Environmental Science: Atmospheres

ROYAL SOCIETY OF CHEMISTRY

PAPER

View Article Online
View Journal

Cite this: DOI: 10.1039/d5ea00167f

Toward reliable air quality simulations over India: optimizing WRF-Chem through comprehensive sensitivity analysis of meteorology, physical parameterizations, and emission inventories

Indranil Nandi * and Dilip Ganguly

Air pollution in India remains a severe environmental and public health challenge, intensified by meteorological conditions, particularly during winter and the post-monsoon season. This study conducts a systematic, multi-component sensitivity evaluation of the Weather Research and Forecasting model coupled with Chemistry (WRF-Chem) to identify an optimal configuration for simulating air quality over India. We assess the influence of meteorological initial and boundary conditions, physical and chemical parameterization schemes, emission inventories, and grid-nudging strengths through simulations for two contrasting periods, the monsoon season (August 2018) and winter (January 2019). Model performance is strongly configuration dependent, with ECMWF Reanalysis 5th Generation (ERA5) meteorology substantially improving the representation of surface temperature, relative humidity, wind speed, and vertical thermodynamic structure. Incorporating moderate grid nudging (0.0005 s^{-1}) reduced relative humidity biases by 13% during January 2019 and improved temperature inversion strength and planetary boundary layer (PBL) representation. The Yonsei University PBL (YSU) and Lin microphysics schemes produced the most realistic PBL heights, rainfall fields, and wet-deposition-driven pollutant removal. For atmospheric composition, the Model for Ozone And Related chemical Tracers coupled with Model for Simulating Aerosol Interactions and Chemistry (MOZART-MOSAIC) mechanism, combined with the high-resolution Speciated Multi-pollutant Generator – Coordinated Aerosol and Pollution Source Emissions for Chemical Evaluation (SMoG-COALESCE) emission inventory, yielded the most accurate simulations of $PM_{2.5}$, SO_2, O_3, and CO across urban and rural regions. The integrated optimal configuration improved model skill relative to Central Pollution Control Board (CPCB) surface measurements, reanalysis products, and satellite datasets. These results demonstrate the need for coordinated evaluation of meteorological and chemical components in WRF-Chem and provide a robust modelling framework for improving air-quality prediction and supporting policy-relevant assessments of pollution mitigation strategies over India.

Received 16th December 2025
Accepted 9th March 2026

DOI: 10.1039/d5ea00167f

rsc.li/esatmospheres

Environmental significance

Air-quality models are essential for understanding pollution and guiding clean-air policies in India, but their accuracy depends strongly on model configuration. This study reduces uncertainty in air-quality simulations by systematically testing different WRF-Chem settings, including weather inputs, atmospheric processes, chemical mechanisms, emissions, and meteorology data-nudging, across monsoon and winter seasons. Reliable modelling is crucial because air pollution in India causes serious health and environmental impacts. Our findings show that jointly improving meteorological and chemical components significantly enhances model performance. The optimized framework enables more accurate air-quality forecasting, pollution assessment, and evaluation of mitigation strategies, supporting effective environmental planning and public-health protection across India.

1 Introduction

Air pollution in India poses an enduring environmental and public health crisis, particularly in rapidly urbanizing and industrializing regions such as Delhi, Mumbai, and Kolkata, where growing vehicular activity, industrial emissions, construction dust, and widespread biomass burning contribute to persistently high pollutant levels.[1,2] Concentrations of particulate matter ($PM_{2.5}$ and PM_{10}), nitrogen oxides (NO_x), sulfur dioxide (SO_2), carbon monoxide (CO), and ground-level ozone (O_3) often exceed national and international standards, with substantial implications for human health, ecosystems, visibility, agricultural productivity, and the economy.[3] Although

Centre for Atmospheric Sciences, Indian Institute of Technology Delhi, India. E-mail: indranilnandi2011@gmail.com; Indranil.Nandi@cas.iitd.ac.in

urban centres experience the most severe pollution episodes, rural regions are also heavily affected by agricultural residue burning and household biomass fuel use, emphasizing the nationwide extent of the problem. Winter meteorology, marked by strong inversions, weak winds, and shallow boundary layers, frequently exacerbates pollution accumulation near the surface, leading to severe smog events and elevated health risks.[4–6]

