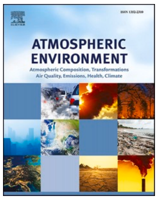
Atmospheric Environment 380 (2026) 122098

Contents lists available at ScienceDirect

Atmospheric Environment

journal homepage: www.elsevier.com/locate/atmosenv



Improving dust aerosol simulation over northern China: Synergy of updated numerical models and machine learning post-processing

Tong Sha[a,b,c,*], Liangqing Li[a,b,c], Zipeng Dong[c], Qingcai Chen[a], Shuqi Yan[d], Huanxin Zhang[e], Jhoon Kim[f], Jun Wang[e]

[a] Shaanxi University Key Laboratory of Industrial Pollution Control and Environmental Health, School of Environmental Science and Engineering, Shaanxi University of Science & Technology, Xi'an, 710021, China

[b] Shanghai Key Laboratory of Atmospheric Particle Pollution and Prevention (LAP^3), Shanghai, 200438, China

[c] China Meteorological Administration Eco-Environment and Meteorology for The Qinling Mountains and Loess Plateau Key Laboratory, Meteorological Institute of Shaanxi Province, Xi'an, 710014, China

[d] Key Laboratory of Transportation Meteorology of China Meteorological Administration, Nanjing Innovation Institute for Atmospheric Sciences, Nanjing, 210041, China

[e] Department of Chemical and Biochemical Engineering, Center for Global and Regional Environmental Research, and Iowa Technology Institute, University of Iowa, Iowa City, IA, 52242, USA

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

HIGHLIGHTS

- Dynamically updated EROD data optimizes dust source spatiotemporal heterogeneity.
- A hybrid framework synergizes numerical modeling with an ensemble ML post-processor.
- Ensemble ML post-processing enhances dust-related PM_{10} simulation accuracy.

ARTICLE INFO

Keywords:
Chemical transportation model
Machine learning
Erodibility
Dust
PM_{10}

ABSTRACT

Frequent dust events in China pose significant challenges to the accurate numerical simulation of dust aerosols. Here, we present a novel framework that synergizes an updated physical parameterization in the Unified Inputs for WRF-Chem (UI-WRF-Chem) model with an ensemble machine learning (ML) post-processor, focusing on Shaanxi Province, a typical Northwest region often affected by dust storms. The key improvement involves a dynamically updated erodibility dataset derived from multi-source satellite observations, thereby better capturing the spatiotemporal heterogeneity of dust sources on a monthly scale. The numerical model provides 5-km resolution dust simulations, which are subsequently bias-corrected using an ensemble ML approach combining extreme gradient Boost (XGBoost), random forest (RF), and neural Net-enhanced LightGBM (NetGBM). All ML models are trained using ground-based dust concentrations calibrated against PM_{10} observations during spring (March-May) from 2018 to 2022, and evaluated with independent data in spring 2023 and PM_{10} as the assessment indicator. Results show that the ML post-processor provides more accurate PM_{10} simulations than the numerical model across the spring months, increasing correlation coefficient (R) by 50~52% and reducing Normalized Standard Deviation (NSD) and Centered Root Mean Square Error (CRMSE) by 58~78% and 48~66%, respectively. The post-processing performance is region-dependent, substantial in northern Shaanxi but limited in the south due to lower dust contribution and uncalibrated non-dust components. Additionally, the ensemble ML model mitigates overestimation of dust-related PM_{10} by the UI-WRF-Chem model during dust episodes, and better captures dust transport processes. This study demonstrates that a robust framework for synergizing numerical modeling with ML approach can substantially enhance dust simulation accuracy in regions influenced by complex dust dynamics.

---

* Corresponding author. Shaanxi University Key Laboratory of Industrial Pollution Control and Environmental Health, School of Environmental Science and Engineering, Shaanxi University of Science & Technology, Xi'an, 710021, China.
E-mail address: tong-sha@sust.edu.cn (T. Sha).

https://doi.org/10.1016/j.atmosenv.2026.122098
Received 12 December 2025; Received in revised form 26 April 2026; Accepted 15 May 2026

1. Introduction

Mineral dust constitutes a major component of atmospheric aerosols, contributing approximately 30% of the global aerosol burden and aerosol optical depth (AOD), and accounting for more than half of the tropospheric aerosol mass (Lee and Lee, 2022; Meng et al., 2022). Globally, dust events inject 200~5000 Tg yr^{-1} of mineral particles into the atmosphere, with Asian dust sources alone contributing ~800 Tg yr^{-1} (Zeng et al., 2020; Zhang et al., 2021a). Dust aerosols directly influence the radiative budget and climate by absorbing and scattering solar radiation. Moreover, by acting as cloud condensation nuclei and ice nuclei, they indirectly alter the microphysical properties of clouds, modulating their albedo and lifetime, and thereby affecting precipitation patterns (Banks et al., 2024; Huang et al., 2014; Li et al., 2024b; Zhao et al., 2022). During long-range transport, dust plumes frequently interact with other air pollutants; this interaction enhances downwind pollutant concentrations via heterogeneous reactions, and exacerbates regional air quality degradation (Fussell and Kelly, 2021; Li et al., 2022a; Mesbahzadeh et al., 2020; Zeng et al., 2020). Furthermore, the inhalation of pathogen-laden dust aerosols poses considerable public health risks (Lian et al., 2025; Lwin et al., 2023). Therefore, accurate simulation and forecasting of dust processes are essential for quantifying their climate impacts, improving air quality management, and protecting human health (Zhang et al., 2021b).

Chemical transport models (CTMs), which utilize meteorological data and complex mathematical algorithms to simulate emissions, transports and removal of dust aerosols, have become essential tools in dust research (Eltahan et al., 2018; Rizza et al., 2017). However, dust events are sporadic phenomena characterized by intricate formation mechanisms, and their simulation remains challenging due to the complex interactions between surface conditions, meteorological fields, and dust aerosol schemes (Darvishi Boloorani et al., 2021; Li et al., 2022b; Srivastava and Blond, 2022). Computational resource constraints often require CTMs to represent complex processes via parameterizations. Such simplified representations, along with incomplete mechanistic understanding, may introduce substantial discrepancies in dust simulations (Kang et al., 2011). Previous studies have developed various approaches to mitigate these discrepancies. These efforts collectively reflect four aspects for CTM improvements: (1) refining parameterization schemes to better represent key processes, such as dust emissions, depositions, chemical reactions, etc. (Lee and Lee, 2022; Yoon et al., 2024); (2) improving the accuracy of key input parameters, such as description of land surface conditions (Han et al., 2021; Ukhov et al., 2021); (3) applying multi-source data assimilation to optimize the dust emission source estimates and initial boundary conditions, including meteorology and dust concentrations (Mei et al., 2024); (4) employing a statistical post-processing model to correct dust simulations (Jin et al., 2019; Raman et al., 2021). Despite these efforts, significant challenges remain, and no single approach has fully resolved the existing simulation biases (Letcher and LeGrand, 2018; Mesbahzadeh et al., 2020).

Compared to traditional CTMs, machine learning (ML) methods benefit from powerful data processing capabilities and strong model generalizability, enabling rapid analysis of large datasets with minimal prior knowledge and high computational efficiency. These advantages have facilitated their widespread application in atmospheric environment research (Chen et al., 2018, 2020; Lee and Lee, 2022; Udristioiu et al., 2023; Weng et al., 2022). Many studies have employed ML models to facilitate the development of CTMs, and the applications include long-term and short-term forecasting (Kim et al., 2019; Ma et al., 2020; Nguyen et al., 2023); the creation of high-precision data products by integrating CTMs with multi-source observations for policymaking and health research (Chelhaoui et al., 2025; Tang et al., 2024); diagnosing model biases and elucidating underlying physicochemical mechanisms (Liu and Xing, 2023; Wang et al., 2024; Ye et al., 2022); and replacing computationally intensive modules to enhance simulation efficiency (Lyu et al., 2025; Xu et al., 2021). However, compared to $PM_{2.5}$ and O_3, which benefit from abundant observational data (Galán-Madruga et al., 2025; Li et al., 2023b; Shi et al., 2024; Sun et al., 2021; Xiao et al., 2022; Zhu et al., 2022), research on utilizing ML to improve dust aerosol simulation performance remains relatively limited. This research gap is further exacerbated by the challenges in distinguishing natural dust from anthropogenic dust.

To refine the accuracy of dust simulation and forecasting, previous studies have commonly utilized ML approaches to assimilate the model initial fields using the ground-based and satellite observations, or directly predict the occurrence of dust events based on observational data (e.g., AOD, historical ground-based air pollution measurements, and meteorological observations) (Jin et al., 2019; Lee et al., 2021, 2022). However, these studies generally adopt a single ML method, making them prone to overfitting and substantial errors. Yi et al. (2025) proposed a multi-model ensemble ML framework that integrates the outputs from several numerical models and fuses various ML algorithms, such as Extreme Gradient Boosting (XGBoost), Light Gradient Boosting Machine (LightGBM), Categorical Boosted Trees (CatBoost), Random Forest (RF), and Multi-Layer Perceptron (MLP). This approach significantly improved the accuracy of dust concentration prediction. Nevertheless, it is important to note that the surface dust concentration data used for model training and validation in China are essentially derived from PM_{10} and $PM_{2.5}$ observations from the China National Environmental Monitoring Center (CNEMC), calculated based on a specific ratio. Additionally, most ground-based air quality monitoring stations in China are concentrated in urban centers, whereas suburban, rural, and desert edges regions, which are critical dust sources or transport pathways, have very sparse coverage or large observational gaps. Consequently, although the ML models based on observations can achieve high-precision forecasting in areas covered by stations, they may not be effectively extrapolated to areas without observations, posing a challenge for generating high spatial resolution dust concentration products.

In this study, we develop a hybrid post-processing framework that combines an ensemble of ML methods with UI-WRF-Chem simulations to improve hourly, 5 km resolution dust simulations over Shaanxi Province, a downwind region frequently affected by sand and dust storms originating from the Gobi Desert and Inner Mongolia of China during spring (Text S1 provides background on dust activity in this region). Key land surface parameters associated with dust aerosols in the UI-WRF-Chem model are also optimized to enhance dust simulation performance; specifically, the soil erodibility (EROD) dataset is dynamically updated using multi-source satellite observations to better represent the spatiotemporal heterogeneity of dust sources. The numerical model enables explicit simulation and output of both natural and anthropogenic dust, and we mainly focus on correcting the deviation of natural dust. Model performance is then evaluated using observed PM_{10} concentrations during spring (March-May) 2023. The hybrid CTM-ML framework is expected to provide an accurate, computationally efficient, and operationally reliable tool for simulating abrupt dust events, thereby providing scientific support for regional dust management and public health protection.

2. Materials and methods

2.1. Datasets

2.1.1. Land surface and soil data

The MODIS yearly land cover type (LCT) product (MCD12Q1, collection 6.1) at 500 m resolution, the MODIS 8-day leaf area index (LAI) product (MCD15A2H, collection 6.1) at 1 km resolution, and the Fraction of Photosynthetically Active Radiation absorbed by vegetation (FPAR) from MODIS 8-day MCD15A2H product at 500 m resolution (https://search.earthdata.nasa.gov/, last access: 1 January 2025), are used to dynamically update the land surface properties including LCT, LAI and green vegetation fraction (GVF) in the UI-WRF-Chem model. These updates are performed in a temporally consistent manner through the WRF Preprocessing System (WPS). The updated annual LCT and monthly LAI and GVF are subsequently employed in the Noah Land Surface Model (LSM) to enhance the representation of land surface processes in the UI-WRF-Chem model.

Soil moisture and texture are key parameters regulating dust emissions. SM directly influences the cohesive forces between soil particles, consequently altering the threshold friction velocity required for particle entrainment (Fécan et al., 1998; Lee et al., 2024; Shao, 2024). Soil texture determines the particle size distribution, which also plays a critical role in dust emissions. Therefore, soil moisture and texture are updated to refine the EROD data in the model. Surface (0~5 cm) SM are from the Soil Moisture Active Passive (SMAP) satellite data at 3-hourly, 9 km resolution (https://nsidc.org/data/smap/data/, last access: 1 January 2025). Soil texture data are from a reprocessed version of the Harmonized World Soil Database v2.0 (HWSD2.0) (https://doi.org/10.4060/cc3823en, last access: 1 January 2025), developed collaboratively by the Food and Agriculture Organization (FAO) and the International Institute for Applied Systems Analysis (IIASA). This dataset provides clay, silt, and sand fractions at a 1 km spatial resolution across seven standard depth layers: 0~20, 20~40, 40~60, 60~80, 80~100, 100~150, and 150~200 cm (IIASA, 2023; Nachtergaele et al., 2008).

2.1.2. Satellite observations

The AOD at 550 nm is obtained from the MODIS MOD08_D3 Collection 6.1 Level-3 daily product, derived from the Terra satellite at a 1° × 1° resolution (https://ladsweb.modaps.eosdis.nasa.gov/). For the evaluation of dust simulation performance, the AOD_550_Dark_Target_Deep_Blue_Combined product was selected. This product incorporates the Deep Blue algorithm, which is particularly suitable for retrieving dust aerosols over bright surfaces, it thus can identify dust storms and mitigate retrieval uncertainties over desert and snowy areas in China (Hsu et al., 2004).

2.1.3. Meteorological data

Hourly meteorological data are from the European Centre for Medium-Range Weather Forecasts (ECMWF) Fifth Generation Global Climate Atmospheric Reanalysis Dataset (ERA5) at a spatial resolution of 0.1° × 0.1° (https://cds.climate.copernicus.eu/datasets/reanalysis-era5-land?tab=download, last access: 1 January 2025) (Hersbach et al., 2020; Hoffmann et al., 2019). The following variables are used to train the first ML model (XGBoost) that separates dust from non-dust components of PM_{10}: the 2-m air temperature (T2, unit: K), 2-m dew point temperature (Td, unit: K), the u and v components of wind speed at 10 m (WS10, unit: $m\ s^{-1}$), and total precipitation (Precip, unit: m). ERA5 reanalysis data were chosen over ground-based meteorological station observations for two main reasons. First, the distribution of the meteorological stations in Shaanxi Province does not fully align with the CNEMC air quality monitoring sites; some stations are located far from the PM_{10} observation sites, which would introduce spatial mismatches if directly paired. In contrast, the gridded ERA5 data allows for precise spatial collocation by matching each air quality monitoring site with its corresponding grid cell, ensuring superior consistency between meteorological drivers and air pollutant measurements. Second, ERA5 provides continuous, high-resolution, and comprehensive spatial coverage, which is essential for capturing the changes in the meteorological fields. Although the 0.1° resolution may be coarser than the spacing between certain adjacent stations, this potential limitation was explicitly considered during the interpretation of the ML results.

2.1.4. Air quality observations

Hourly ground-based observations of air pollutant concentrations, including $PM_{2.5}$, PM_{10}, NO_2, CO, and O_3, are from the China National Environmental Monitoring Center (CNEMC) (http://www.cnemc.cn/, last access: 1 January 2025) to train the ML model and evaluate the simulation performance.

