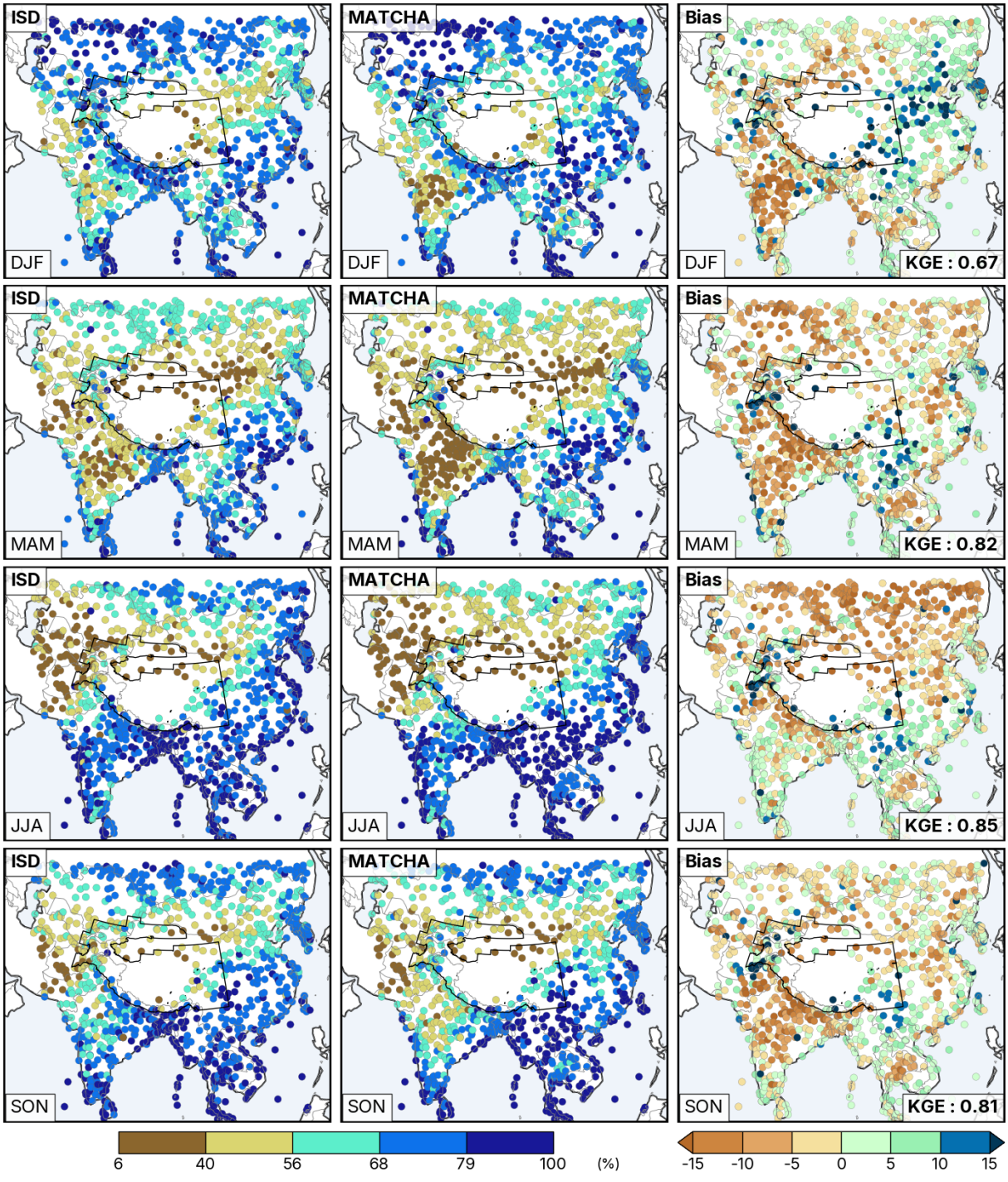


Figure 5: Seasonal averages of relative humidity at 2 m (in %) from the ISD archive (left column), MATCHA (middle column), and the associated model bias (MATCHA - ISD, right column).

https://doi.org/10.5194/essd-2025-275
Preprint. Discussion started: 11 September 2025

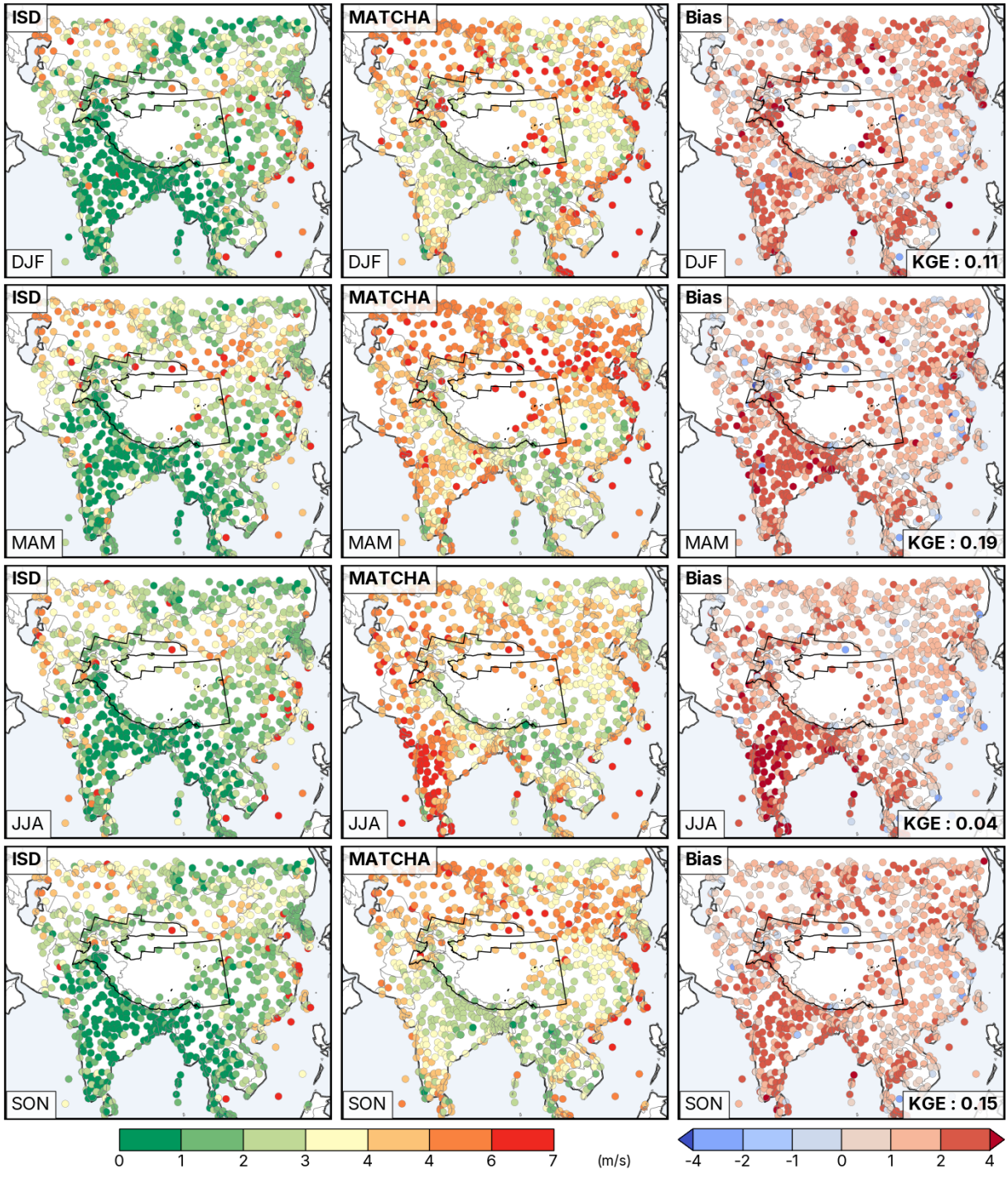


Figure 6: Seasonal averages of wind speed at 10 m (in m/s) from the ISD archive (left column), MATCHA (middle), and the associated model bias (MATCHA - ISD, right column).