The complexity of India's air pollution arises from the confluence of diverse emission sources and highly heterogeneous meteorological conditions. Numerical weather prediction (NWP) models, particularly the Weather Research and Forecasting model coupled with Chemistry (WRF-Chem), play a central role in simulating these coupled processes.[7] WRF-Chem resolves emissions, transport, chemical transformation, deposition, and meteorology within a unified modelling framework, enabling process-level understanding of pollutant behaviour. However, the accuracy of WRF-Chem simulations is highly sensitive to choices in meteorological initial and boundary conditions, physical and chemical parameterization schemes, and emission inventories.[8] These sensitivities are especially consequential in India, where monsoonal circulations, winter stagnation, pre-monsoon dust activity, and steep spatial gradients in terrain, from coastal regions to the Himalayas, create complex meteorological regimes that strongly modulate pollutant dispersion and transformation.[9]

Accurate meteorological simulation is particularly vital in India, where extreme events such as heatwaves, lightning, cold surges, and heavy precipitation are key drivers of socioeconomic impacts, with 25% of accidental deaths between 2001 and 2014 linked to such events.[5,10] WRF-Chem applications in the region have expanded rapidly, driven by its ability to represent monsoon dynamics, moisture fluxes, and convective processes that shape rainfall patterns and pollutant distribution.[11,12] The selection of planetary boundary layer (PBL) schemes, microphysics schemes, and radiation parameterizations critically influences simulations of temperature, wind speed, turbulence, and pollutant concentrations.[13] PBL schemes are especially important for resolving diurnal mixing and vertical transport of pollutants, with substantial impacts on surface ozone, PM, and other regulated species.[2,14–16]

Despite major advances, WRF-Chem performance over India continues to exhibit substantial uncertainties, driven by the region's meteorological diversity, complex chemistry, and emission heterogeneity.[17] Sensitivity studies have shown that improved model accuracy requires careful combination of parameterization schemes, high-quality meteorological inputs, and region-specific emissions.[18] Data assimilation approaches such as Four-Dimensional Data Assimilation (FDDA) or grid nudging can enhance large-scale meteorological realism by constraining model fields toward observations, thereby improving the representation of temperature gradients, wind profiles, and pollutant transport.[19] Prior studies have also demonstrated that microphysics choices significantly influence rainfall, wet deposition, and cloud–aerosol interactions,[20,21] while aerosol schemes such as MADE-SORGAM (Modal Aerosol Dynamics Model for Europe – Secondary Organic Aerosol Model)[22] and MOSAIC (Model for Simulating Aerosol Interactions and Chemistry)[23] can lead to divergent predictions of $PM_{2.5}$.[24] Boundary layer and land surface schemes similarly modulate surface fluxes and turbulence, affecting pollutant mixing depth and vertical structure.[25]

Although numerous studies have examined individual components of WRF-Chem have over India, such as boundary layer schemes, chemical mechanisms, or emission inventories, few have performed a comprehensive, multi-component sensitivity assessment spanning contrasting meteorological regimes. This gap limits our understanding of how different modelling choices interact to influence pollutant simulations across India's diverse atmospheric environment. To address this need, the present study systematically evaluates WRF-Chem sensitivity to meteorological and chemical configurations during two distinct periods: the monsoon season (August 2018), characterized by strong convection, high humidity, and efficient wet deposition; and the winter season (January 2019), dominated by stagnant conditions, shallow PBLs, and pollution buildup.[2,6] Through coordinated perturbations to meteorological initial and boundary conditions, PBL schemes, microphysics schemes, chemical mechanisms, emission inventories, and grid-nudging strengths, we identify an optimal WRF-Chem configuration capable of more accurately reproducing observed meteorological fields and air pollutant concentrations across India.