2.2. UI-WRF-Chem experimental setup

The UI-WRF-Chem model, developed from the standard version of WRF-Chem 3.9.1 (Grell et al., 2005), was used in this study. Meteorological and chemical initial and boundary conditions are provided by the Modern-Era Retrospective analysis for Research and Applications, Version 2 (MERRA-2) dataset with a spatial resolution of 0.625° × 0.5°. MERRA-2 is generated using the Goddard Earth Observing System (GEOS) atmospheric model and data assimilation system, incorporating simultaneously assimilates both meteorological and aerosol observations (Gelaro et al., 2017). Specifically, chemical initial and boundary conditions derived from MERRA-2 include the concentrations of key species, i.e., sulfate, black carbon, organic carbon, dust aerosol, and SO_2. Soil properties initial and boundary conditions, namely soil moisture and temperature, are provided by the Global Land Data Assimilation System (GLDAS) dataset with a spatial resolution of 0.25° × 0.25° (Rodell et al., 2004). This dataset has demonstrated robust performance in land surface process simulations (Araki et al., 2023; Bi et al., 2016; Kim et al., 2024), making it suitable for dust emission simulations. Further details of meteorological and chemical composition data for UI-WRF-Chem can be found in our previous studies (Li et al., 2024a; Sha et al., 2024; Sha et al., 2025; Zhang et al., 2026).

Anthropogenic emissions are obtained from the Multi-resolution Emission Inventory model for Climate and air pollution research (MEIC: http://www.meicmodel.org/, last access: 1 January 2025) with a spatial resolution of 0.25° × 0.25° for the specific study year. Since MEIC inventory is only publicly available until 2020, simulations after 2020 still use MEIC in 2020. Biomass burning emissions are from the Fire INventory from NCAR (FINN, version 1.5, https://www.acom.ucar.edu/Data/fire/, last access: 1 January 2025). Biogenic VOCs emissions (BVOCs) are calculated using the Model of Emissions of Gases and Aerosols from Nature (MEGAN) version 2.1 (Guenther et al., 2012). Soil reactive nitrogen (NO_x and HONO) emissions are calculated using the Berkeley-Dalhousie-Iowa Soil NO Parameterization (BDISNP) (Sha et al., 2021, 2024).

We select the Regional Acid Deposition Model, Version 2 (RADM2) for gas-phase chemistry (Stockwell et al., 1990) coupled with the Modal Aerosol Dynamics Model for Europe (MADE) (Ackermann et al., 1998) and the Secondary ORGanic Aerosol Model (SORGAM) (Schell et al., 2001) for aerosol chemistry, including aqueous-phase reactions. The MADE/SORGAM aerosol module adopts a modal approach (Binkowski and Shankar, 1995), which incorporates three size modes: Aitken, accumulation, and coarse mode. Notably, the default MADE/SORGAM scheme does not explicitly simulate dust aerosols, instead implicitly blends them into other aerosol species. To explicitly output the dust proportion of PM_{10} concentrations, the dust species simulated by the GOCART dust emission with the Air Force Weather Agency (AFWA) modifications (LeGrand et al., 2019) are coupled with the MADE/SORGAM aerosol scheme. The GOCART-AFWA emission scheme is selected because it releases dust in five size bins with lower and upper radius ranges of 0.1~1, 1~1.8, 1.8~3, 3~6, 6~10 μm, consistent with the dust size distribution used in the MERRA-2. Dust density is assumed to be 2500 $kg\ m^{-3}$ for the first bin and 2650 $kg\ m^{-3}$ for bins 2~5 in the model (Marticorena and Bergametti, 1995; Su and Fung, 2015; Wang et al., 2015). To better characterize East Asian dust properties, the size-resolved dust emission fractions (*dust_emis_fact*) of the GOCART-AFWA scheme are adjusted according to Su and Fung (2015). The revised emission fractions for each dust size bin are 0.034, 0.187, 0.327, 0.163, and 0.309, respectively.

For the physical parameterization schemes, the unified Noah land surface model (Mukul Tewari et al., 2004) and the revised MM5 Monin–Obukhov scheme (Pahlow et al., 2001) are selected to represent land surface processes and surface layer physics. The Yonsei University scheme is selected for planetary boundary layer (PBL) parameterization

(Hong et al., 2006). The WRF single-moment Purdue Lin microphysics scheme is used for the treatment of cloud microphysics (Lin et al., 1983). The Grell 3-D cumulus scheme is used for cumulus parameterization (Grell and Freitas, 2013). The Rapid Radiative Transfer Model (RRTMG) for both shortwave and longwave radiation is used for radiative transfer calculations (Iacono et al., 2000; Mlawer et al., 1997). Two nested domains are used in this study, outer domain (d01) covers China with a horizontal resolution of 25 km and 194×148 grid cells, and the inner domain (d02) covers Shaanxi province and its surrounding area with a horizontal resolution of 5 km and 209×246 grid cells (study region shown in Fig. 1). Both domains have 47 vertical levels from the surface to 50 hPa and 4 levels of soils. Simulations are conducted separately for each month (March, April, and May) from 2018 to 2023. For each month, the model runs continuously for the entire month without re-initialization, with the first 48 h of each simulation treated as spin-up and excluded from analysis. The same configuration is applied independently to each month of each year.

2.3. *Update of soil erodibility data*

To accurately identify the dust emission sources, the soil EROD data are dynamically updated utilizing the reprocessed LCT and GVF data from the WPS, SM from the SMAP, and soil texture from the HWSD2.0. All datasets used are monthly mean values. First, soil moisture and texture datasets are resampled to match the simulation grid resolution. Bilinear interpolation and averaging are applied for upscaling and downscaling variables, respectively, to ensure consistent spatial resolution across all variables. A preliminary EROD map is generated using LCT data to classify erodible surfaces, i.e., dryland or desert and grasslands are assigned a value of 1 (erodible), whereas all other areas are assigned a value of 0. This process yields initial EROD data for bare soil surfaces. Detailed MODIS LCT classifications are shown in Table S1. Next, GVF is used to mask vegetated areas, and the preliminary EROD data are multiplied by (1-GVF) to quantify vegetation-induced reductions in soil erodibility at the pixel scale. Following vegetation adjustment, potential dust source areas are identified using the sand fraction (SF). Subsequently, areas with SM (volumetric moisture content) below 0.1 $m^3\ m^{-3}$ are delineated to indicate surfaces susceptible to wind erosion. The monthly dynamic EROD data are finally calculated using Equation (1), followed by the study of Li et al., 2023a with adjustments to input datasets and spatial processing methods to better align with the specific objectives of this study. Fig. S1 illustrates the spatiotemporal variations in soil EROD across China in 2023 relative to 2018, driven by changes in LCT, GVF, and SM. Specifically, the EROD variations range from −1.91%~3.74% in March, −3.82%~4.85% in April, and −2.25%~3.74% in May (Fig. S1).

$$EROD = LCT \cdot (1\text{- } GVF) \cdot SF \cdot \exp[\text{-}\min(0.1, SM)] \tag{1}$$

2.4. *Ensemble ML model*

The ML framework comprises two modules (Fig. 2). The first module aims to separate dust concentrations (PM_{10}_dust) from the observed PM_{10}. During the dust events, PM_{10} concentrations often increase sharply, whereas $PM_{2.5}$ levels generally remain stable and unchanged (Jin et al., 2018; Lin et al., 2008; Xu et al., 2017); thus, ground-based PM_{10} observations have become one of the key indicators for monitoring dust events. However, these observations contain particles from multiple sources, including natural dust and other anthropogenic sources (e.g., vehicle emissions, industrial emissions). Therefore, the observed PM_{10} cannot be directly used to calibrate the model for dust aerosol simulations. In this study, the XGBoost model was employed to estimate PM_{10}_dust during dust events by establishing a statistical relationship between PM_{10} in non-dust periods (PM_{10}_nondust) and concentrations of other air pollutants as well as meteorological factors. First, based on ERA5 reanalysis meteorological data and ground-based air pollutant concentrations (NO_2, CO, O_3, $PM_{2.5}$, CO, PM_{10}) from CNEMC, dust periods are identified following the method of Jin et al. (2019) using the two criteria: (1) PM_{10} concentrations exceed five times the monthly average, or the ratio of $PM_{2.5}/PM_{10}$ is below 1/4; (2) visibility is less than 3 km, relative humidity (RH) is lower than 80%, WS10 is greater than 3 $m\ s^{-1}$, and there is no precipitation. Periods satisfying both criteria are classified as “dust days”, while all other periods are treated as non-dust days and used to develop a prediction model for PM_{10}_nondust. Hourly data from the spring months during non-dust

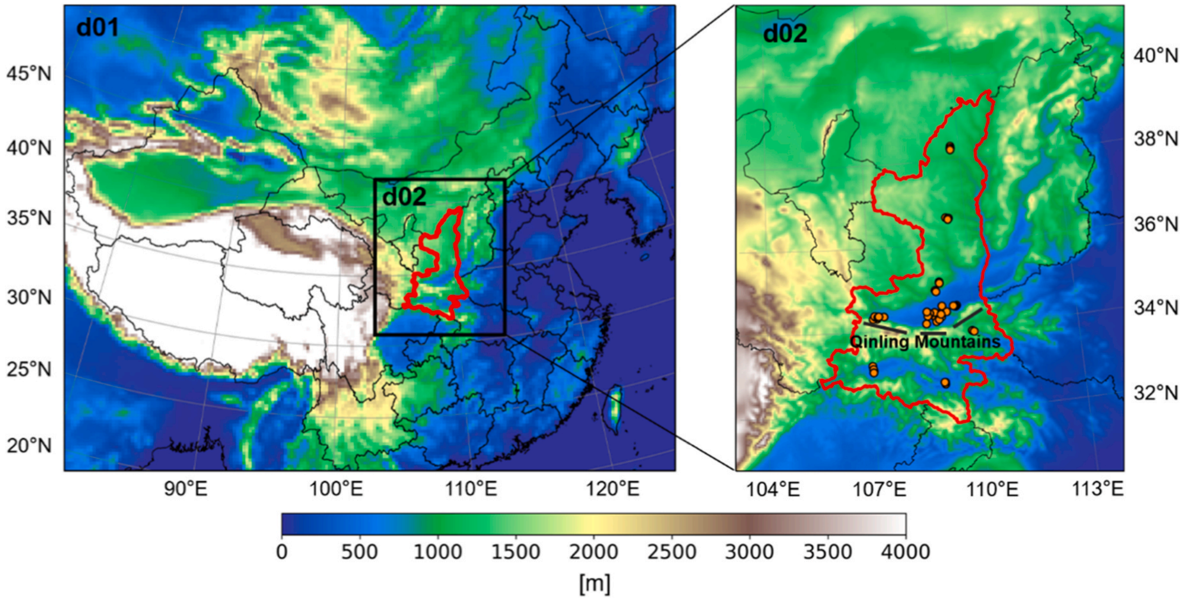


Fig. 1. Simulation domain of the UI-WRF-Chem model depicting the terrain elevation (units: m) across the study region. The dark orange dots in the right panel denote the locations of 51 national monitoring stations within Shaanxi Province, and the black dotted line represents location of the Qinling Mountains, which divides Shaanxi Province into northern and southern regions. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

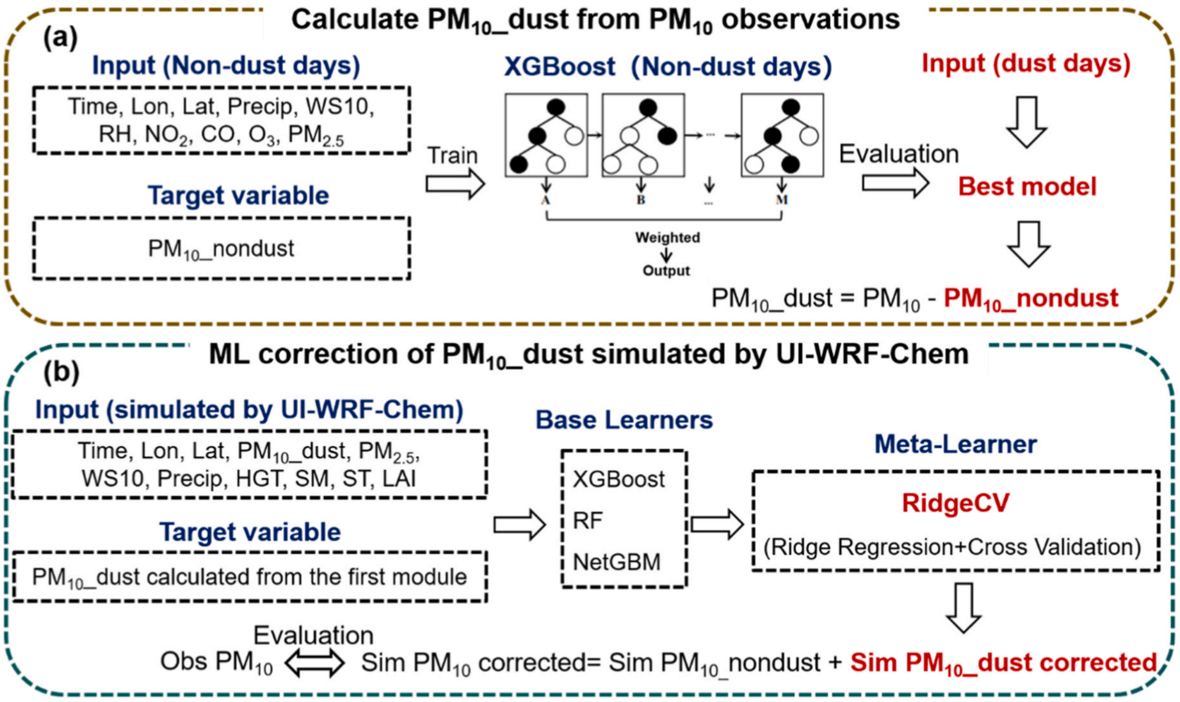


Fig. 2. Flow diagram of PM_{10}_dust concentrations simulated by the UI-WRF-Chem model corrected via a ML ensemble model.

periods in 2018~2023, including NO_2, CO, O_3, $PM_{2.5}$, CO, and PM_{10} concentrations from CNEMC, together with Precip, WS10, RH from ERA5, and auxiliary features such as day of year (Time), longitude (Lon), and latitude (Lat), are used as input to train a station-specific ML model for predicting PM_{10}_nondust (80% for training, 20% for testing). Stations with missing or invalid data on any given day are excluded from the training set to ensure data integrity and model reliability. Subsequently, the trained ML model is then applied to dust periods, and PM_{10}_dust is calculated by subtracting the predicted PM_{10}_nondust from the observed total PM_{10}. Finally, a network of ground-based PM_{10}_dust concentrations covering the study region is established, serving as the training target for the subsequent post-processing ML model (Fig. 2a).