520

https://doi.org/10.5194/essd-2025-275
Preprint. Discussion started: 11 September 2025

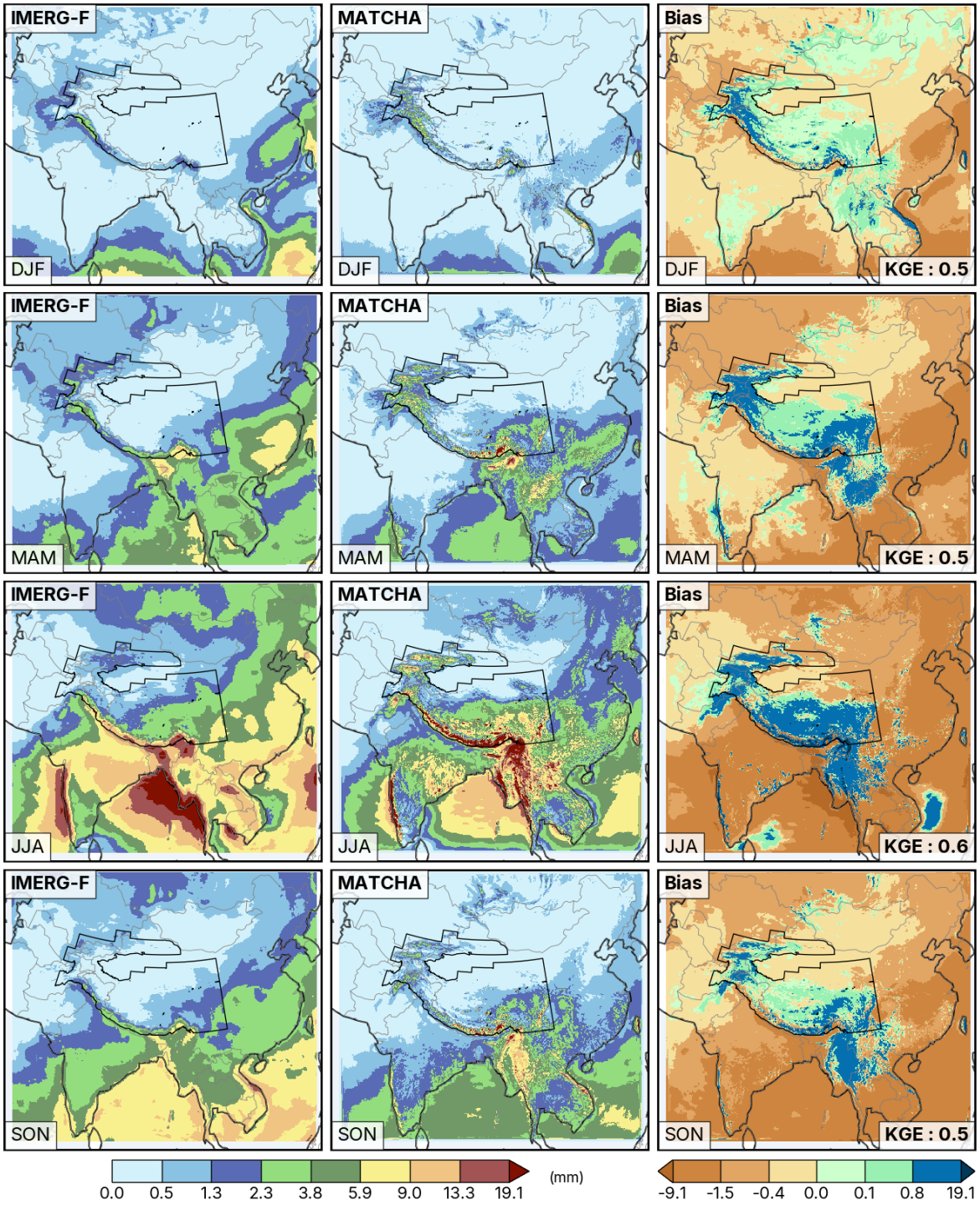


Figure 7: Seasonal averages of daily accumulated precipitation (in mm) across the domain from IMERG (left column), MATCHA (middle column), and the associated model bias (MATCHA-IMERG, right column).

https://doi.org/10.5194/essd-2025-275
Preprint. Discussion started: 11 September 2025

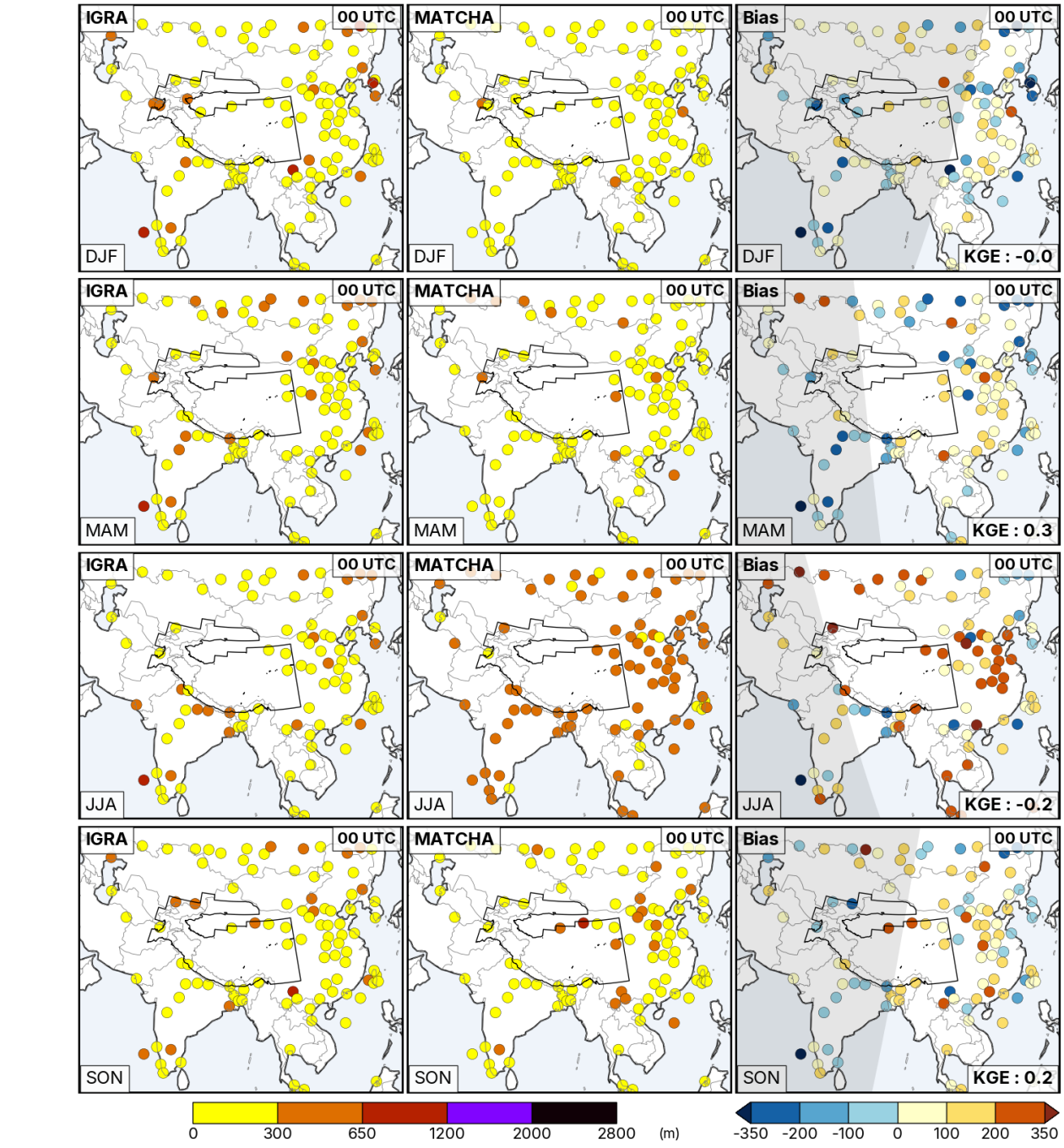


Figure 8: Seasonal averages of planetary boundary layer height at approximately 00 UTC from the IGRA archive (left column), MATCHA (middle column), and the associated model bias (MATCHA - IGRA, right column). The shadow zones show parts of the domain with nighttime conditions depending on the solar elevation angle on the 15th of the months; January (for DJF), April (for MAM), July (for JJA), and October (for SON), for the year 2003.