This work advances regional air-quality modelling by evaluating the combined effects of physical and chemical modelling choices rather than assessing individual components in isolation. The resulting optimized configuration has substantial implications for air-quality forecasting, long-term air pollution assessments, and policy design. Although the present analysis focuses on August (peak monsoon) and January (peak winter), we recognize that India also experiences distinct pre-monsoon and post-monsoon regimes. Pre-monsoon conditions are characterized by high temperatures, deep daytime boundary layers, and frequent dust transport from western arid regions, while post-monsoon conditions over northern India are often influenced by enhanced stagnation and episodic agricultural biomass burning. These regimes introduce additional sensitivities related to dust emission and transport, aerosol optical properties and mixing state, and the temporal allocation and vertical injection of biomass-burning emissions. A full extension of the present multi-component sensitivity matrix to these seasons would require a substantial additional set of simulations and is therefore beyond the scope of this work. Instead, we use the two most meteorologically contrasting regimes, peak monsoon and peak winter, to provide a stringent test of WRF-Chem meteorology–chemistry coupling and to identify a robust baseline configuration intended for broad air-quality applications across India. The paper is organized as follows: Section 2 describes the datasets, model configuration, sensitivity experiments, and validation methodology. Section 3 presents the results, identifying the optimal WRF-Chem configuration and discussing its strengths and limitations. Section 4 synthesizes key findings and proposes a modelling framework for improving air-quality simulations across India and other regions with complex meteorology and diverse emission sources. SI includes additional methodological

details, supplementary figures (Fig. S1–S14), and tables (*e.g.*, Tables S1–S5) supporting the main results.

2 Methodology

2.1 Model configuration

The Weather Research and Forecasting model coupled with Chemistry (WRF-Chem, v4.2.2)[26] was used to simulate meteorological and atmospheric composition over the Indian subcontinent. WRF-Chem integrates prognostic meteorological fields generated by the Weather Research and Forecasting (WRF) model with fully interactive gas-phase chemistry and aerosol processes, enabling the coupled simulation of emissions, transport, chemical transformation, and deposition of atmospheric pollutants within a single modelling system.[7] This fully coupled approach is particularly well suited for India, where strong interactions between complex emission patterns and highly variable meteorological conditions occur across seasons, and terrain types.

The model domain spans 60°E to 106°E and 1°N to 46°N, encompassing India and surrounding regions to account for long-range transport influences. The domain is centred over Bilaspur (22.38°N and 82.13°E), providing balanced spatial coverage of India's diverse climatic zones while maintaining computational efficiency. Simulations were conducted at a horizontal resolution of 27 × 27 km using a grid of 321 × 331 points, with 39 vertical levels extending from the surface to 50 hPa (Fig. S1). This vertical configuration provides enhanced resolution within the planetary boundary layer (PBL), where pollutant mixing and removal processes dominate, while also allowing realistic representation of free-tropospheric and upper-level dynamical features.[27,28]

Two simulation periods were selected to represent contrasting meteorological regimes. August 2018 corresponds to the monsoon season, characterized by deep convection, strong moisture transport, high humidity, and frequent wet deposition, whereas January 2019 represents winter conditions marked by weak winds, strong temperature inversions, shallow boundary layers, and severe pollutant accumulation. A 10 day spin-up period was applied prior to each simulation to minimize initialization effects, and spin-up output was excluded from the analyses.

To systematically assess model sensitivity and identify an optimal configuration for air-quality applications over India, a comprehensive suite of sensitivity experiments was conducted. Seven distinct sensitivity experiment groups were designed to isolate the influence of individual modelling components, including meteorological initial and boundary conditions, physical parameterization schemes, chemical mechanisms, emission inventories, and grid-nudging strategies. These groups collectively comprise a total of 18 WRF-Chem simulations, following a controlled, stepwise experimental design in which one model component is varied at a time while all other settings are held constant. This approach enables a clear attribution of model performance differences to specific configuration choices.