For the second module, we employ XGBoost, RF, and NetGBM (the neural Net-enhanced LightGBM) as the base learner, with the RidgeCV serving as meta learner, to develop the ensemble ML model. XGBoost is a supervised boosting algorithm known for its effectiveness in mitigating overfitting, capturing nonlinear relationships among predictor variables, and providing rapid and accurate solutions to scientific problems (Chen and Guestrin, 2016). It builds tree models incrementally by optimizing an objective function, adding a new tree in each iteration based on the previous one to minimize the loss function. Furthermore, XGBoost incorporates a regularization term to prevent model overfitting, making it particularly effective for high-dimensional data. RF is an ensemble learning algorithm based on the Bagging strategy. It constructs multiple decision trees through bootstrap sampling and random feature selection during node splitting, thereby reducing inter-tree correlation and improving generalization performance (Belgiu and Drăguţ, 2016; Breiman, 2001; Rigatti, 2017). NetGBM, another gradient boosting-based ensemble method, is conceptually similar to LightGBM but with neural network structures integrated into its feature processing module—this strengthens the model's ability to capture complex nonlinear relationships. These three ML algorithms have been widely adopted in air pollutant concentration estimations (Li et al., 2023c; Nababan et al., 2022; Singh et al., 2024; Zhang et al., 2023).

The ensemble model is applied to post-process the simulated PM_{10}_dust from the UI-WRF-Chem model. It is trained on a dataset covering spring seasons during 2018~ 2022, using a 5-fold cross-validation method for validation. Input features include the simulated variables from the UI-WRF-Chem model, i.e., PM_{10}_dust, $PM_{2.5}$, WS10, Precip, RH, terrain height (HGT), SM, soil temperature (ST), and LAI. The target variable is the predicted ground-based PM_{10}_dust concentrations from the first ML model. An independent dataset during spring 2023 is reserved for the ensemble model evaluation. Simulations from the UI-WRF-Chem model are matched to station observations using the nearest-neighbor method. The optimal base ML models (XGBoost, RF, NetGBM) are selected through five-fold cross-validation based on the smallest root mean square error (RMSE). The best predictions from these three ML models are then integrated into a new training set and fed into the final ensemble model, which is built on the Stacking framework to leverage its synergistic advantages. Additionally, we quantitatively evaluate the performance of the numerical and ML models via multiple statistical metrics, including RMSE, mean absolute error (MAE), and Pearson correlation coefficient (R) (Fig. 2b).

2.5. *Simulation experiment design*

A comparative sensitivity experiment was conducted to evaluate the impact of two key improvements on dust aerosol simulations: (1) updating land surface properties in the UI-WRF-Chem model, and (2) applying the ensemble ML to post-process the dust aerosol simulations. Detailed experiment descriptions are summarized in Table S2. The baseline simulation (referred to as Default) uses the standard model configuration without any modifications. In the Updated experiment, key dust-related land surface properties are refined, including LCT, LAI, and GVF, which collectively influence surface moisture and energy exchange. Soil properties (e.g., soil texture) and the monthly dynamic EROD data are also updated to better represent dust emission processes. To further optimize the model performance on simulating dust aerosols, an ensemble ML framework is applied to the Updated outputs, denoted as Updated_ML, which aims to minimize prediction bias and improve the reliability of simulated dust aerosol concentrations through a post-processing procedure.

3. Results

3.1. Impacts of land surface updates on dust emissions

3.1.1. Comparison of soil erodibility

Fig. 3 compares the EROD datasets used in the Default and Updated simulations. Both datasets identify major dust source regions, such as the Taklimakan Desert (TD), Gobi Desert (GD), and Inner Mongolia (IM) in China. The Default dataset exhibits relatively limited spatial heterogeneity, with the EROD value ranging from 0 to 0.17. It fails to capture the erodible areas from southern Mongolia, a region characterized as part of the GD with semi-desert vegetation and sandy soil, as well as the erodibility of the Qilian Desert within the Qaidam Basin (QB) in Qinghai Province. These omissions may lead to inaccurate dust emission predictions. In contrast, the Updated EROD dataset is derived from the monthly satellite observations of LCT, GVF, and SM, enabling better representation of some potential dust source regions and their variations under climate change and human activities. The spring-averaged Updated EROD dataset for 2018∼2023 shows an expanded extent of erodible areas, incorporating areas overlooked by the Default, such as Qinghai, Gansu, northern Tibet, and southern Mongolia, thus providing a more comprehensive identification of potential dust source regions, with EROD values ranging from 0 to 0.15.

The TD is the second largest mobile desert in the world, characterized by frequent sandstorms and high erosion rates. In contrast, the Gurbantunggut Desert (GRD) in northern Xinjiang is a semi-fixed desert with a lower erosion probability than TD. However, the Default inaccurately represents their relative erodibility. The Updated EROD corrects this discrepancy and more realistically depicts the oases surrounding the TD, which appear as discontinuously distributed within the desertification zone. Overall, the monthly dynamic EROD dataset provides a more accurate and reasonable representation of dust sources and exhibits higher spatial heterogeneity than the Default.

3.1.2. Comparison of dust emissions, AOD, and PM$_{10}$

EROD is a critical factor governing dust aerosol formation. To evaluate the impact of a monthly-updated dynamic EROD dataset on dust emission simulations, we select March 2021 as the study period. During this month, northern China experienced two of the most intense dust events in the past decade (dust event 1: March 15 to 20, dust event 2: March 27 to 29), affecting an area of approximately 3.8 × 10^{6} km^{2} (accounting for ∼40% of China's total land area). These events were recognized by the World Meteorological Organization as "the strongest sandstorms in Asia in the past decade" (Filonchyk, 2022; Gui et al., 2022; Liang et al., 2022). The simulated AOD and PM$_{10}$ from the Updated are further evaluated, and the first ML model for predicting PM$_{10}$_dust is also validated during the same period.

Fig. 4 compares the distribution of simulated monthly averaged dust emissions using the Default and Updated EROD datasets. The results demonstrate that the Updated EROD significantly alters both the spatial pattern and intensity of dust emissions. In Default, the simulated monthly averaged dust emission fluxes exceeding 50 μg m^{-2} s^{-2} are mainly limited to southern Mongolia, the eastern Taklamakan Desert in Xinjiang, and parts of Inner Mongolia. In contrast, the Updated simulation exhibits more prominent high-emission centers (fluxes exceeding 50 μg m^{-2} s^{-2}) in the Gobi Desert of southern Mongolia. Furthermore, the simulated monthly averaged emission flux from dust source areas in northwestern China increased to 20 μg m^{-2} s^{-2}. These findings indicate that the Updated EROD substantially enhanced the simulated dust emission intensity in regions characterized by sparse vegetation and extensive dunes, such as Mongolia and northwestern China, whereas the Default simulation fails to capture the emission characteristics of these source areas.

To further quantify the impact of the Updated EROD on dust emission simulations, we divided the high-emission regions into three subregions, namely Northwest, North-central, and Northeast China. As shown in Fig. 4c, the Updated simulation exhibits significantly higher dust emissions over Mongolia and Gobi, with monthly averaged emissions in North-central China surpassing those in the Default by 3.43 μg m^{-2} s^{-1}. The stronger dust emissions over this region in the Updated simulation can be attributed to the dynamic EROD dataset, which captures seasonal variations in vegetation and soil moisture more

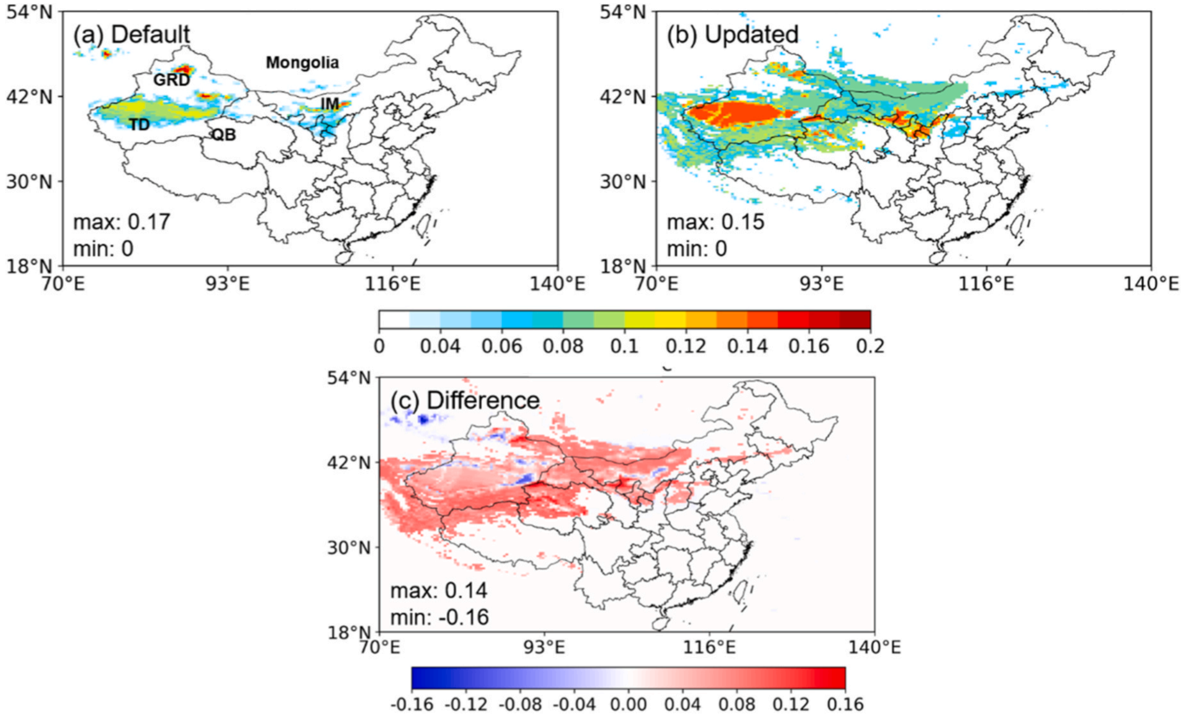


Fig. 3. Distribution of soil erodibility factor in China from the (a) Default and (b) Updated simulation, and (c) the absolute difference between the two simulations (Updated–Default). The Updated EROD represents the average over the spring season (March to May) from 2018 to 2023.

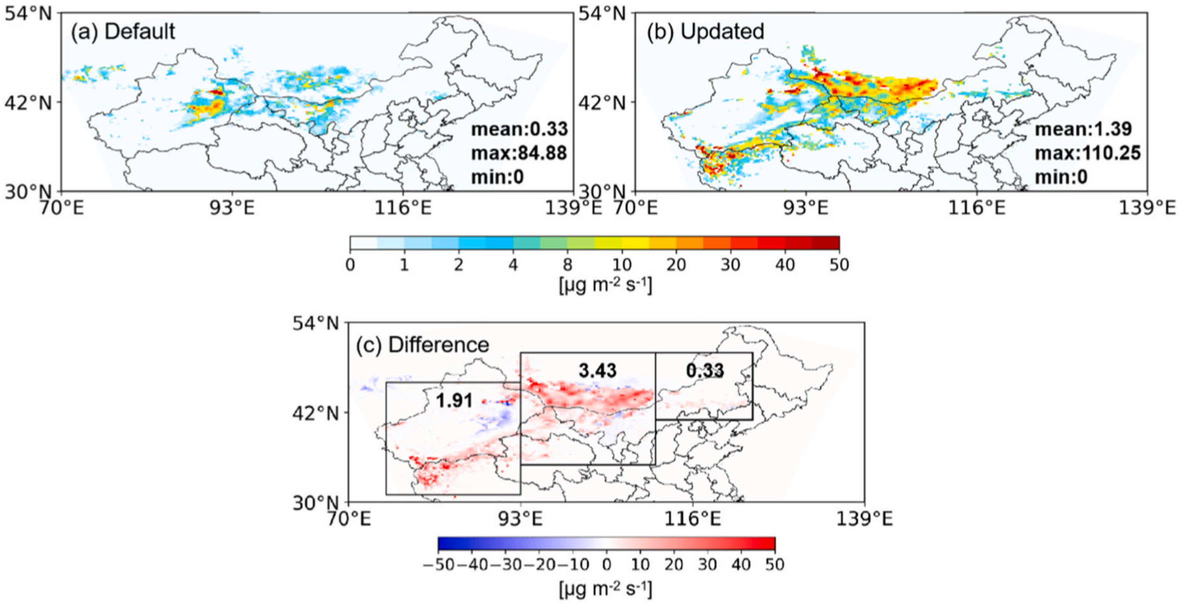


Fig. 4. Distribution of the monthly averaged dust emissions in March 2021 from (a) the Default and (b) Updated simulations, and (c) the absolute difference between the two simulations (Updated–Default). Rectangles represent the three major dust sources over northwest China, Mongolia, and northeast China.

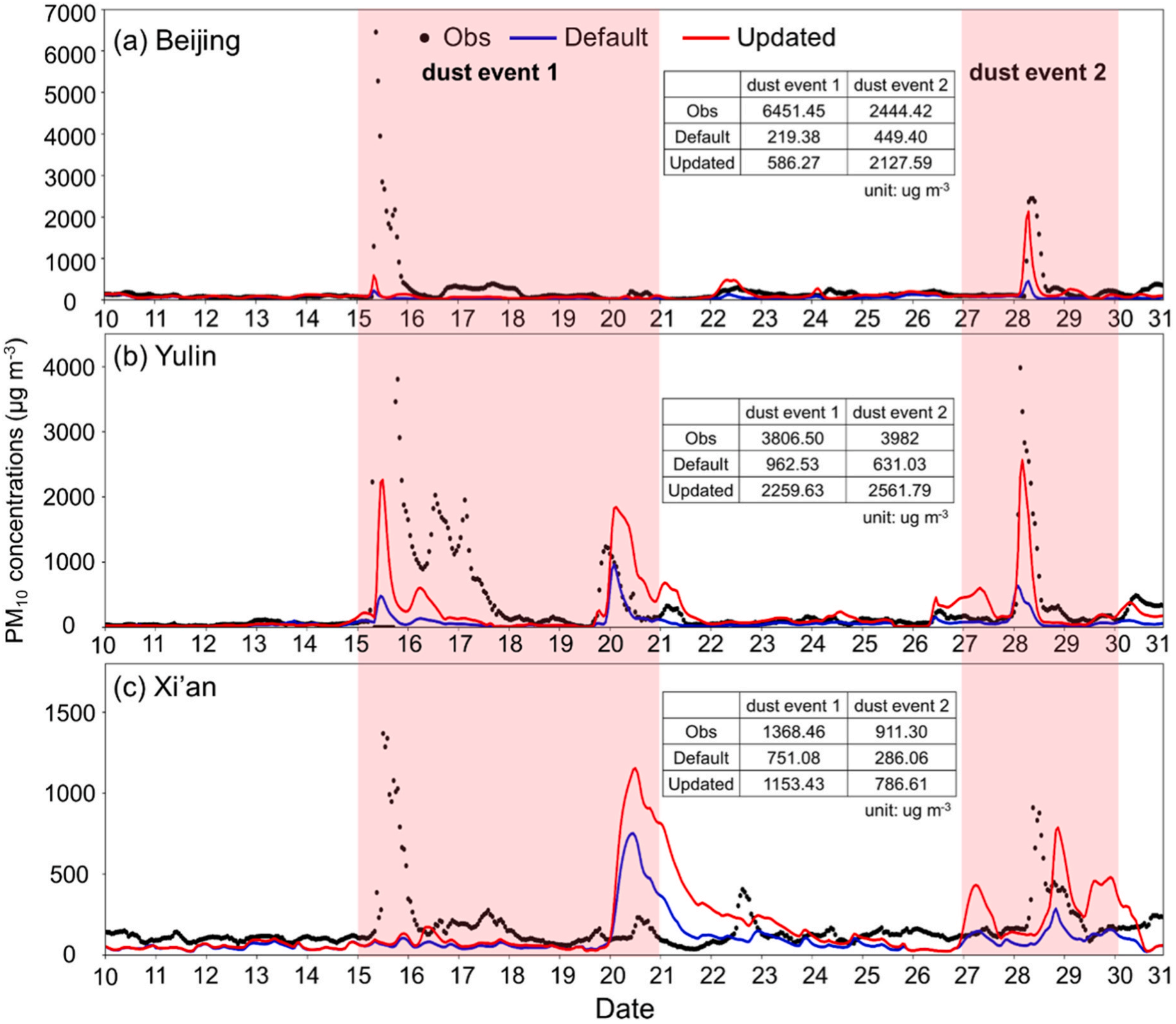
realistically than the static default. This finding is consistent with recent research by Zhang et al. (2024), who demonstrated that dynamic dust source data can effectively improve WRF-Chem simulations in Mongolia, leading to better agreement with observed dust emission patterns. Their study further emphasizes that a detailed representation of dust source distributions in Mongolia and their dynamic monthly variations is important for accurately quantifying dust emissions in China. In contrast, dust emissions in the TD from the Updated simulation are substantially lower than those from the Default, indicating that the Updated EROD dataset not only confirms the dominant contribution of Mongolia to total dust emissions but also suggests a reduced relative contribution from northwestern China. Moreover, the Updated EROD dataset reveals previously unidentified potential dust sources in eastern China, where the monthly averaged emission flux in the Updated simulation is 0.33 μg m^{-2} s^{-1} higher than that in the Default.