https://doi.org/10.5194/essd-2025-275
Preprint. Discussion started: 11 September 2025

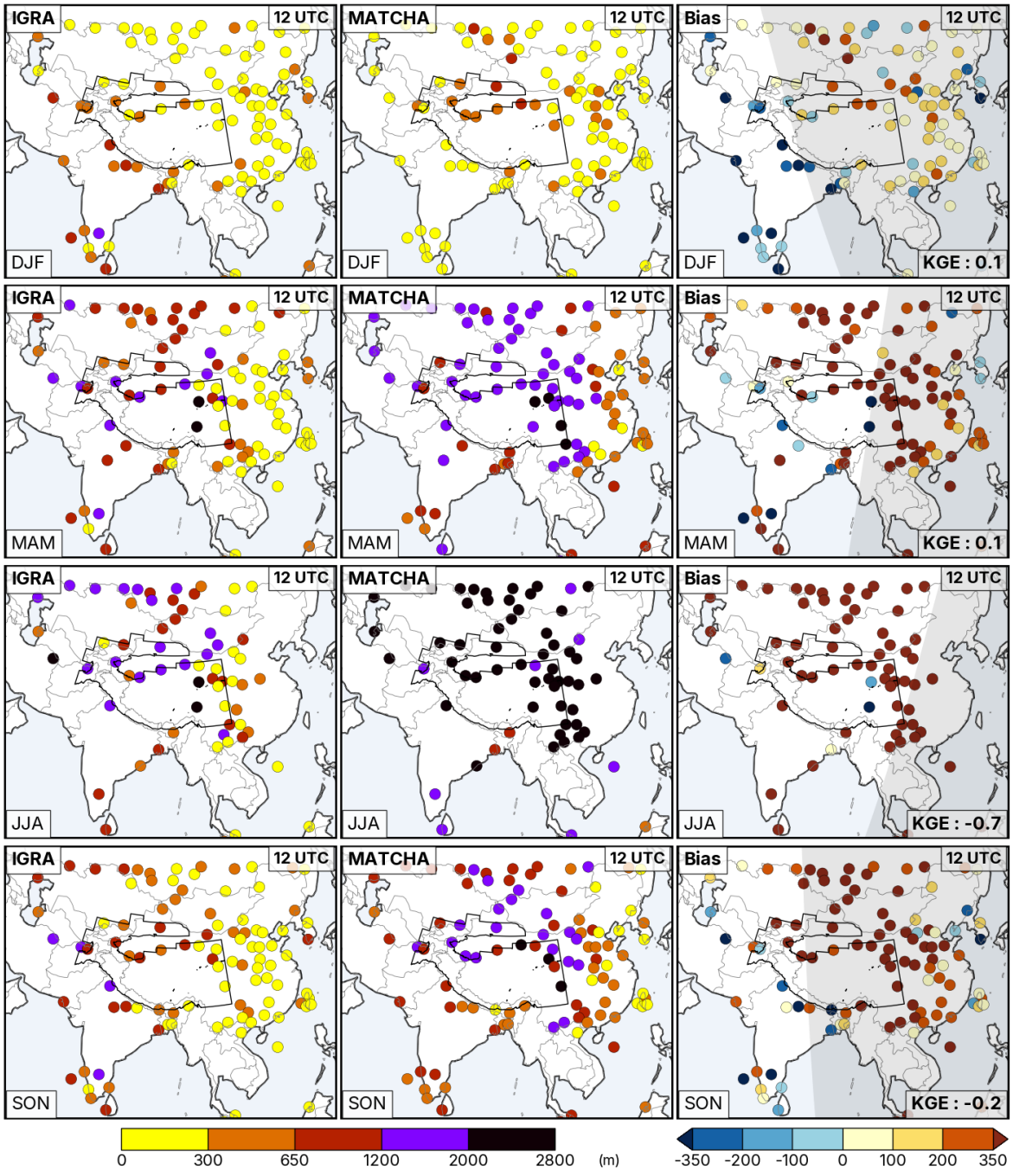


Figure 9: Seasonal mean of boundary layer height at approximately 12 UTC from the IGRA archive (left column), MATCHA (middle column), and the associated model bias (MATCHA - IGRA, right column). The shadow zones show parts of the domain with nighttime conditions depending on the solar elevation angle on the 15th of the months; January (for DJF), April (for MAM), July (for JJA), and October (for SON), for the year 2003.

https://doi.org/10.5194/essd-2025-275
Preprint. Discussion started: 11 September 2025

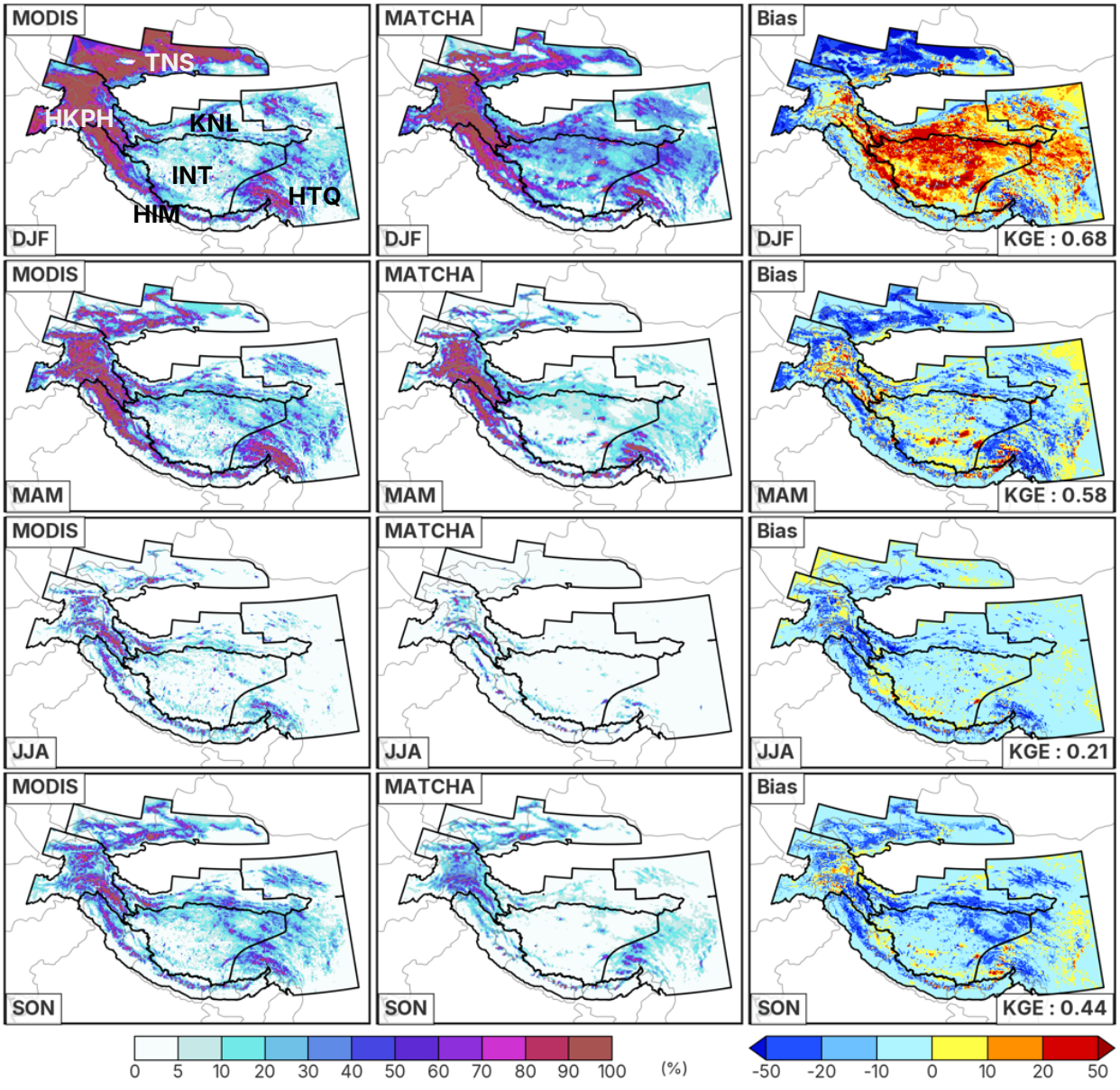


Figure 10: Seasonal averages of snow cover fraction (%) during 2003-2018 for the six aggregated first-order glacier regions of HMA from RGI v6 (INT: Inner Tibet, S and E Tibet; HTQ: Hengduan Shan, Qilian Shan; TNS: W and E Tien Shan; KNL: W and E Kun Lun; HIM: W, C, and E Himalayas, HKPH: Hindu Kush, Karakoram, Pamir, and Hissar Alay). The spatial averages are shown for MATCHA (left column), for MATCHA (middle column), and the associated model bias (MATCHA-MODIS, right column).