Specifically, the first set of simulations (control, CFSWRF, and GDASWRF) evaluated sensitivity to meteorological initial and boundary conditions. This was followed by experiments assessing planetary boundary layer and microphysics parameterizations (MYJWRF, MYNWRF, THOMWRF, and MDMWRF) to identify optimal representations of boundary-layer mixing and precipitation processes. Subsequently, Four-Dimensional Data Assimilation (FDDA) was introduced, and model performance was evaluated using two chemical mechanisms (CBMZWRF and Nudged03). Using the best-performing chemical configuration, four anthropogenic emission inventories (CMIPWRF, ED5WRF, ED6WRF, and HTAPWRF) were then tested for their ability to reproduce pollutant concentrations at the surface and 850 hPa. Finally, five additional simulations (Nudged05, Nudged1, NoNudPBL03, NoNudPBL05, and NoNudPBL1) were conducted to examine the sensitivity of model performance to grid-nudging strength and boundary-layer-specific nudging strategies. Details of all simulations, including their physical and chemical configurations, emission inventories, meteorological forcings, and nudging settings, are summarized in SI Table S1, which serves as a comprehensive reference for the experimental design and supports reproducibility of the study.

Unless otherwise stated, all sensitivity experiments are evaluated relative to a common baseline (control) configuration. The control simulation employs ERA5 meteorological initial and boundary conditions, the Yonsei University (YSU) planetary boundary layer scheme, Lin microphysics, the MOZART-MOSAIC chemical mechanism, the SMoG-COALESCE anthropogenic emission inventory, and no grid nudging. In each sensitivity experiment, only one model component is modified relative to this baseline, while all other physical, chemical, emission, and meteorological settings are held fixed.

2.2 Sensitivity experiments

The sensitivity experiments were designed to quantify how WRF-Chem responds to changes in key modelling components that collectively shape the simulation of meteorology, atmospheric chemistry, and pollutant distribution. These components include the selection of meteorological initial and boundary conditions, PBL schemes, microphysics schemes, chemical mechanisms, and emission inventories. Experiments were conducted for August 2018 and January 2019, enabling evaluation under monsoon and winter meteorological regimes, respectively. These seasons represent sharply contrasting atmospheric conditions, allowing a robust assessment of how different configurations influence the model's ability to reproduce observed meteorological variables and pollutant concentrations across India. For each sensitivity test, comparisons are conducted against the baseline configuration described in Section 2.1, ensuring that all reported differences arise solely from the targeted model component under investigation.

2.2.1 Meteorological initial and boundary conditions. Meteorological initial and boundary conditions exert a dominant influence on model performance because errors or biases introduced at the lateral and initial boundaries propagate

throughout the simulation, affecting cloud development, moisture transport, temperature gradients, wind fields, and ultimately pollutant transport and transformation.[29] To evaluate this influence, three widely used global meteorological datasets were tested.

ECMWF Reanalysis 5th Generation reanalysis (ERA5), with a horizontal resolution of 31 km, provides high-quality fields capable of resolving synoptic and mesoscale features critical for representing pollutant dispersion and meteorological variability across the Indian region.[30] Its higher resolution and improved representation of monsoonal dynamics, moisture fluxes, and temperature structure offer potential advantages for both monsoon and winter simulations.

The National Centers for Environmental Prediction – Global Data Assimilation System (NCEP-GDAS) Final Analysis (0.25° resolution) offers high-frequency updates that capture rapid temporal variability but is limited by its coarser horizontal resolution. This can weaken the representation of fine-scale meteorological features such as localized convergence, orographic lifting, or nocturnal low-level jets, which influence pollutant transport and dispersion.[31]

National Centers for Environmental Prediction Climate Forecast System Version 2 (NCEP CFSv2, 0.5° resolution) is widely used for climate applications, but its relatively coarse horizontal resolution and weaker representation of regional-scale features make it less suitable for resolving important processes such as shallow winter inversions, weak wind regimes, and humidity stratification, which are essential for pollution buildup and persistence.[32]

The performance of these datasets was assessed by comparing simulated surface temperature, relative humidity, rainfall, and wind fields against observations, satellite products, and reanalysis datasets, as shown in Fig. 1, S2–S4 and Table S2 within the SI.