To further evaluate the impacts of refined dust-related land surface properties and EROD data on simulated AOD and surface PM_{10} concentrations, we focused on assessing the Updated simulation performance for two dust episodes in March 2021 (dust event 1: March 15 to 20, dust event 2: March 27 to 29). Fig. S2 shows the spatial distributions of AOD from the Default and Updated simulations during the two dust events, compared with MODIS observations. Although the spatial pattern of AOD in the Default simulation is generally consistent with MODIS, it fails to reproduce the high AOD regions observed by MODIS during the dust events. For dust event 1, MODIS recorded high AOD over northwest and north-central China, with the AOD of 0.78 and 1.1, whereas the Default simulation substantially underestimates AOD in these regions (northwest: 0.27; north-central: 0.23). Although the Updated simulation still underestimates AOD (northwest: 0.39; north-central: 0.32), it shows improved agreement with MODIS in terms of the spatial extent of high AOD (Table S3). Additionally, the Updated simulation can capture the evolution of dust event 2, i.e., the high AOD over north-central China on March 27 (MODIS: 0.58; Updated: 0.27), followed by southeastward transport. It is worth noting that the Updated simulation still systematically underestimates AOD when observed AOD exceeds 0.3, while much better agreement is achieved for lower AOD. This indicates that despite the improvements in the dynamic EROD dataset, the model still fails to capture the magnitude of dust emissions during intense dust storms. One possible reason is that the dust emission scheme may not adequately represent the high dust emission fluxes associated with extreme meteorological conditions (e.g., strong winds over dry and unvegetated surfaces). Additionally, uncertainties in deposition, specifically an overestimation of dry and wet scavenging efficiency under high dust loading, may result in excessive dust removal along the transport pathway.

Fig. 5 compares the simulated and observed hourly PM_{10} concentrations in three northern cities in China (Beijing, Yulin, and Xi'an) during the two dust events. We use hourly data for comparison because dust events are characterized by rapid temporal variations, and high-frequency data is essential to capture the critical phases of onset, peak, and dissipation. Accordingly, all subsequent PM_{10} evaluations in this study are conducted at an hourly scale to ensure a robust assessment of model performance. The Default simulation significantly underestimates PM_{10} by 97% and 82% in Beijing, 75% and 84% in Yulin, and 45% and 69% in Xi'an, respectively, across the two events. This is attributed to the limitations of the default EROD dataset, which failed to accurately characterize dust emissions in north-central China. Although discrepancies between simulations and observations still remain in the Updated, the updated EROD dataset implemented in the UI-WRF-Chem model enhances the peak PM_{10} concentrations during dust events. Thus, the Updated simulation improves the PM_{10} simulation performance, with the bias decreased to 91% and 13% in Beijing, 41% and 36% in Yulin, and 16% and 14% in Xi'an, respectively, for the two dust events. Overall, the simulated dust emissions and AOD exhibit similar spatially synchronized changes following the EROD updates. The changes in simulated dust emissions are largely governed by the updated EROD values; regions with increased erodibility (e.g., southern Mongolia) show a corresponding increase in both dust emission fluxes and AOD during dust storms. However, in areas where EROD decreases (e.g., parts of the Taklamakan Desert), the reduction in local emissions is occasionally offset by the long-range transport of dust from upwind source regions, resulting in a less pronounced decrease in AOD.

3.2. *p.m.*$_{10}$*_dust prediction*

Before applying the ensemble ML model, it is essential to derive observed PM_{10}_dust concentrations to serve as the training target. To achieve this, we use an XGBoost model to separate dust from non-dust components of PM_{10}. The section presents the validation and performance results of this initial separation process.



(a) Beijing

	dust event 1	dust event 2
Obs	6451.45	2444.42
Default	219.38	449.40
Updated	586.27	2127.59

unit: ug m^{-3}

(b) Yulin

	dust event 1	dust event 2
Obs	3806.50	3982
Default	962.53	631.03
Updated	2259.63	2561.79

unit: ug m^{-3}

(c) Xi'an

	dust event 1	dust event 2
Obs	1368.46	911.30
Default	751.08	286.06
Updated	1153.43	786.61

unit: ug m^{-3}

Fig. 5. Time series of hourly observed and simulated PM$_{10}$ concentrations in (a) Beijing, (b) Yulin, and (c) Xi'an city during March 10~30, 2021. The black dots represent the observation results, while the blue and red lines represent the Default and Updated simulation results, respectively. The tables within each panel provide the average PM$_{10}$ concentrations for the two dust events (dust event 1: March 15 to 20, dust event 2: March 27 to 29). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

Although the performances of the station-specific XGBoost models for PM$_{10}$_nondust prediction vary across national monitoring stations in Shaanxi Province during the validation period, their overall performances are robust. From Fig. S3, the models achieve an average coefficient of determination (R^2) of 0.84, with RMSE and MAE being relatively low and stable levels at 16.01 μg m^{-3} and 11.26 μg m^{-3}, respectively. These results indicate that the trained XGBoost model (i.e., the first module in Fig. 2) can effectively predict PM$_{10}$_nondust concentrations during both non-dust and dust episodes across different regions, thereby providing a reliable foundation for subsequent dust-related PM$_{10}$ prediction.

Fig. S4 presents the time series of observed hourly PM$_{10}$ concentrations together with the XGBoost-predicted PM$_{10}$_nondust and PM$_{10}$_dust concentrations for three cities at different latitudes in Shaanxi Province, Yulin (38.18 °N, 109.47 °E), Xi'an (34.17 °N, 108.57 °E), and Hanzhong (33.04 °N, 107.01 °E), during the spring seasons from 2018 to 2023. The PM$_{10}$_dust peaks predicted by the model align closely with the observed PM$_{10}$ peaks, indicating that the XGBoost model effectively captures the increases in PM$_{10}$ concentrations caused by dust events. Moreover, PM$_{10}$_dust concentrations rise sharply during these dust episodes, whereas PM$_{10}$_nondust levels remain relatively stable. This phenomenon is consistent with the findings of Jin et al. (2018), who also reported an increase in PM$_{10}$ concentrations during dust events, whereas PM$_{2.5}$ concentrations are less affected and generally remain unchanged.

To further demonstrate the accuracy and reliability of the XGBoost model in predicting PM$_{10}$_dust, we conducted a case study using a typical dust episode (March 15~20 in 2021), consistent with Section 3.1. Distinct PM$_{10}$ concentration peaks were sequentially observed in Yulin, Xi'an, and Hanzhong (Fig. 6). Located at the edge of the Mu Us Desert in northern Shaanxi Province, Yulin was the first to be affected by the dust event, with the observed PM$_{10}$ peak of 4613.30 μg m^{-3} and the XGBoost predicted PM$_{10}$_dust peak of 4467.40 μg m^{-3}. Both peaks occurred simultaneously at 18:00 local time on March 15, demonstrating temporal synchronization (Table S4). Additionally, the ratio of PM$_{10}$_dust to PM$_{10}$ remained stable between 0.73 and 0.96 (Fig. S5), suggesting that PM$_{10}$ in Yulin city was predominantly composed of dust, which is consistent with the city's geographical location adjacent to dust source regions. For Xi'an, which is located at a lower latitude than Yulin, the PM$_{10}$ peak lagged behind Yulin city by 18 h. The observed PM$_{10}$ and predicted PM$_{10}$_dust peaks were 917.70 μg m^{-3} and 761.80 μg m^{-3}, respectively. The ratio of PM$_{10}$_dust to PM$_{10}$ ranged from 0.50 to 0.80, suggesting a lower contribution of dust to total PM$_{10}$ compared to Yulin, likely resulting from the downwind transport of dust combined with local emissions. In Hanzhong, located at the southern foot of the Qinling Mountains, a much lower PM$_{10}$ peak of 417 μg m^{-3} was observed, with the predicted PM$_{10}$_dust peak of only 291 μg m^{-3}. The PM$_{10}$_dust to

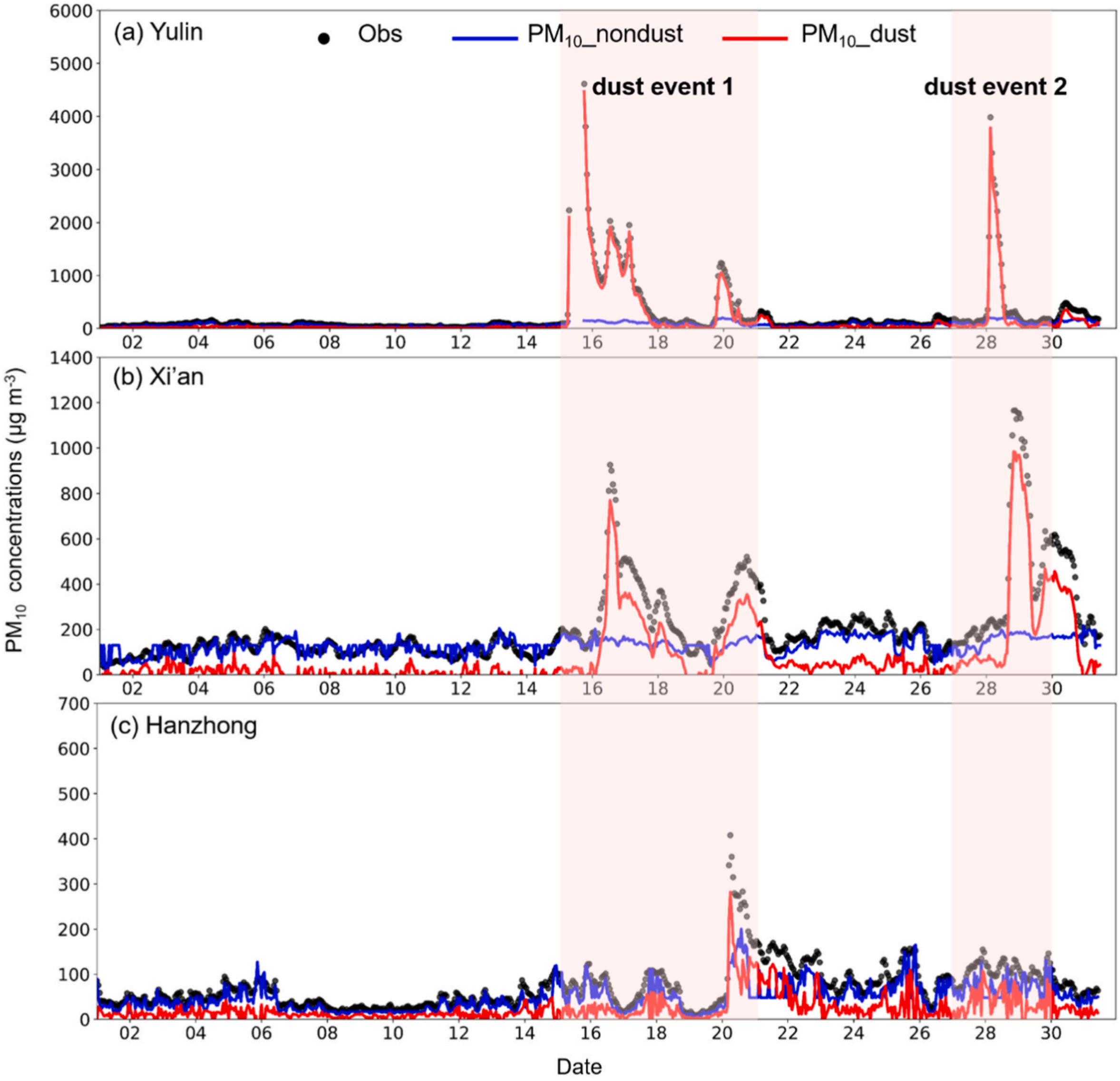


Fig. 6. Time series of hourly PM_{10} concentrations in (a) Yulin, (b) Xi'an, and (c) Hanzhong during March 2021. The black dots denote the observed PM_{10} concentrations, the blue line represents the PM_{10}_nondust concentrations predicted by the XGBoost model, the red line represents the PM_{10}_dust concentrations, calculated as the difference between observed PM_{10} and predicted PM_{10}_nondust concentrations, and the red shadings represent the two major dust events from March 15~20 and 27~29. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

PM_{10} ratio ranged from 0.34 to 0.69, reflecting a further decreased dust contribution due to the blocking effect of the Qinling Mountains and the along-path loss during the transportation of dust aerosols, which enhances dust deposition and increases the relative contribution of non-dust sources such as biomass burning and secondary aerosols (Table S4). Furthermore, the XGBoost-predicted PM_{10}_nondust concentrations in the three cities during this period remained stable within the range of 103~131.40 μg m^{-3}, which is comparable to the background concentrations during non-dust periods (80.98~~147.59 μg m^{-3}). This suggests that anthropogenic emissions and secondary aerosols are only marginally affected by the dust events. Overall, these results confirm that the XGBoost model can effectively distinguish between dust and non-dust contributions to PM_{10} and non-dust sources, and accurately capture regional variations in dust impacts.