https://doi.org/10.5194/essd-2025-275
Preprint. Discussion started: 11 September 2025

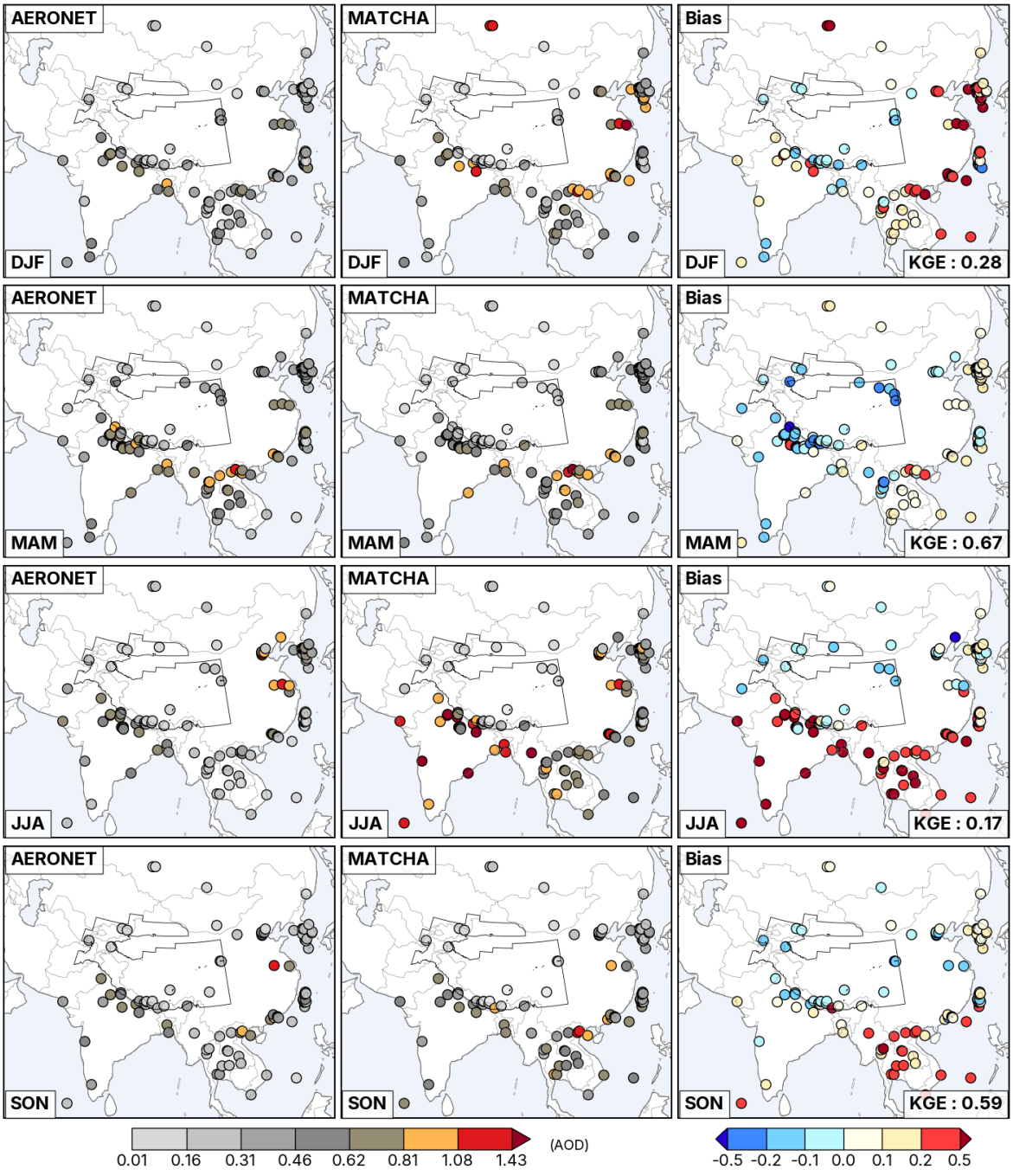


Figure 11: Seasonal averages of AOD at 550 nm (unitless) from AERONET sites (circles) across the domain (left column), MATCHA (middle), and the associated model bias (MATCHA-Observations, right column).

https://doi.org/10.5194/essd-2025-275
Preprint. Discussion started: 11 September 2025

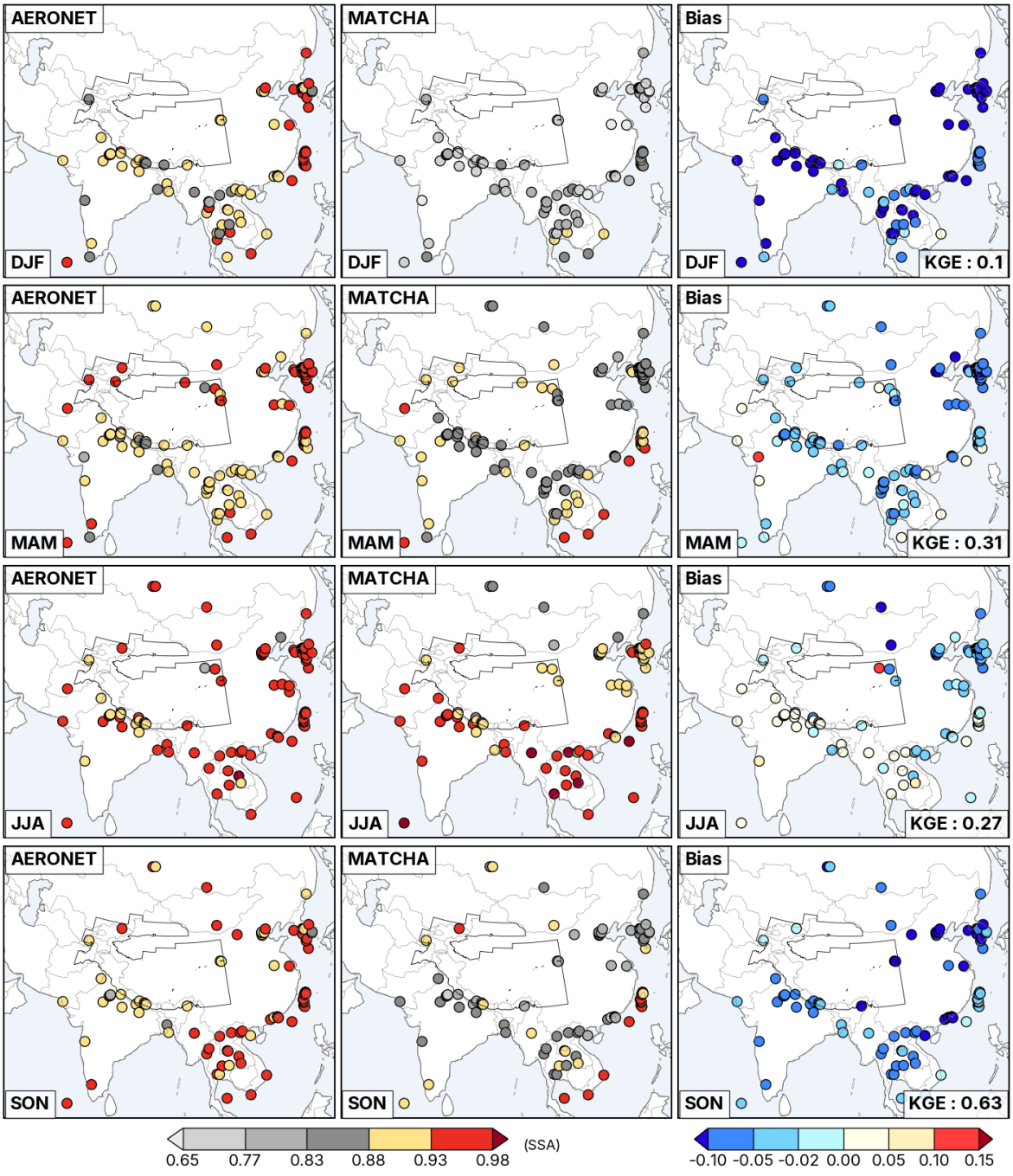


Figure 12: Seasonal averages of SSA at 550 nm (unitless) from AERONET sites across the domain (left column), from MATCHA (middle), and the associated model bias (MATCHA-AERONET, right column).