3.3. Comparison and evaluation of the ML models' results

Fig. 7 compares the performance of the UI-WRF-Chem model against four ML models, i.e., NetGBM, XGBoost, RF, and an ensemble model, in simulating PM_{10}_dust concentrations during the validation period. The UI-WRF-Chem model shows limited performance in capturing dust-related PM_{10} variability, with a low R^2 of 0.21 and relatively high RMSE and MAE of 219.45 μg m^{-3} and 107.29 μg m^{-3}, respectively. To address these limitations, we employed multiple ML models for post-processing, leveraging UI-WRF-Chem outputs as ML model inputs to enhance the simulation performance. As shown in Fig. 7b–e, all ML models effectively reduce the simulation bias present in the UI-WRF-Chem results, but their post-processing performances are different. The XGBoost and RF models show the best performance, with R^2 of 0.90 and 0.89, RMSE of 29.08 μg m^{-3} and 30.58 μg m^{-3}, and MAE of 11.87 μg m^{-3} and 12.69 μg m^{-3}, respectively. The post-processing performance of the NetGBM model is relatively poor, with an R^2 of 0.85, RMSE of 38.02 μg m^{-3}, and MAE of 16.77 μg m^{-3}. The superior performance of the tree-based model (XGBoost and RF) over NetGBM likely stems from the regional scale of the dataset and its feature distribution. As a gradient boosting model that integrates neural network components, NetGBM may be more sensitive to the imbalanced distribution of sandstorm event samples (Ma et al., 2024). The previous study

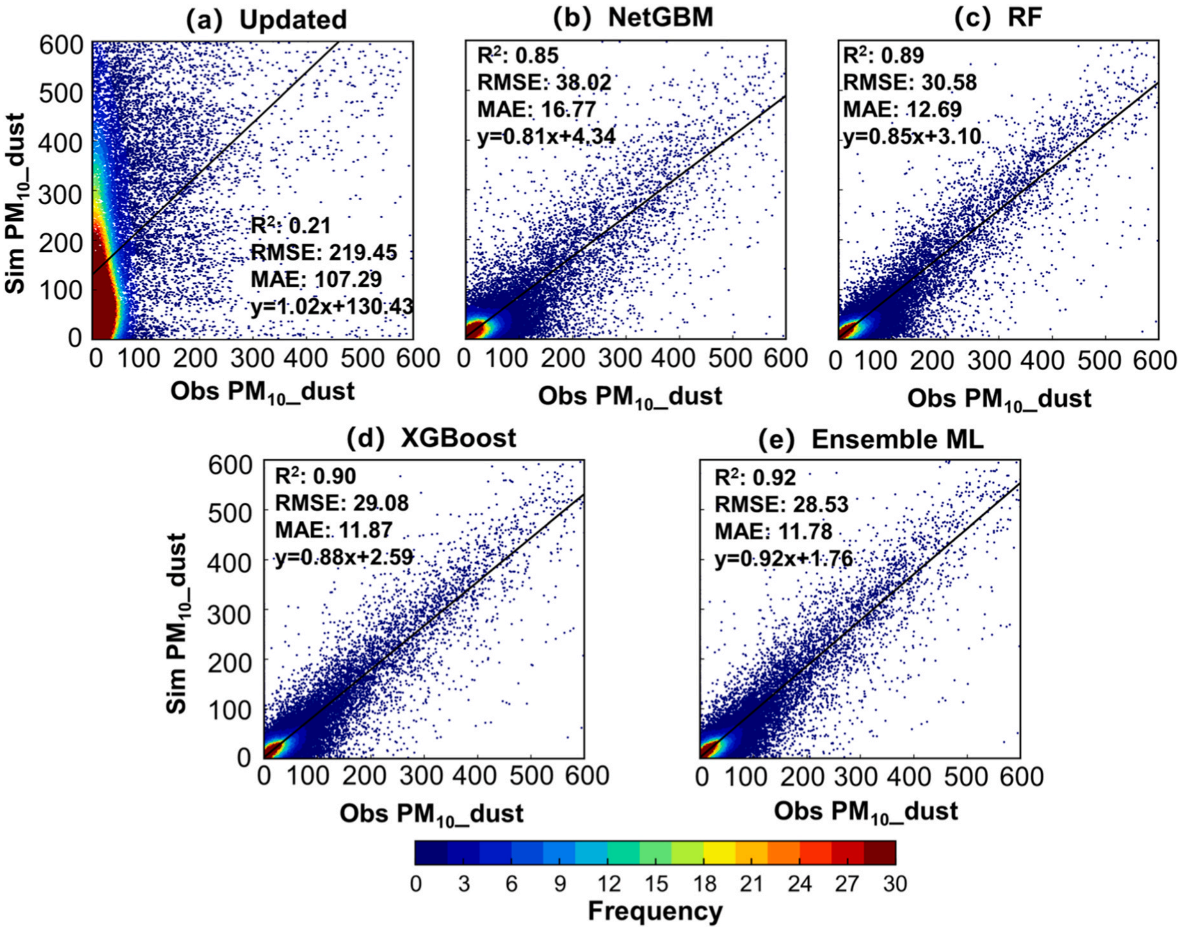


Fig. 7. Density scatter plots of five-fold cross-validation results for simulated and observed PM_{10}_dust during spring (March to May) from 2018 to 2022 using the UI-WRF-Chem (Updated), NetGBM, RF, XGBoost, and Ensemble ML models. The frequencies in the legend at the bottom denote the density of the data distribution. The regression line equation and statistics (R^2, RMSE, MAE) are provided in each panel.

also demonstrated that the RF model shows greater robustness in sandstorm prediction (Kazemi et al., 2026). Furthermore, our ensemble model does not simply average outputs, it leverages the strengths of diverse base learners to reduce individual model uncertainty and variance. This mechanism allows the ensemble framework to capture extreme dust peaks more effectively than any single constituent model (Shi et al., 2024). By integrating the strengths of these individual ML models, the ensemble model further reduces uncertainties associated with single models and thereby achieves the optimal performance, with the highest R^2 of 0.92 and the lowest RMSE of 28.53 μg m^{-3} and MAE of 11.78 μg m^{-3} among all ML models. This ensemble approach not only significantly improves the dust-related PM_{10} simulation performance but also enhances the model stability, thereby optimizing overall correction performance. Therefore, the ensemble ML model is selected as the post-processing tool for the PM_{10}_dust simulation results, aiming to accurately characterize dust-related PM_{10} concentrations during dust events.

3.4. *Performance on PM_{10} simulation*

To validate the post-processing performance of the ensemble model in refining PM_{10}_dust simulation results from the UI-WRF-Chem model, we selected the spring of 2023 as the evaluation period, during which seven dust events occurred. Details of these events and associated observed PM_{10} concentrations are summarized in Table S5. Fig. 8 compares the observed and simulated hourly PM_{10} concentrations, averaged across 51 national monitoring sites in Shaanxi Province from March to May 2023. We choose to directly compare total PM_{10} concentrations to intuitively assess whether the ensemble ML model can capture the impacts of the dust events' onset and dissipation processes on PM_{10} variations. Although the Updated simulation could capture the overall temporal variation of PM_{10} concentrations, it shows relatively poorly in reproducing PM_{10} peak levels during dust events. Specifically, the Updated simulation overestimates the PM_{10} levels across all seven spring dust events, exhibiting an overestimation of 203.20 μg m^{-3}, RMSE of 410.40 μg m^{-3}, and MAE of 307.72 μg m^{-3}, respectively. In contrast, the Updated_ML simulation better reproduces the PM_{10} variations during dust events, reducing the bias to -123.66 μg m^{-3}, RMSE to 173.23 μg m^{-3}, and MAE to 131.62 μg m^{-3}, respectively. On the whole, the ensemble model effectively improves the accuracy of the PM_{10} simulation, particularly during intense dust transport processes. Nevertheless, the ensemble model still has limitations and fails to accurately capture the PM_{10} peak levels during certain dust events (e.g., April 6 to 7 and 19 to 21). This shortcoming may be attributed to uncertainties in the input data, which leads to inconsistent performance of the ML model when predicting extreme PM_{10} peaks. As reported by Chen et al. (2024), who also suggested that significant discrepancies between ML predictions and observations often arise when air pollutant concentrations abruptly increase. Furthermore, dust events are low-frequency phenomena, thereby the training dataset of the ensemble model lacks sufficient samples of extreme sandstorm events with high PM_{10} concentrations, and the corresponding meteorological-surface configurations (e.g., specific combinations of soil moisture, wind fields, and vegetation cover). Since ML models are fundamentally based on historical distributions, their extrapolation capabilities for rare or unobserved conditions remain constrained. This limitation has been

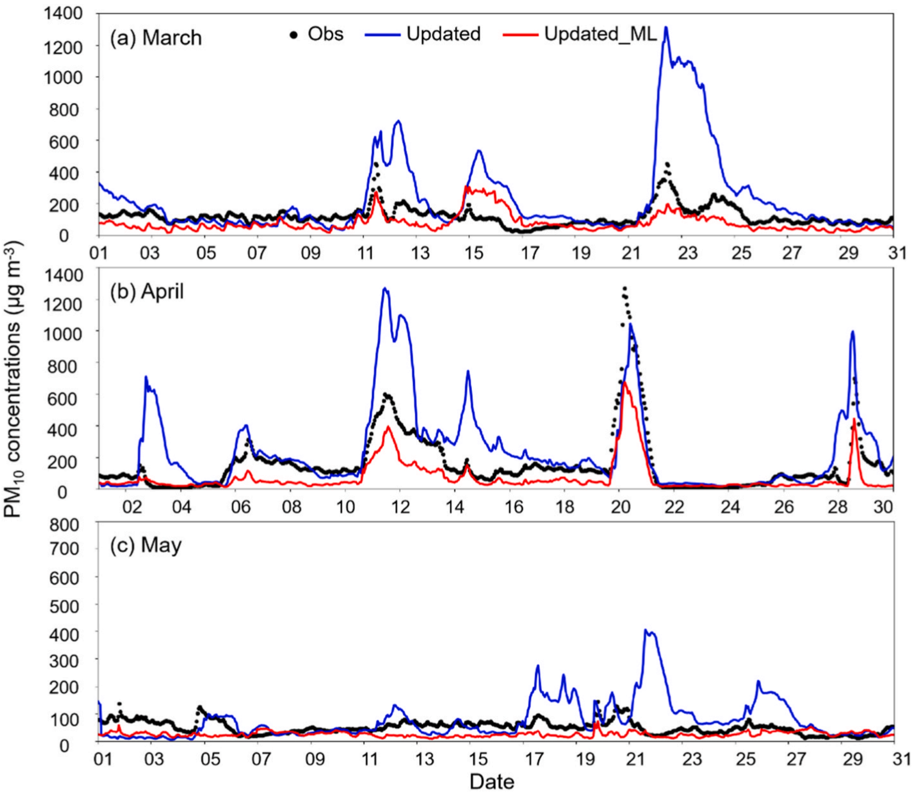


Fig. 8. Time series of observed and simulated hourly PM_{10} concentrations from Updated and Updated_ML in Shaanxi Province averaged across 51 monitoring sites during (a) March, (b) April, and (c) May 2023. The black dots denote the observations, and the red and blue lines denote the simulation from the Updated and Updated_ML, respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

clearly reported in a previous study, which found that ML models systematically underestimate high concentrations of PM_{10} when predicting dust events in the southwestern United States (Aryal, 2022).

A Taylor diagram was used to evaluate the performance of the Updated and Updated_ML simulations against monthly mean PM_{10} concentrations observed at 51 monitoring sites during the spring (March–May) of 2023 (Fig. S6). The correlation coefficient (R) characterizes the models' ability to reproduce the temporal variations in PM_{10} concentrations. The Normalized Standard Deviation (NSD) quantifies the consistency of dispersion between simulations and observations, and the Centered Root Mean Square Error (CRMSE), represented by purple arcs, denotes the overall discrepancy between the simulations and observations ("Obs" point) across the three months. Specifically, the Updated_ML simulation shows much higher R (0.36~0.85) and lower NSD (0.57~0.86) and CRMSE (0.59~0.96) than the Updated (R: 0.24~0.56, NSD: 1.35~3.05, CRMSE: 1.14~2.70) across these three months (detailed statistics are shown in Table S6). Notably, the performance of Updated_ML simulation exhibits monthly variability: April shows the best agreement with observations (R = 0.85, NSD = 0.57, CRMSE = 0.59), while May yields the most modest improvement (R = 0.36, NSD = 0.62, CRMSE = 0.96). The monthly variation is likely attributable to differences in the frequency and intensity of dust events. April typically experiences more frequent and stronger dust storms in Shaanxi (Wang et al., 2022), which dominate the PM_{10} signal. This allows the ML model, trained on abundant dust event samples, to perform more effectively. In contrast, May is characterized by fewer and weaker dust events with non-dust sources (e.g., local anthropogenic emissions, secondary aerosols) contributing more to total PM_{10}. Since our ensemble model is specifically designed to correct the dust component, its improvement over the raw simulation is inherently limited during periods dominated by non-dust sources. Additionally, inter-monthly variations in meteorological conditions (e.g., wind fields, precipitation) influence dust transport and deposition, potentially leading to fluctuations in model performance. Overall, these results demonstrate that the Updated_ML simulation agrees more closely with observations and exhibits better overall performance across the spring months compared to the Updated simulation.

To analyze the post-correction performance of the ensemble ML model for the UI-WRF-Chem PM_{10} simulations in different regions across the Shaanxi Province during spring 2023, we compared the distribution of simulated bias, RMSE and R for PM_{10} between the Updated and Updated_ML (Fig. 9). In northern Shaanxi (north of the Qinling Mountains as shown in Fig. 1), a region frequently affected by dust events, the Updated simulation overestimates PM_{10} concentrations by approximately 99.67~136.55 μg m^{-3} relative to observations, with RMSE ranging from 299.76 to 385.93 μg m^{-3}. In contrast, the Updated_ML reduces the bias to −57.42~-30.88 μg m^{-3}, and the RMSE decreases to 99.95~174.72 μg m^{-3} (Table S7). These improvements indicate that Updated_ML effectively mitigates the overestimation of dust-related PM_{10} components in the UI-WRF-Chem model. However, the post-processing performance of Updated_ML is more limited in southern Shaanxi (south of the Qinling Mountains as shown in Fig. 1), a non-dust source region. Compared to the Updated simulation, the bias and RMSE of Updated_ML are only reduced by 24.24 μg m^{-3} and 87.05 μg m^{-3}, respectively. This can be attributed to the relatively low contribution of dust to total PM_{10} in this region, combined with the fact that the ML models do not incorporate calibration for non-dust PM_{10} components. Regarding the distribution of R, the Updated_ML simulation shows higher R (0.55~0.72) in northern Shaanxi than the Updated (R: 0.52~0.66). However, in downwind areas influenced by long-range dust transport, the values of R for the two experiments are comparable with no significant differences. The above verifies the reasonable performance of the ensemble ML model in simulating PM_{10} concentrations at 51 national monitoring stations across Shaanxi during the spring of

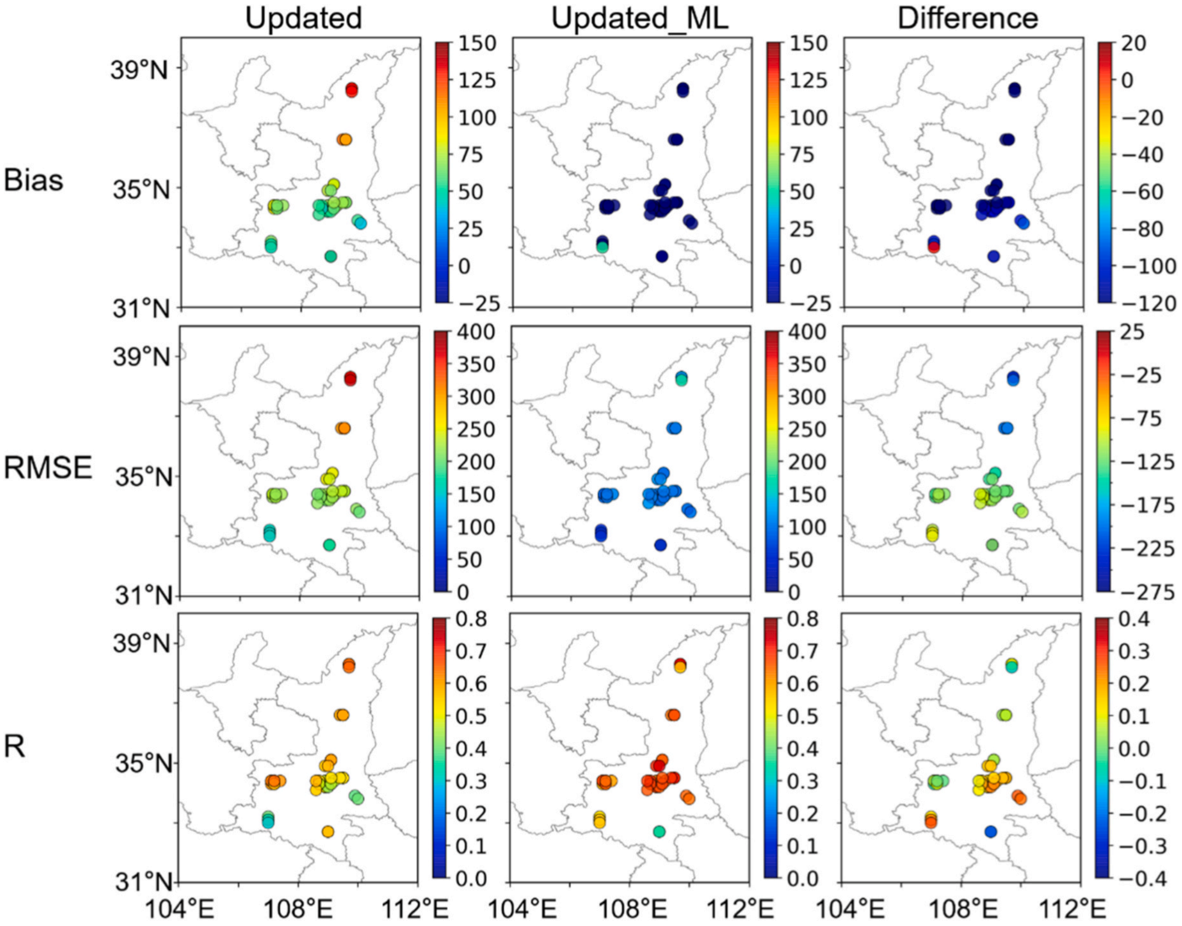


Fig. 9. Spatial distribution of simulation bias (first row), RMSE (second row), and R (third row) from Updated (left column) and Updated_ML (middle column) simulations, and their differences (Updated_ML – Updated, right column) of the three metrics at 51 national monitoring stations in Shaanxi Province from March to May 2023. (units of bias and RMSE are μg m^{-3}).

2023.

To further assess whether the Updated_ML could accurately reproduce the dust transport process and capture the spatiotemporal distribution characteristics of PM$_{10}$ across Shaanxi during dust events, a typical dust episode from April 9 to 12 in 2023 was selected as a case study. This dust storm originated in southern Mongolia and western Inner Mongolia, China, and was transported southeastward under the influence of Mongolian cyclones and northwest airflows (Chen et al., 2024). By 12:00 local time on April 9th, the dust storm reached the northern Shaanxi Province. The observed hourly concentrations of PM$_{10}$ in Yulin city reached 1000~1200 μg m^{-3}, resulting in significantly reduced visibility. On April 11, a dust storm spread southward into the Guanzhong Basin, where observed hourly PM$_{10}$ concentrations in Xi'an and Baoji cities exceeded 600 μg m^{-3}. From April 12 to 13, the dust storm continued to move southward. Subsequently, as cold air weakened and precipitation occurred, the dust storm dissipated in Shaanxi Province on April 13, with PM$_{10}$ concentrations decreasing below 300 μg m^{-3} (Fig. S7).

Fig. 10 shows the distribution of simulated hourly PM$_{10}$ from Updated and Updated_ML against the observations across the Shaanxi Province during the dust event from April 9 to 12, 2023. Although the Updated can capture the southeastward dust transport pattern and the rapid increase in PM$_{10}$ concentrations, it significantly overestimates PM$_{10}$ concentrations in the northern study region. Specifically, the simulated peak PM$_{10}$ concentrations exceed 2000 μg m^{-3}, whereas the observed peak is only 754 μg m^{-3} (Table S8). Moreover, the variation range of simulated PM$_{10}$ concentrations across the entire Shaanxi Province from the Updated is 25.90~2061.90 μg m^{-3}, which is much higher than the observations (54~753.71 μg m^{-3}). The ensemble ML model effectively reduces the simulated peak PM$_{10}$ concentrations and yields a variation range of 110.30~360.51 μg m^{-3} during this dust event. It also reproduces the spatiotemporal variation in high PM$_{10}$ concentrations, thereby enabling the simulations to better align with the observations. Consequently, the ensemble ML model significantly enhances the PM$_{10}$ simulation performance of the UI-WRF-Chem model during extreme dust storm episodes.

4. Conclusion and discussion

In this study, we developed a comprehensive framework for improving dust aerosol simulations over Shaanxi Province, a typical Northwest region downwind of the Gobi Desert and prone to dust storms, through the synergistic integration of an updated UI-WRF-Chem model and a ML-based post-processor. To address key limitations in the traditional dust model (GOCART-AFWA), we introduced a dynamic EROD data derived from multi-satellite observations on a monthly scale.

The integration of updated land surface parameters within the UI-WRF-Chem model and the post-correction approach from the ensemble ML model collectively form a robust framework that significantly enhances the simulation of PM$_{10}$ concentrations for the spring of 2023. The ensemble ML model can yield markedly higher R (0.36~0.85) and lower NSD (0.57~0.86) and CRMSE (0.59~0.96) of the simulated PM$_{10}$ compared to the UI-WRF-Chem outputs (R: 0.24~0.56, NSD: 1.35~3.05, CRMSE: 1.14~2.70) across these three months in 2023. Compared with previous CTM-only studies (e.g., Han et al., 2021; Lee and Lee, 2022), our hybrid framework achieves a substantial reduction

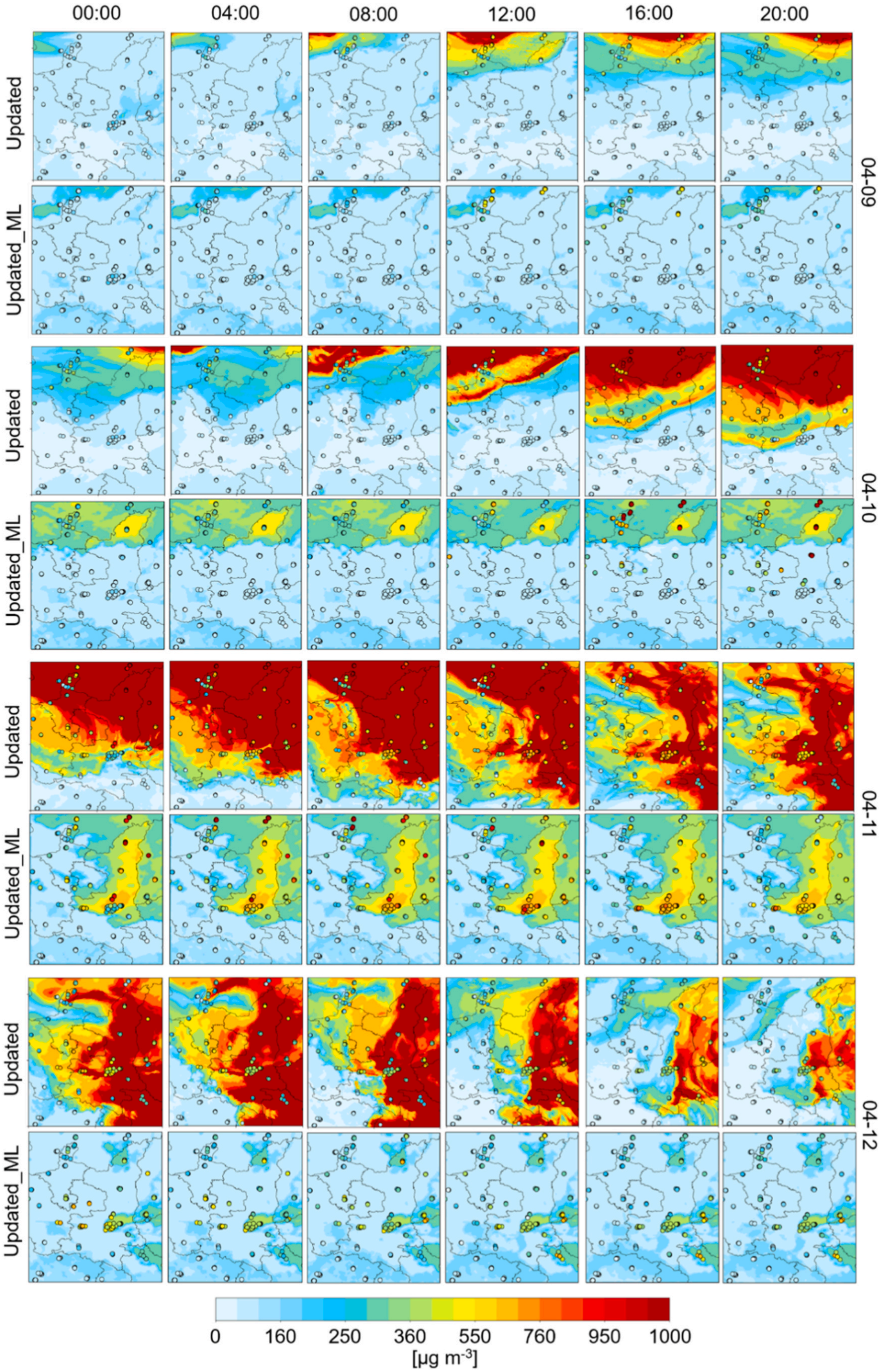


Fig. 10. Spatial distribution of observed and simulated (Updated and Updated_ML) hourly PM_{10} concentrations at 00:00, 04:00, 08:00, 12:00, 16:00, and 20:00 local time during the dust event from April 9 to 12, 2023. The circles represent the observations.

in PM_{10} simulation bias (from 203.2 μg m^{-3} to −123.66 μg m^{-3} for spring 2023). This improvement is comparable to or even exceeds results achieved using ensemble Kalman filtering (Raman et al., 2021), but with the distinct advantage of significantly lower computational costs. Furthermore, the post-correction performance of the ensemble ML model exhibits significant spatial heterogeneity. Across the Shaanxi Province, the ML post-processing reduces the bias and RMSE by 24.24∼70.24 μg m^{-3} and 87.05∼214.10 μg m^{-3}, respectively, while the improvement in R is more modest, particularly in southern Shaanxi. The superior performance in northern Shaanxi aligns with a previous study that model simulations are more effective in areas close to dust source regions (Jin et al., 2018). Unlike the statistical post-processing approach, our framework explicitly distinguishes between dust and non-dust components, thereby providing results with clear physical interpretability.

However, there are still several limitations that require improvement. First, dust storms occur infrequently, leading to a scarcity of high PM_{10} concentration samples in the ML model training dataset. Consequently, the models cannot fully learn the concentration variation patterns under extreme dust storm scenarios, which affects the accuracy of predicting the peak dust concentrations during such episodes. Second, the current model only corrects PM_{10}_dust and does not include PM_{10}_nondust, such as particles from anthropogenic sources. In regions with low dust storm frequency where total PM_{10} is dominated by anthropogenic pollutants, the model fails to effectively reduce simulation biases, limiting its regional applicability. Differences in surface properties (e.g., soil texture and vegetation coverage) and meteorological fields across regions may cause spatial heterogeneity in simulation performance, and the model's cross-regional generalizability requires further verification. Accordingly, future research should intensify ground-based monitoring in high-frequency dust storm regions to obtain high-accuracy PM_{10} observations, which will refine the ML model's training dataset for extreme dust events and improve its predictive robustness. Additionally, the ML post-processing framework should be expanded to incorporate both dust and non-dust PM_{10} components, integrating anthropogenic emission inventories, such as those from industrial, vehicular, and biomass burning sources, to enable accurate total PM_{10} simulation across different regions.

Data statement

The MODIS land cover type product (MCD12Q1, collection 6.1), the leaf area index, and the Fraction of Photosynthetically Active Radiation absorbed by vegetation product (MCD15A2H, collection 6.1) can be downloaded from https://search.earthdata.nasa.gov/. Soil Moisture Active Passive (SMAP) satellite data can be downloaded from https://nsidc.org/data/smap/data/. Harmonized World Soil Database v2.0 (HWSD2.0) can be downloaded from https://doi.org/10.4060/cc3823en. MODIS daily AOD data are available at https://www.ladsweb.modaps.eosdis.nasa.gov/. Hourly meteorological data from ERA5 can be downloaded from https://cds.climate.copernicus.eu/datasets/reanalysis-era5-land?tab=download. Hourly ground-based observation of air pollutant concentrations from CNEMC can be downloaded at http://www.cnemc.cn/.

CRediT authorship contribution statement

Tong Sha: Conceptualization, Funding acquisition, Methodology, Writing – original draft, Writing – review & editing. **Liangqing Li:** Formal analysis, Investigation, Methodology, Writing – original draft, Writing – review & editing. **Zipeng Dong:** Data curation, Writing – review & editing. **Qingcai Chen:** Supervision, Writing – review & editing. **Shuqi Yan:** Writing – review & editing. **Huanxin Zhang:** Writing – review & editing. **Jhoon Kim:** Supervision, Writing – review & editing. **Jun Wang:** Conceptualization, Supervision, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This study is supported by the National Natural Science Foundation of China (grant nos. 42205107), Opening Project of Shanghai Key Laboratory of Atmospheric Particle Pollution and Prevention (LAp^3) (FDLAP24016), and Opening Project of China Meteorological Administration Eco-Environment and Meteorology for The Qinling Mountains and Loess Plateau Key Laboratory (Development of a Dust Aerosol Forecasting System for Shaanxi Province Based on Machine Learning Methods and Regional Numerical Models (2024G-3)). Jun Wang's participation is made possible via the in-kind support from the University of Iowa.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.atmosenv.2026.122098.

Data availability

Data will be made available on request.

References

Ackermann, I.J., Hass, H., Memmesheimer, M., Ebel, A., Binkowski, F.S., Shankar, U., 1998. Modal aerosol dynamics model for Europe: development and first applications. Atmos. Environ. 32, 2981–2999. https://doi.org/10.1016/S1352-2310(98)00006-5.