https://doi.org/10.5194/essd-2025-275
Preprint. Discussion started: 11 September 2025

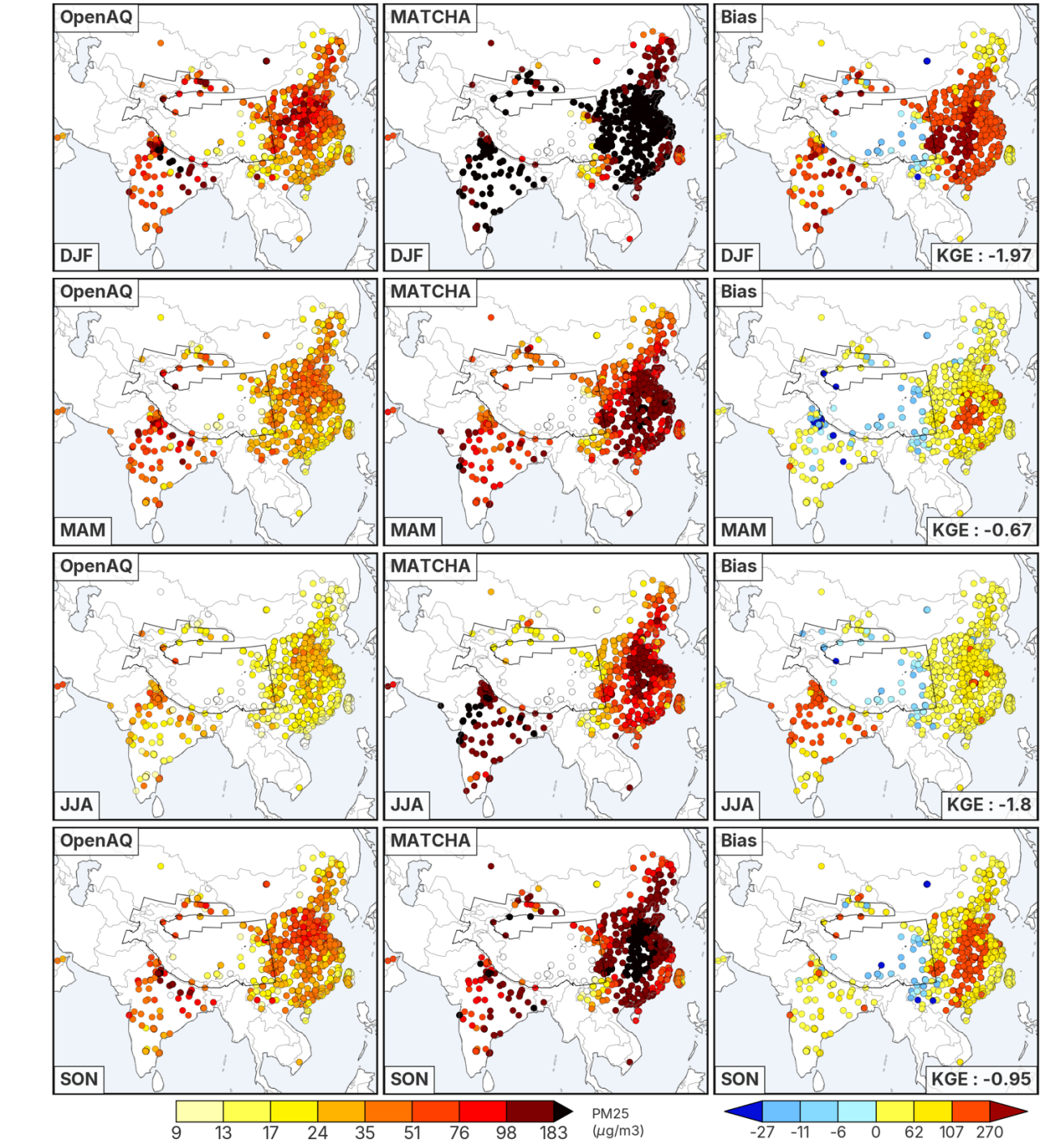


Figure 13: Seasonal averages of surface $PM_{2.5}$ (in µg/m^3) across the domain from OpenAQ (left column), MATCHA (middle), and the associated model bias (MATCHA-OpenAQ, right column).

https://doi.org/10.5194/essd-2025-275
Preprint. Discussion started: 11 September 2025

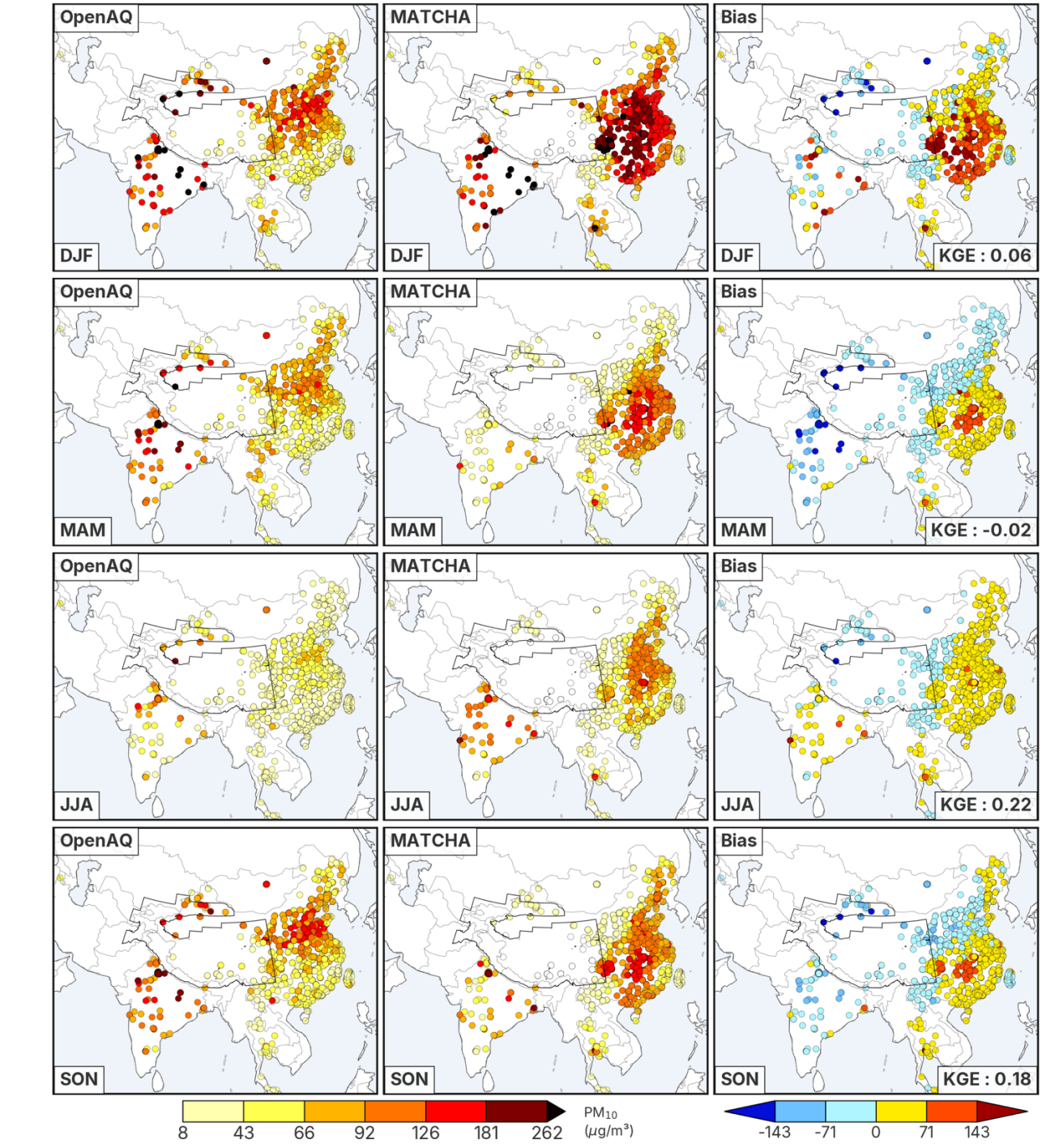


Figure 14: Seasonal averages of adjusted surface PM_{10} (after removing surface sea salt and dust) (in µg/m^3) across the domain from OpenAQ (left column), MATCHA (middle), and the associated model bias (MATCHA-OpenAQ, right column).

https://doi.org/10.5194/essd-2025-275
Preprint. Discussion started: 11 September 2025

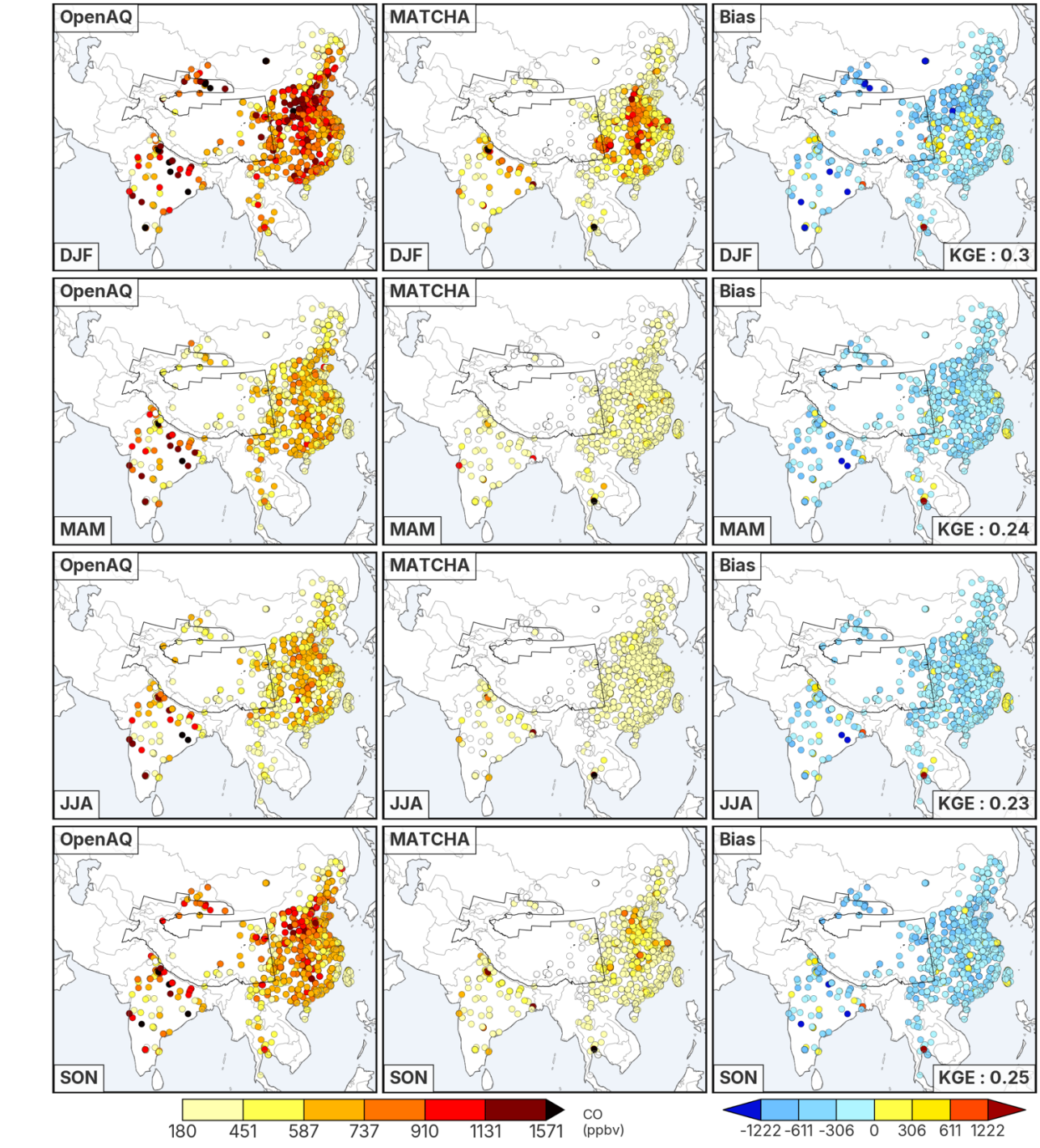
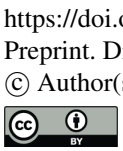


Figure 15: Seasonal mean of surface CO (in ppbv) across the domain from OpenAQ (left column), MATCHA (middle), and the associated model bias (MATCHA-OpenAQ, right column).

https://doi.org/10.5194/essd-2025-275
Preprint. Discussion started: 11 September 2025



Open Access Earth System Science Data Discussions

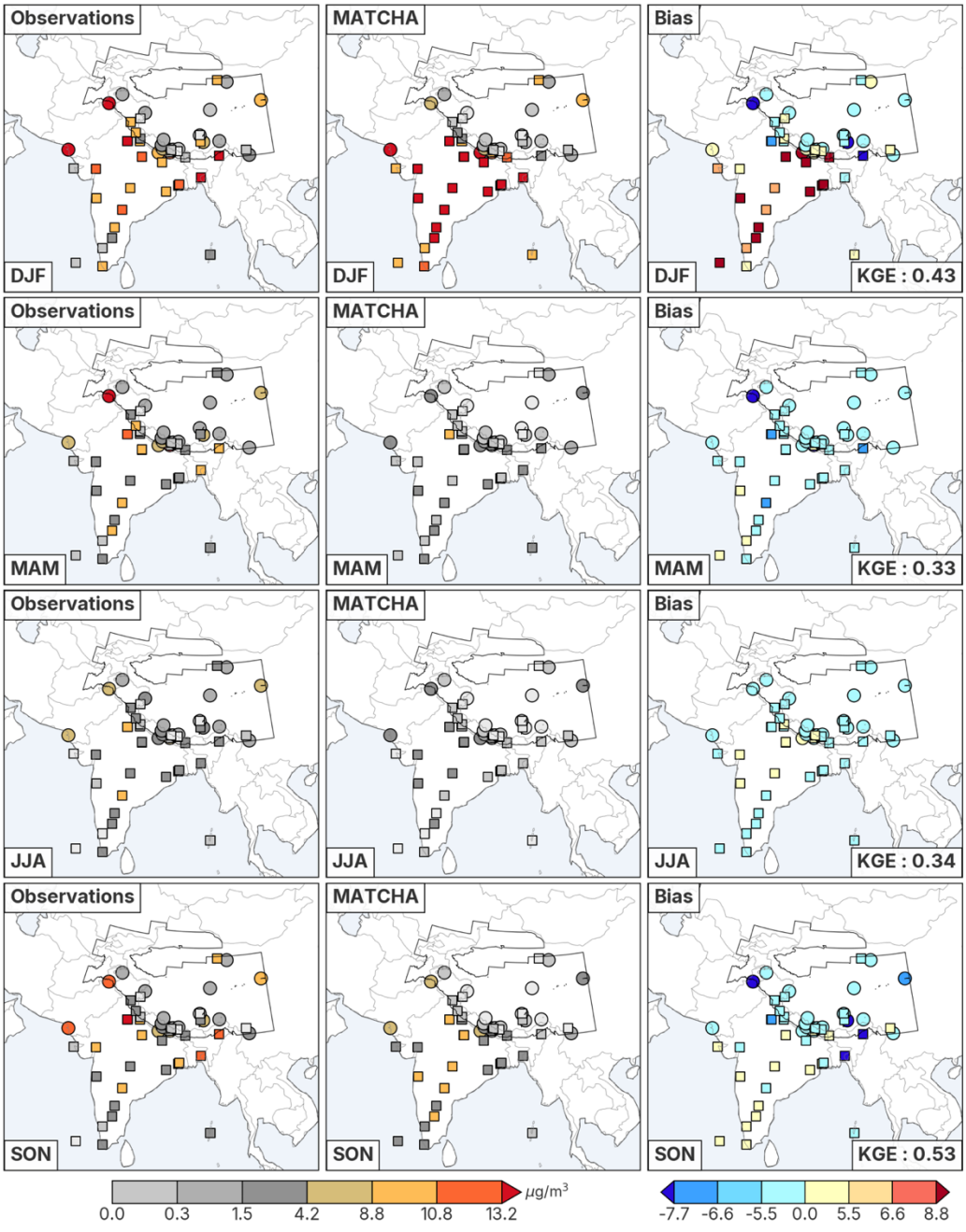


Figure 16: Seasonal mean of surface BC mass concentrations (in μg/m^3) across observation sites in the domain from ground-based observation sites (left column), MATCHA (middle), and the median model bias (MATCHA-sites, right column). Circle markers denote sites from the APCC network with daily observations, and square markers denote sites with monthly or seasonal measurements (see main text for details).