Araki, R., Mu, Y., McMillan, H., 2023. Evaluation of GLDAS soil moisture seasonality in arid climates. Hydrol. Sci. J. 68, 1109–1126. https://doi.org/10.1080/02626667.2023.2206032.

Aryal, Y., 2022. Evaluation of machine-learning models for predicting aeolian dust: a case study over the southwestern USA. Climate 10, 78. https://doi.org/10.3390/cli10060078.

Banks, J.R., Heinold, B., Schepanski, K., 2024. Dust aerosol from the Aralkum Desert influences the radiation budget and atmospheric dynamics of central Asia. Atmos. Chem. Phys. 24, 11451–11475. https://doi.org/10.5194/acp-24-11451-2024.

Belgiu, M., Drăguţ, L., 2016. Random forest in remote sensing: a review of applications and future directions. ISPRS. J. Photogramm. 114, 24–31. https://doi.org/10.1016/j.isprsjprs.2016.01.011.

Bi, H., Ma, J., Zheng, W., Zeng, J., 2016. Comparison of soil moisture in GLDAS model simulations and in situ observations over the Tibetan Plateau. J. Geophys. Res. Atmos. 121, 2658–2678. https://doi.org/10.1002/2015JD024131.

Binkowski, F.S., Shankar, U., 1995. The regional particulate matter model: 1. Model description and preliminary results. J. Geophys. Res. Atmos. 100, 26191–26209. https://doi.org/10.1029/95JD02093.

Breiman, L., 2001. Random forests. Mach. Learn. 45, 5–32. https://doi.org/10.1023/A:1010933404324.

Chelhaoui, Y., El Ass, K., Khomsi, K., El Moussaoui, T., Chelhaoui, K., 2025. Evaluation of three chemical transport models for extreme PM_{10} events in Morocco and improvement of forecasts using hybrid CTM–random forest models. Sci. Total Environ. 1003, 180668. https://doi.org/10.1016/j.scitotenv.2025.180668.

Chen, G., Li, S., Knibbs, L.D., Hamm, N.A., Cao, W., Li, T., Guo, J., Ren, H., Abramson, M.J., Guo, Y., 2018. A machine learning method to estimate $PM_{2.5}$ concentrations across China with remote sensing, meteorological and land use information. Sci. Total Environ. 636, 52–60. https://doi.org/10.1016/j.scitotenv.2018.04.251.

Chen, W., Ran, H., Cao, X., Wang, J., Teng, D., Chen, J., Zheng, X., 2020. Estimating $PM_{2.5}$ with high-resolution 1-km AOD data and an improved machine learning model over Shenzhen. China. Sci. Total Environ. 746, 141093. https://doi.org/10.1016/j.scitotenv.2020.141093.

Chen, Y., Chen, S., Bi, H., Zhou, J., Zhang, Y., 2024. Where is the dust source of 2023 several severe dust events in China? Bull. Am. Meteorol. Soc. 105, E2085–E2096. https://doi.org/10.1175/BAMS-D-23-0121.1.

Chen, T., Guestrin, C., 2016. Xgboost: a scalable tree boosting system. Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining 785–794. https://doi.org/10.1145/2939672.2939785.

Darvishi Boloorani, A., Najafi, M.S., Mirzaie, S., 2021. Role of land surface parameter change in dust emission and impacts of dust on climate in Southwest Asia. Nat. Hazards 109, 111–132. https://doi.org/10.1007/s11069-021-04828-0.

Eltahan, M., Shokr, M., Sherif, A.O., 2018. Simulation of severe dust events over Egypt using tuned dust schemes in weather research forecast (WRF-Chem). Atmosphere 9, 246. https://doi.org/10.3390/atmos9070246.

Fécan, F., Marticorena, B., Bergametti, G., 1998. Parametrization of the increase of the aeolian erosion threshold wind friction velocity due to soil moisture for arid and semi-arid areas. Ann. Geophys. 17, 149–157. https://doi.org/10.1007/s00585-999-0149-7.

Filonchyk, M., 2022. Characteristics of the severe March 2021 gobi desert dust storm and its impact on air pollution in China. Chemosphere 287, 132219. https://doi.org/10.1016/j.chemosphere.2021.132219.

Fussell, J.C., Kelly, F.J., 2021. Mechanisms underlying the health effects of desert sand dust. Environ. Int. 157, 106790. https://doi.org/10.1016/j.envint.2021.106790.

Galán-Madruga, D., Broomandi, P., Satyanaga, A., Jahanbakhshi, A., Bagheri, M., Fathian, A., Sarvestan, R., Cárdenas-Escudero, J., Cáceres, J.O., Kumar, P., 2025. A methodological framework for estimating ambient $PM_{2.5}$ particulate matter concentrations in the UK. J. Environ. Sci. 150, 676–691. https://doi.org/10.1016/j.jes.2023.11.019.

Gelaro, R., McCarty, W., Suárez, M.J., Todling, R., Molod, A., Takacs, L., Randles, C.A., Darmenov, A., Bosilovich, M.G., Reichle, R., 2017. The modern-era retrospective analysis for research and applications, version 2 (MERRA-2). J. Clim. 30, 5419–5454. https://doi.org/10.1175/JCLI-D-16-0758.1.

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

Grell, G.A., Freitas, S.R., 2013. A scale and aerosol aware stochastic convective parameterization for weather and air quality modeling. Atmos. Chem. Phys. 14, 5233–5250. https://doi.org/10.5194/acp-14-5233-2014.

Guenther, A., Jiang, X., Heald, C.L., Sakulyanontvittaya, T., Duhl, Emmons, L., Wang, X., 2012. The model of emissions of gases and aerosols from nature version 2.1 (MEGAN2. 1): an extended and updated framework for modeling biogenic emissions. Geosci. Model Dev. 5, 1471–1492. https://doi.org/10.5194/gmd-5-1471-2012.

Gui, K., Yao, W., Che, H., An, L., Zheng, Y., Li, L., Zhao, H., Zhang, L., Zhong, J., Wang, Y., 2022. Record-breaking dust loading during two mega dust storm events over northern China in March 2021: aerosol optical and radiative properties and meteorological drivers. Atmos. Chem. Phys. 22, 7905–7932. https://doi.org/10.5194/acp-22-7905-2022.

Han, T., Pan, X., Wang, X., 2021. Evaluating and improving the sand storm numerical simulation performance in Northwestern China using WRF-Chem and remote sensing soil moisture data. Atmos. Res. 251, 105411. https://doi.org/10.1016/j.atmosres.2020.105411.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., 2020. The ERA5 global reanalysis. Q. J. Roy. Meteorol. Soc. 146, 1999–2049. https://doi.org/10.1002/qj.3803.

Hoffmann, L., Günther, G., Li, D., Stein, O., Wu, X., Griessbach, S., Heng, Y., Konopka, P., Müller, R., Vogel, B., 2019. From ERA-interim to ERA5: the considerable impact of ECMWF's next-generation reanalysis on Lagrangian transport simulations. Atmos. Chem. Phys. 19, 3097–3124. https://doi.org/10.5194/acp-19-3097-2019.

Hong, S.-Y., Noh, Y., Dudhia, J., 2006. A new vertical diffusion package with an explicit treatment of entrainment processes. Mon. Weather Rev. 134, 2318–2341. https://doi.org/10.1175/MWR3199.1.

Hsu, N.C., Tsay, S.-C., King, M.D., Herman, J.R., 2004. Aerosol properties over bright-reflecting source regions. IEEE. T. Geosci. Remote. 42, 557–569. https://doi.org/10.1109/TGRS.2004.824067.

Huang, J., Wang, T., Wang, W., Li, Z., Yan, H., 2014. Climate effects of dust aerosols over East Asian arid and semiarid regions. J. Geophys. Res. Atmos. 119 (11). https://doi.org/10.1002/2014JD021796, 398-311,416.

Iacono, M.J., Mlawer, E.J., Clough, S.A., Morcrette, J.J., 2000. Impact of an improved longwave radiation model, RRTM, on the energy budget and thermodynamic properties of the NCAR community climate model, CCM3. J. Geophys. Res. Atmos. 105, 14873–14890. https://doi.org/10.1029/2000JD900091.

Iiasa, F., 2023. Harmonized world soil database version 2.0. rome and laxenburg. https://doi.org/10.4060/cc3823en.

Jin, J., Lin, H.X., Heemink, A., Segers, A., 2018. Spatially varying parameter estimation for dust emissions using reduced-tangent-linearization 4DVar. Atmos. Environ. 187, 358–373. https://doi.org/10.1016/j.atmosenv.2018.05.060.

Jin, J., Lin, H.X., Segers, A., Xie, Y., Heemink, A., 2019. Machine learning for observation bias correction with application to dust storm data assimilation. Atmos. Chem. Phys. 19, 10009–10026. https://doi.org/10.5194/acp-19-10009-2019.

Kang, J.Y., Yoon, S.C., Shao, Y., Kim, S.W., 2011. Comparison of vertical dust flux by implementing three dust emission schemes in WRF. Chem. J. Geophys. Res. Atmos. 116. https://doi.org/10.1029/2010JD014649.

Kazemi, M., Rezaei, M., Mousaei, S., Kariminejad, N., 2026. Modeling dust storms using machine learning and deep learning techniques. J. Arid Environ. 233, 105542. https://doi.org/10.1016/j.jaridenv.2025.105542.

Kim, H.S., Park, I., Song, C.H., Lee, K., Yun, J.W., Kim, H.K., Jeon, M., Lee, J., Han, K.M., 2019. Development of a daily PM_{10} and $PM_{2.5}$ prediction system using a deep long short-term memory neural network model. Atmos. Chem. Phys. 19, 12935–12951. https://doi.org/10.5194/acp-19-12935-2019.

Kim, S., Lee, J.-H., Lee, S.-H., 2024. Significant impacts of GLDAS soil moisture data on severe Asian dust storm simulations by the WRF-Chem model. AGU Fall Meeting Abstracts A43J–2126.

Lee, J., Shi, Y.R., Cai, C., Ciren, P., Wang, J., Gangopadhyay, A., Zhang, Z., 2021. Machine learning based algorithms for global dust aerosol detection from satellite images: inter-comparisons and evaluation. Remote Sens. 13, 456. https://doi.org/10.3390/rs13030456.

Lee, J.A., Jimenez, P.A., Kumar, R., He, C., 2024. Impact of direct insertion of SMAP soil moisture retrievals in WRF-Chem for dust storm events in the western U.S. Atmos. Environ. 321. https://doi.org/10.1016/j.atmosenv.2024.120349.

Lee, S., Park, S., Lee, M.I., Kim, G., Im, J., Song, C.K., 2022. Air quality forecasts improved by combining data assimilation and machine learning with satellite AOD. Geophys. Res. Lett. 49. https://doi.org/10.1029/2021GL096066 e2021GL096066.

Lee, J.-H., Lee, S.-H., 2022. Modeling a severe wintertime Asian dust event observed in the east Asia region: sensitivity of the WRF-chem dust emission schemes. Atmos. Pollut. Res. 13, 101599. https://doi.org/10.1016/j.apr.2022.101599.

LeGrand, S.L., Polashenski, C., Letcher, T.W., Creighton, G.A., Peckham, S.E., Cetola, J. D., 2019. The AFWA dust emission scheme for the GOCART aerosol model in WRF-Chem v3. 8.1. Geosci. Model Dev. 12, 131–166. https://doi.org/10.5194/gmd-12-131-2019.

Letcher, T.W., LeGrand, S.L., 2018. A comparison of simulated dust produced by three dust-emission schemes in WRF-Chem: case Study assessment. https://doi.org/10.21079/11681/28868.

Li, C., Wang, J., Zhang, H., Diner, D.J., Hasheminassab, S., Janechek, N., 2024a. Improvement of surface $PM_{2.5}$ diurnal variation simulations in East Africa for the MAIA satellite mission. ACS Es&t Air 1, 223–233. https://doi.org/10.1021/acsestair.3c00008.

Li, H., Wang, C., Wang, M., Liu, Z., Mamtimin, A., Pan, X., 2023a. A new dataset of erodibility in dust source for WRF-chem model based on remote sensing and soil texture-Application and validation. Atmos. Environ. 315, 120156. https://doi.org/10.1016/j.atmosenv.2023.120156.

Li, H., Zhang, M., Wang, L., Lu, Y., Yu, L., Ma, Y., Gong, W., 2024b. Effects of aerosol on downward diffuse radiation under blowing dust and haze conditions. Atmos. Environ. 334, 120682. https://doi.org/10.1016/j.atmosenv.2024.120682.

Li, J., Hao, X., Liao, H., Yue, X., Li, H., Long, X., Li, N., 2022a. Predominant type of dust storms that influences air quality over northern China and future projections. Earths Future 10. https://doi.org/10.1029/2022EF002649 e2022EF002649.

Li, J., Jang, J.-c., Zhu, Y., Lin, C.-J., Wang, S., Xing, J., Dong, X., Li, J., Zhao, B., Zhang, B., 2023b. Development of a recurrent spatiotemporal deep-learning method coupled with data fusion for correction of hourly ozone forecasts. Environ. Pollut. 335, 122291. https://doi.org/10.1016/j.envpol.2023.122291.

Li, L., Mahowald, N.M., Kok, J.F., Liu, X., Wu, M., Leung, D.M., Hamilton, D.S., Emmons, L.K., Huang, Y., Sexton, N., 2022b. Importance of different parameterization changes for the updated dust cycle modeling in the community atmosphere model (version 6.1). Geosci. Model Dev. 15, 8181–8219. https://doi.org/10.5194/gmd-15-8181-2022.

Li, T., Zhang, Q., Peng, Y., Guan, X., Li, L., Mu, J., Wang, X., Yin, X., Wang, Q., 2023c. Contributions of various driving factors to air pollution events: interpretability analysis from machine learning perspective. Environ. Int. 173, 107861. https://doi.org/10.1016/j.envint.2023.107861.

Lian, L., Huang, J., Chen, S., Du, S., Zhang, L., Yang, J., 2025. A comprehensive review of dust events: characteristics, climate feedbacks, and public health risks. Curr. Pollut. Rep. 11, 18. https://doi.org/10.1007/s40726-025-00347-9.

Liang, P., Chen, B., Yang, X., Liu, Q., Li, A., Mackenzie, L., Zhang, D., 2022. Revealing the dust transport processes of the 2021 mega dust storm event in northern China. Sci. Bull. 67, 21–24. https://doi.org/10.1016/j.scib.2021.08.014.

Lin, C., Wang, Z., Zhu, J., 2008. An ensemble kalman filter for severe dust storm data assimilation over China. Atmos. Chem. Phys. 8, 2975–2983. https://doi.org/10.5194/acp-8-2975-2008.

Lin, Y.-L., Farley, R.D., Orville, H.D., 1983. Bulk parameterization of the snow field in a cloud model. J. Appl. Meteorol. Climatol. 22, 1065–1092. https://doi.org/10.1175/1520-0450(1983)022<1065:BPOTSF>2.0.CO;2.

Liu, J., Xing, J., 2023. Identifying contributors to $PM_{2.5}$ simulation biases of chemical transport model using fully connected neural networks. J. Adv. Model. Earth Syst. 15. https://doi.org/10.1029/2021MS002898 e2021MS002898.