https://doi.org/10.5194/essd-2025-275
Preprint. Discussion started: 11 September 2025

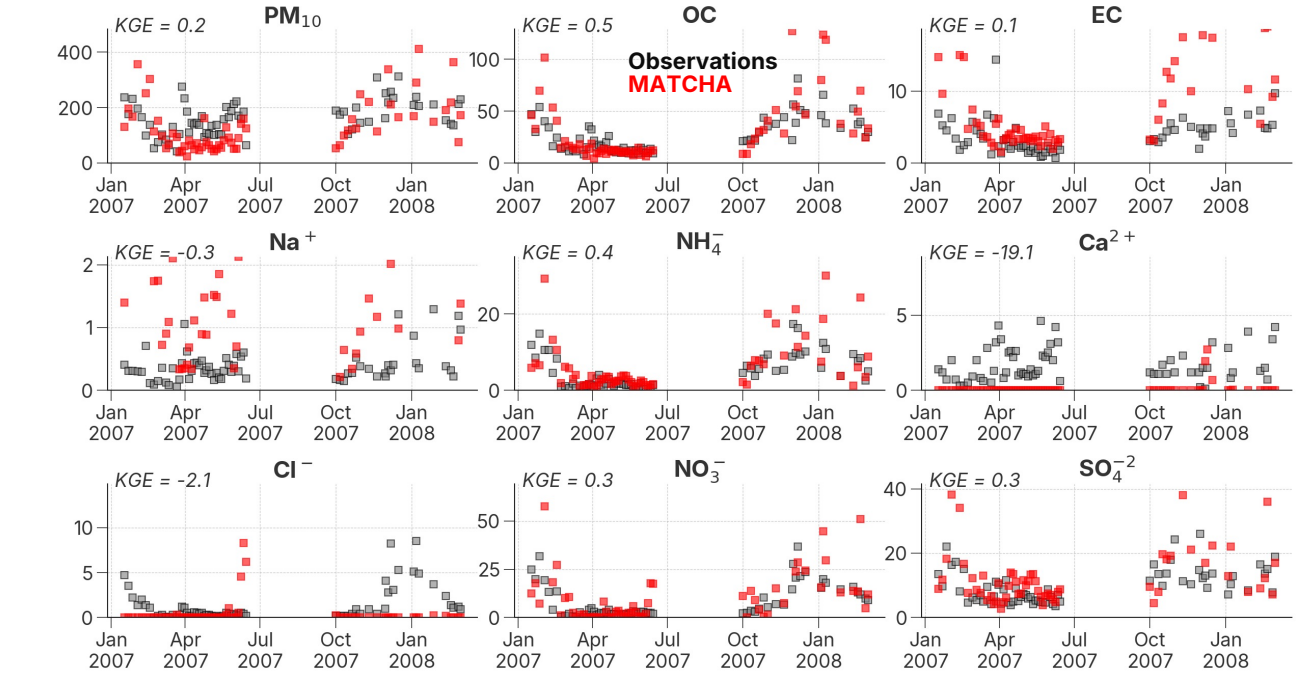


Figure 17: Observed (black) versus MATCHA-simulated (red) daily averaged (between 0600–1800 IST) concentrations (in μg/m^3) of total PM_{10} and eight chemical species at Kanpur (26.5° N, 80.3° E) from January 2007 to March 2008.

https://doi.org/10.5194/essd-2025-275
Preprint. Discussion started: 11 September 2025

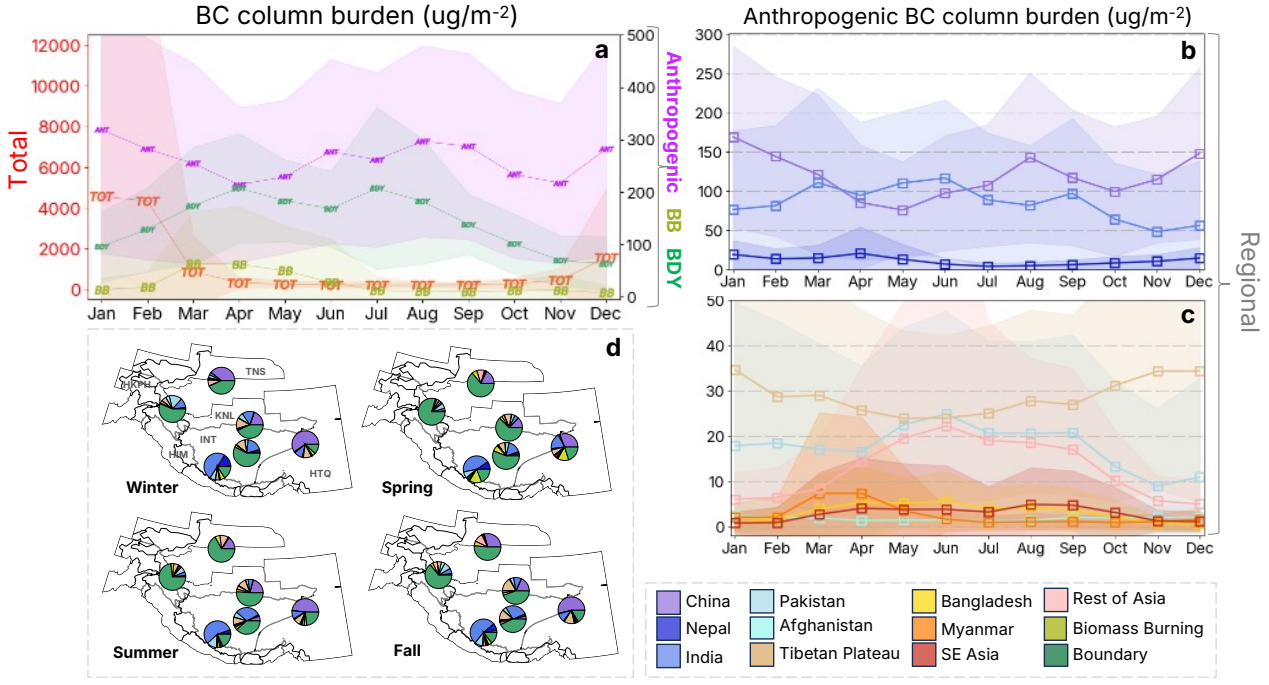


Figure 18: Average total column BC burden in µg/m^2 across sectors (a), across ten tagged regions of anthropogenic total column BC (b-c), and the seasonal contribution (%) of the 12 tags (d) (ten tagged regions + biomass burning + trans-boundary inflow) at six HMA first-order glacier regions from RGI v6 (INT: Inner Tibet, S and E Tibet; HTQ: Hengduan Shan, Qilian Shan; TNS: W and E Tien Shan; KNL: W and E Kun Lun; HIM: W, C, and E Himalayas, HKPH: Hindu Kush, Karakoram, Pamir, and Hissar Alay). The shaded regions for (a-c) refer to the average standard deviation across all six regions. The legend refers to the colors of the regional and the sectoral tags applicable for (b-d).

https://doi.org/10.5194/essd-2025-275
Preprint. Discussion started: 11 September 2025

Open Access Earth System Science Data Discussions

Tables

Table 1. Domain configuration and parameterizations used in the MATCHA model set-up.

Model Set-up	
Domain	High Mountain Asia and surrounding countries
Domain size	466 grid cells by 524 grid cells
Horizontal grid spacing	12 km by 12 km
Vertical grid	35 eta levels from surface up to 50 hPa
Simulation period	1st January 2003 – 31st August 2019
Temporal Resolution	Hourly and 3-hourly
Spin-up	1st December – 31st December of every year from 2002 to 2018
Meteorological IC/BC	ERA-Interim (Dee, 2011)
Chemical IC/BC	CAMS-EAC4 (Inness et al., 2019) CAM-Chem (Buchholz et al., 2019)
Physical Processes	
Cloud microphysics	Morrison double moment (Morrison et al., 2009)
Surface layer	MM5 similarity scheme (Zhang and Anthes, 1982)
Land surface model	Community Land Model v4 (Lawrence et al., 2011)
Planetary Boundary layer height	Yonsei University Scheme (Hong, 2010)
Cumulus parameterization	Grell-Freitas (Grell and Freitas, 2014)
Radiation (both short and long-wave)	RRTMG (Iacono et al., 2008)
Chemical Processes	
Gas-phase chemistry	MOZART-4 (Emmons et al., 2010)
Photolysis	Troposphere Ultraviolet Visible (TUV)
Dry deposition	Wesely, 1989
Wet deposition	Neu and Prather, 2012 for gases; Easter et al., 2004 for aerosols
Aerosol model	MOSAIC (Zaveri et al., 2008)
Aerosol optical properties	Mie-theory parameterization (Fast et al., 2006)
Biogenic emissions	MEGAN (Guenther et al., 2006)
Sea-salt emissions	Gong, 2003; Gong et al., 1997
Snow radiation model	SNICAR (Flanner et al., 2012; Zhao et al., 2014)
Anthropogenic emissions	CAMS-ANT-GLOBv4.2 (Granier et al., 2019)
Plume rise of fire emissions	Freitas et al., 2007
Fire emissions	FINN v2.5 (Wiedinmyer et al., 2023)

590

https://doi.org/10.5194/essd-2025-275
Preprint. Discussion started: 11 September 2025

Table 2. Evaluation metrics for all variables related to meteorology from MATCHA evaluated in this study across four seasons.

Variable	Season	MB	RMSE	R	KGE
Temperature at 2 m (°C)	DJF	-0.38	4.03	0.97	0.82
	MAM	-0.19	3.57	0.95	0.95
	JJA	0.07	2.9	0.89	0.89
	SON	-0.28	2.51	0.97	0.95
Temperature profiles (°C)	DJF	-1.05	3.67	0.99	0.99
	MAM	-1.26	4.09	0.98	0.98
	JJA	-1.12	3.61	0.98	0.98
	SON	-0.95	3.49	0.99	0.98
Temperature (°C)	DJF	-0.12	1.02	1	1
(lower troposphere, > 500 hPa)	MAM	-0.13	0.81	1	1
	JJA	0.04	0.83	0.99	0.99
	SON	-0.09	2.16	0.99	0.99
Temperature (°C)	DJF	-1.4	4.27	0.97	0.97
(upper troposphere, < 500 hPa)	MAM	-1.69	4.78	0.96	0.96
	JJA	-1.56	4.21	0.95	0.95
	SON	-1.28	3.88	0.96	0.96
Relative humidity at 2 m (%)	DJF	0.97	10.07	0.67	0.67
	MAM	-2.40	10.25	0.82	0.82
	JJA	-2.18	10.24	0.85	0.85
	SON	-1.60	8.63	0.81	0.81
Relative humidity profiles (%)	DJF	-4.20	14.66	0.67	0.64
	MAM	-2.41	12.89	0.78	0.76
	JJA	-0.83	12.41	0.87	0.85
	SON	-3.14	13.36	0.78	0.75
Relative humidity (%)	DJF	0.10	6.72	0.94	0.94
(lower troposphere, > 500 hPa)	MAM	0.82	5.89	0.94	0.94
	JJA	1.08	5.97	0.94	0.94
	SON	0.63	6.04	0.94	0.94
Relative humidity (%)	DJF	-6.16	17.09	0.23	0.17

https://doi.org/10.5194/essd-2025-275
Preprint. Discussion started: 11 September 2025

(upper troposphere, < 500 hPa)	MAM	-3.89	15.04	0.47	0.42
	JJA	-1.71	14.44	0.70	0.66
	SON	-4.90	15.64	0.45	0.40
Wind speed at 10 m (m/s)	DJF	1.64	2.33	0.65	0.11
	MAM	1.64	2.04	0.59	0.19
	JJA	1.67	2.18	0.41	0.04
	SON	1.55	1.83	0.68	0.15
Wind speed profiles (m/s)	DJF	-0.1	2.87	0.97	0.97
	MAM	-0.31	2.41	0.96	0.96
	JJA	-0.41	1.75	0.96	0.95
	SON	-0.53	2.26	0.97	0.95
Wind speed profiles (m/s)	DJF	0.08	1.8	0.93	0.93
(lower troposphere, > 500 hPa)	MAM	-0.05	1.48	0.93	0.93
	JJA	-0.14	1.04	0.9	0.9
	SON	0.01	1.64	0.9	0.9
Wind speed profiles (m/s)	DJF	-0.18	3.24	0.96	0.96
(upper troposphere, < 500 hPa)	MAM	-0.43	2.72	0.96	0.95
	JJA	-0.53	1.98	0.96	0.95
	SON	-0.78	2.5	0.96	0.94

595

https://doi.org/10.5194/essd-2025-275
Preprint. Discussion started: 11 September 2025

Table 3. Evaluation metrics for all variables related to land-atmosphere interactions from MATCHA evaluated in this study across four seasons.

Variable	Season	MB	RMSE	R	KGE
PBLH (m) (00 UTC)	DJF	-13.57	167.04	0.01	-0.01
	MAM	-2.47	151.73	0.29	0.28
	JJA	122.97	209.92	0	-0.17
	SON	49.79	148.46	0.21	0.16
PBLH (m) (12 UTC)	DJF	5.83	259.11	0.08	0.06
	MAM	525.3	743.91	0.63	0.06
	JJA	1398.53	1552.58	0.44	-0.67
	SON	411.06	637.43	0.41	-0.17
Daily accumulated precipitation (mm)	DJF	-0.43	1.2	0.72	0.45
	MAM	-0.68	1.85	0.74	0.54
	JJA	-1.75	5.17	0.82	0.61
	SON	-1.29	2.54	0.84	0.54
SCF (%)	DJF	-2.07	23.67	0.7	0.68
	MAM	-8.89	18.23	0.86	0.58
	JJA	-5.21	12.2	0.74	0.21
	SON	-8.08	14.28	0.85	0.44

https://doi.org/10.5194/essd-2025-275
Preprint. Discussion started: 11 September 2025

Table 4. Evaluation metrics for all variables related to atmospheric composition from MATCHA evaluated in this study across four seasons.

Variable	Season	MB	RMSE	R	KGE
AOD550	DJF	0.18	0.4	0.53	0.28
	MAM	-0.01	0.18	0.67	0.67
	JJA	0.3	0.6	0.61	0.17
	SON	0.08	0.21	0.67	0.59
SSA550	DJF	-0.16	0.19	0.11	0.1
	MAM	-0.05	0.07	0.31	0.31
	JJA	-0.02	0.04	0.27	0.27
	SON	-0.07	0.08	0.64	0.63
Surface $PM_{2.5}$ (μg/m³)	DJF	187.64	211.47	0.57	-1.97
	MAM	58.66	76.17	0.25	-0.67
	JJA	63.74	77.43	0.62	-1.8
	SON	88.73	104.22	0.56	-0.95
Surface PM_{10} (μg/m³)	DJF	203.02	235.36	0.5	-1.28
	MAM	38.91	73.54	0.12	-0.03
	JJA	42.23	58.33	0.48	-0.11
	SON	73.29	98.74	0.47	-0.06
Surface PM_{10} (μg/m³) (surface sea salt and dust removed)	DJF	62.33	100.67	0.36	0.06
	MAM	9.48	62.62	-0.01	-0.02
	JJA	25.51	42.47	0.49	0.22
	SON	5.64	55.97	0.19	0.18
Surface CO (ppbv)	DJF	-221.79	624.26	0.35	0.3
	MAM	-204	466.89	0.33	0.24
	JJA	-196.77	462.4	0.33	0.23
	SON	-245.97	591.11	0.34	0.25
Surface BC (μg/m³)	DJF	2.68	7.63	0.75	0.43
	MAM	-2.13	3.45	0.74	0.33
	JJA	-0.96	1.68	0.59	0.34
	SON	-1.27	3.18	0.68	0.53