Lwin, K.S., Tobias, A., Chua, P.L., Yuan, L., Thawonmas, R., Ith, S., Htay, Z.W., Yu, L.S., Yamasaki, L., Roqué, M., 2023. Effects of desert dust and sandstorms on human health: a scoping review. GeoHealth 7. https://doi.org/10.1029/2022GH000728 e2022GH000728.

Lyu, B., Huang, R., Wang, X., Wang, W., Hu, Y., 2025. FastCTM (v1. 0): atmospheric chemical transport modelling with a principle-informed neural network for air quality simulations. Geosci. Model Dev. 18, 6295–6312. https://doi.org/10.5194/gmd-18-6295-2025.

Ma, C., Song, J., Ran, M., Wan, Z., Guo, Y., Gao, M., 2024. Machine learning-driven spatiotemporal analysis of ozone exposure and health risks in China. J. Geophys. Res. Atmos. 129. https://doi.org/10.1029/2024JD041593 e2024JD041593.

Ma, R., Liu, N., Xu, X., Wang, Y., Noh, H.Y., Zhang, P., Zhang, L., 2020. Fine-grained air pollution inference with mobile sensing systems: a weather-related deep autoencoder model. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies 4, 1–21. https://doi.org/10.1145/3397322.

Marticorena, B., Bergametti, G., 1995. Modeling the atmospheric dust cycle: 1. Design of a soil-derived dust emission scheme. J. Geophys. Res. Atmos. 100, 16415–16430. https://doi.org/10.1029/95JD00690.

Mei, S., You, W., Zhong, W., Zang, Z., Guo, J., Xiang, Q., 2024. Optimizing the numerical simulation of the dust event of March 2021: integrating aerosol observations through multi-scale 3D variational assimilation in the WRF-chem model. Remote Sens. 16, 1852. https://doi.org/10.3390/rs16111852.

Meng, L., Zhao, T., He, Q., Yang, X., Mamtimin, A., Wang, M., Pan, H., Huo, W., Yang, F., Zhou, C., 2022. Dust radiative effect characteristics during a typical springtime dust storm with persistent floating dust in the tarim basin, northwest China. Remote Sens. 14, 1167. https://doi.org/10.3390/rs14051167.

Mesbahzadeh, T., Salajeghe, A., Sardoo, F.S., Zehtabian, G., Ranjbar, A., Marcello Miglietta, M., Karami, S., Krakauer, N.Y., 2020. Spatial-temporal variation characteristics of vertical dust flux simulated by WRF-chem model with GOCART and AFWA dust emission schemes (case study: central plateau of Iran). Appl. Sci. 10, 4536. https://doi.org/10.3390/app10134536.

Mlawer, E.J., Taubman, S.J., Brown, P.D., Iacono, M.J., Clough, S.A., 1997. Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave. J. Geophys. Res. Atmos. 102, 16663–16682. https://doi.org/10.1029/97JD00237.

Mukul Tewari, N., Tewari, M., Chen, F., Wang, W., Dudhia, J., LeMone, M., Mitchell, K., Ek, M., Gayno, G., Wegiel, J., 2004. Implementation and verification of the unified NOAH land surface model in the WRF model (formerly paper number 17.5). In: Proceedings of the 20th Conference on Weather Analysis and forecasting/16th Conference on Numerical Weather Prediction, vol. 14, pp. 11–15. Seattle, WA, USA.

Nababan, A.A., Zarlis, M., Nababan, E.B., 2022. Air quality prediction based on air pollution emissions in the city environment using xgboost with smote. In: 2022 IEEE International Conference of Computer Science and Information Technology (ICOSNIKOM), pp. 1–6. https://doi.org/10.1109/ICOSNIKOM56551.2022.10034887.

Nachtergaele, F., Van Velthuizen, H., Verelst, L., Batjes, N., Dijkshoorn, J., Van Engelen, V., Fischer, G., Jones, A., Montanarella, L., Petri, M., 2008. Harmonized World Soil Database. FAO version 1.0.

Nguyen, H.A., Ha, Q.P., Duc, H., Azzi, M., Jiang, N., Barthelemy, X., Riley, M., 2023. Long short-term memory Bayesian neural network for air pollution forecast. IEEE Access 11, 35710–35725. https://doi.org/10.1109/ACCESS.2023.3265725.

Pahlow, M., Parlange, M.B., Porté-Agel, F., 2001. On Monin–Obukhov similarity in the stable atmospheric boundary layer. Bound-Lay. Meteorol 99, 225–248. https://doi.org/10.1023/A:1018909000098.

Raman, A., Arellano, A.F., Delle Monache, L., Alessandrini, S., Kumar, R., 2021. Exploring analog-based schemes for aerosol optical depth forecasting with WRF-Chem. Atmos. Environ. 246, 118134. https://doi.org/10.1016/j.atmosenv.2020.118134.

Rigatti, S.J., 2017. Random forest. J. Insur. Med. 47, 31–39. https://doi.org/10.17849/insm-47-01-31-39.1.

Rizza, U., Barnaba, F., Miglietta, M.M., Mangia, C., Di Liberto, L., Dionisi, D., Costabile, F., Grasso, F., Gobbi, G.P., 2017. WRF-Chem model simulations of a dust outbreak over the central mediterranean and comparison with multi-sensor desert dust observations. Atmos. Chem. Phys. 17, 93–115. https://doi.org/10.5194/acp-17-93-2017.

Rodell, M., Houser, P., Jambor, U., Gottschalck, J., Mitchell, K., Meng, C., Arsenault, K., Cosgrove, B., Radakovich, J., Bosilovich, M., 2004. The global land data assimilation system. Bull. Am. Meteorol. Soc. 84, 381–394. https://doi.org/10.1175/BAMS-85-3-381.

Schell, B., Ackermann, I.J., Hass, H., Binkowski, F.S., Ebel, A., 2001. Modeling the formation of secondary organic aerosol within a comprehensive air quality model system. J. Geophys. Res. Atmos. 106, 28275–28293. https://doi.org/10.1029/2001JD000384.

Sha, T., Ma, X., Zhang, H., Janechek, N., Wang, Y., Wang, Y., Castro García, L., Jenerette, G.D., Wang, J., 2021. Impacts of soil NO_x emission on O_3 air quality in rural California. Environ. Sci. Technol. 55, 7113–7122. https://doi.org/10.1021/acs.est.0c06834.

Sha, T., Yang, S., Chen, Q., Li, L., Ma, X., Zhang, Y.-L., Feng, Z., Boersma, K.F., Wang, J., 2024. Large contributions of soil emissions to the atmospheric nitrogen budget and their impacts on air quality and temperature rise in North China. Atmos. Chem. Phys. 24, 8441–8455. https://doi.org/10.5194/acp-24-8441-2024.

Sha, T., Yang, S., Chen, Q., Wei, J., Ma, M., Gao, Y., Zhu, Y., Hu, Y., Boersma, K.F., Wang, J., 2025. Elucidating contributions of anthropogenic and soil NO_x emissions changes to O_3 trends over China. J. Geophys. Res. Atmos. 130. https://doi.org/10.1029/2025JD044623 e2025JD044623.

Shao, Y., 2024. Adhesion theory and model for air humidity impact on dust emission. Aeolian Res. 66. https://doi.org/10.1016/j.aeolia.2024.100898.

Shi, S., Chen, R., Wang, P., Zhang, H., Kan, H., Meng, X., 2024. An ensemble machine learning model to enhance extrapolation ability of predicting coarse particulate matter with high resolutions in China. Environ. Sci. Technol. 58, 19325–19337. https://doi.org/10.1021/acs.est.4c08610.

Singh, A., Islam, M., Dinh, N., 2024. Forecasting indoor air quality using machine learning models. In: 2024 IEEE International Conference on Consumer Electronics (ICCE), pp. 1–6. https://doi.org/10.1109/ICCE59016.2024.10444387.

Srivastava, N., Blond, N., 2022. Impact of meteorological parameterization schemes on CTM model simulations. Atmos. Environ. 268, 118832. https://doi.org/10.1016/j.atmosenv.2021.118832.

Stockwell, W.R., Middleton, P., Chang, J.S., Tang, X., 1990. The second generation regional acid deposition model chemical mechanism for regional air quality modeling. J. Geophys. Res. Atmos. 95, 16343–16367. https://doi.org/10.1029/JD095iD10p16343.

Su, L., Fung, J.C., 2015. Sensitivities of WRF-Chem to dust emission schemes and land surface properties in simulating dust cycles during springtime over East Asia. J. Geophys. Res. Atmos. 120 (11). https://doi.org/10.1002/2015JD023446, 215-211,230.

Sun, H., Shin, Y.M., Xia, M., Ke, S., Wan, M., Yuan, L., Guo, Y., Archibald, A.T., 2021. Spatial resolved surface ozone with urban and rural differentiation during 1990–2019: a space–time Bayesian neural network downscaler. Environ. Sci. Technol. 56, 7337–7349. https://doi.org/10.1021/acs.est.1c04797.

Tang, B., Stanier, C.O., Carmichael, G.R., Gao, M., 2024. Ozone, nitrogen dioxide, and $PM_{2.5}$ estimation from observation-model machine learning fusion over S. Korea: influence of observation density, chemical transport model resolution, and geostationary remotely sensed AOD. Atmos. Environ. 331, 120603. https://doi.org/10.1016/j.atmosenv.2024.120603.

Udristioiu, M.T., Mghouchi, Y.E., Yildizhan, H., 2023. Prediction, modelling, and forecasting of PM and AQI using hybrid machine learning. J. Clean. Prod. 421, 138496. https://doi.org/10.1016/j.jclepro.2023.138496.

Ukhov, A., Ahmadov, R., Grell, G., Stenchikov, G., 2021. Improving dust simulations in WRF-Chem v4. 1.3 coupled with the GOCART aerosol module. Geosci. Model Dev. 14, 473–493. https://doi.org/10.5194/gmd-14-473-2021.

Wang, K., Zhang, Y., Yahya, K., Wu, S.-Y., Grell, G., 2015. Implementation and initial application of new chemistry-aerosol options in WRF/Chem for simulating secondary organic aerosols and aerosol indirect effects for regional air quality. Atmos. Environ. 115, 716–732. https://doi.org/10.1016/j.atmosenv.2014.12.007.

Wang, L., Meng, R., Cao, L., Gao, F., Chen, H., Li, Y., Zhang, J., Yan, D., Liu, X., Niu, T., 2022. Influence of dust events on air quality of Shaanxi province from 2016 to 2020. J. Desert Res. 42, 130. https://doi.org/10.7522/j.issn.1000-694X.2022.000064.

Wang, S., Zhang, M., Gao, Y., Wang, P., Fu, Q., Zhang, H., 2024. Diagnosing drivers of $PM_{2.5}$ simulation biases in China from meteorology, chemical composition, and emission sources using an efficient machine learning method. Geosci. Model Dev. 17, 3617–3629. https://doi.org/10.5194/gmd-17-3617-2024.

Weng, X., Forster, G.L., Nowack, P., 2022. A machine learning approach to quantify meteorological drivers of ozone pollution in China from 2015 to 2019. Atmos. Chem. Phys. 22, 8385–8402. https://doi.org/10.5194/acp-22-8385-2022.

Xiao, Q., Geng, G., Liu, S., Liu, J., Meng, X., Zhang, Q., 2022. Spatiotemporal continuous estimates of daily 1 km $PM_{2.5}$ from 2000 to present under the Tracking air pollution in China (TAP) framework. Atmos. Chem. Phys. 22, 13229–13242. https://doi.org/10.5194/acp-22-13229-2022.

Xu, L., Batterman, S., Chen, F., Li, J., Zhong, X., Feng, Y., Rao, Q., Chen, F., 2017. Spatiotemporal characteristics of $PM_{2.5}$ and PM_{10} at urban and corresponding background sites in 23 cities in China. Sci. Total Environ. 599, 2074–2084. https://doi.org/10.1016/j.scitotenv.2017.05.048.

Xu, M., Jin, J., Wang, G., Segers, A., Deng, T., Lin, H.X., 2021. Machine learning based bias correction for numerical chemical transport models. Atmos. Environ. 248, 118022. https://doi.org/10.1016/j.atmosenv.2020.118022.

Ye, X., Wang, X., Zhang, L., 2022. Diagnosing the model bias in simulating daily surface ozone variability using a machine learning method: the effects of dry deposition and cloud optical depth. Environ. Sci. Technol. 56, 16665–16675. https://doi.org/10.1021/acs.est.2c05712.

Yi, Z., Zeng, Z., Wang, Y., Li, W., Zhang, B., Gui, H., Guo, B., Chen, W., Che, H., Zhang, X., 2025. Improving short-term forecasting of surface dust concentration in northern China: integrating machine learning with multi-numerical models. Atmos. Pollut. Res. 16, 102480. https://doi.org/10.1016/j.apr.2025.102480.

Yoon, J.W., Lee, E., Park, S.K., 2024. Improving the Asian dust storm prediction using WRF-Chem through combinational optimization of physical parameterization schemes. Atmos. Environ. 326, 120461. https://doi.org/10.1016/j.atmosenv.2024.120461.

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

Zhang, H., Wang, J., Janechek, N., Ge, C., Zhou, M., García, L.C., Sha, T., Wang, Y., Deng, W., Xue, Z., 2026. Development of UI-WRF-Chem (v1. 0) for the MAIA satellite mission: case demonstration. Geosci. Model Dev. https://doi.org/10.5194/egusphere-2025-1360.

Zhang, J., Zhou, C.-H., Gui, H.-L., Zhang, X.-Y., 2021a. Biological crust in sand and dust storm source areas of Asia and its impact on dust emission. Adv. Clim. Change Res. 12, 395–408. https://doi.org/10.1016/j.accre.2021.05.007.

Zhang, S., Xu, H., Lan, J., Goldsmith, Y., Torfstein, A., Zhang, G., Zhang, J., Song, Y., Zhou, K., Tan, L., 2021b. Dust storms in northern China during the last 500 years. Sci. China Earth Sci. 64, 813–824. https://doi.org/10.1007/s11430-020-9730-2.

Zhang, X., Jiang, X., Li, Y., 2023. Prediction of air quality index based on the SSA-BiLSTM-LightGBM model. Sci. Rep. 13, 5550. https://doi.org/10.1038/s41598-023-32775-2.

Zhang, Y., Chen, Y., Chen, S., Alam, K., Chen, J., 2024. Mongolia dust transport across borders under the background of global warming. Global Planet. Change 239, 104509. https://doi.org/10.1016/j.gloplacha.2024.104509.

Zhao, X., Zhao, C., Yang, Y., Sun, Y., Xia, Y., 2022. Dust aerosol impacts on the time of cloud formation in the badain jaran desert area. J. Geophys. Res. Atmos. 127. https://doi.org/10.1029/2022JD037019 e2022JD037019.

Zhu, Q., Bi, J., Liu, X., Li, S., Wang, W., Zhao, Y., Liu, Y., 2022. Satellite-based long-term spatiotemporal patterns of surface ozone concentrations in China: 2005–2019. Environ. Health Perspect. 130, 027004. https://doi.org/10.1289/EHP9406.