2.2.2 Planetary boundary layer (PBL) schemes. The planetary boundary layer governs vertical exchange of heat, momentum, and atmospheric constituents, and thus strongly influences the dispersion, dilution, and near-surface accumulation of pollutants. Accurate representation of PBL processes is therefore critical for air-quality simulations, especially in regions such as northern India where shallow winter boundary layers exacerbate pollution episodes.[33] Three PBL schemes were examined to capture the diversity of boundary-layer conditions characteristic of monsoon and winter seasons.

The Yonsei University PBL (YSU),[34] a nonlocal first-order closure scheme, is known for effectively representing convective boundary-layer dynamics and has demonstrated strong performance over Indian conditions during both monsoon and winter. Its ability to capture entrainment processes and the diurnal evolution of the PBL makes it a strong candidate for simulating pollutant mixing under varying atmospheric stability. The Mellor–Yamada–Janjic PBL scheme (MYJ)[35] employs a local 1.5-order turbulence kinetic energy closure and is better suited to representing stable boundary layers. This makes it particularly relevant for winter conditions over regions such as the Indo-Gangetic Plain, where persistent temperature inversions and weak winds trap pollutants near the surface. The Mellor–Yamada–Nakanishi–Niino Level 3 PBL scheme (MYNN3)[36] offers an advanced turbulence parameterization designed for complex terrain and urban environments. It attempts to capture a broader spectrum of turbulent processes and may, in principle, improve pollutant dispersion simulations in heterogeneous landscapes.

These three schemes were evaluated using PBL height (PBLH), vertical temperature profiles, and surface pollutant concentrations. Fig. 2, S5 and Table S2 provide detailed comparisons illustrating how each scheme influences pollutant accumulation and vertical mixing across regions and seasons.

2.2.3 Microphysics schemes. Microphysics schemes determine the representation of cloud formation, precipitation processes, and wet deposition, all of which exert direct control over atmospheric composition. This influence is especially pronounced during the monsoon, when rainfall is the dominant removal mechanism for aerosol and gas-phase pollutants.[37] To evaluate these processes, three microphysics schemes were examined.

The Lin scheme[38] is a widely used single-moment scheme that offers computational efficiency while providing reasonable performance for regional-scale meteorological and air-quality applications. Its representation of cloud microphysical processes and precipitation formation is often suitable for capturing large-scale monsoon rainfall patterns. The Thompson scheme[39] is a more advanced double-moment formulation that improves the representation of cloud–aerosol interactions by predicting both mass and number concentrations for hydrometeors. Its enhanced microphysical detail may improve the simulation of cloud fraction, liquid water path, and aerosol–cloud interactions. The Morrison scheme,[40] also a double-moment scheme, is designed to capture a broad spectrum of hydrometeor species and is well suited for convective environments. Its ability to represent convective microphysical processes may contribute to improved precipitation fields and enhanced removal of pollutants.

Rainfall distributions, cloud fraction fields, and wet deposition estimate from these schemes were compared using India Meteorological Department (IMD) rainfall observations, Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA2) datasets, and satellite measurements. Fig. S6, S7 and Table S3 summarize their performance during the simulation periods.

2.2.4 Chemical mechanisms. Chemical mechanisms govern the formation, transformation, and removal of gas-phase species and aerosols, including secondary pollutants such as ozone, sulfate, nitrate, and secondary organic aerosols.[41] Two mechanisms were evaluated to assess their suitability for air-quality simulations over India's complex emission and chemical environments.

The Carbon Bond Mechanism version Z – Model for Simulating Aerosol Interactions and Chemistry (CBMZ-MOSAIC) mechanism[42] is computationally efficient and widely implemented in regional air-quality studies. However, its simplified representation of secondary chemistry may limit its ability to capture detailed oxidant chemistry and multiphase processes, especially under low-photolysis winter conditions or in regions

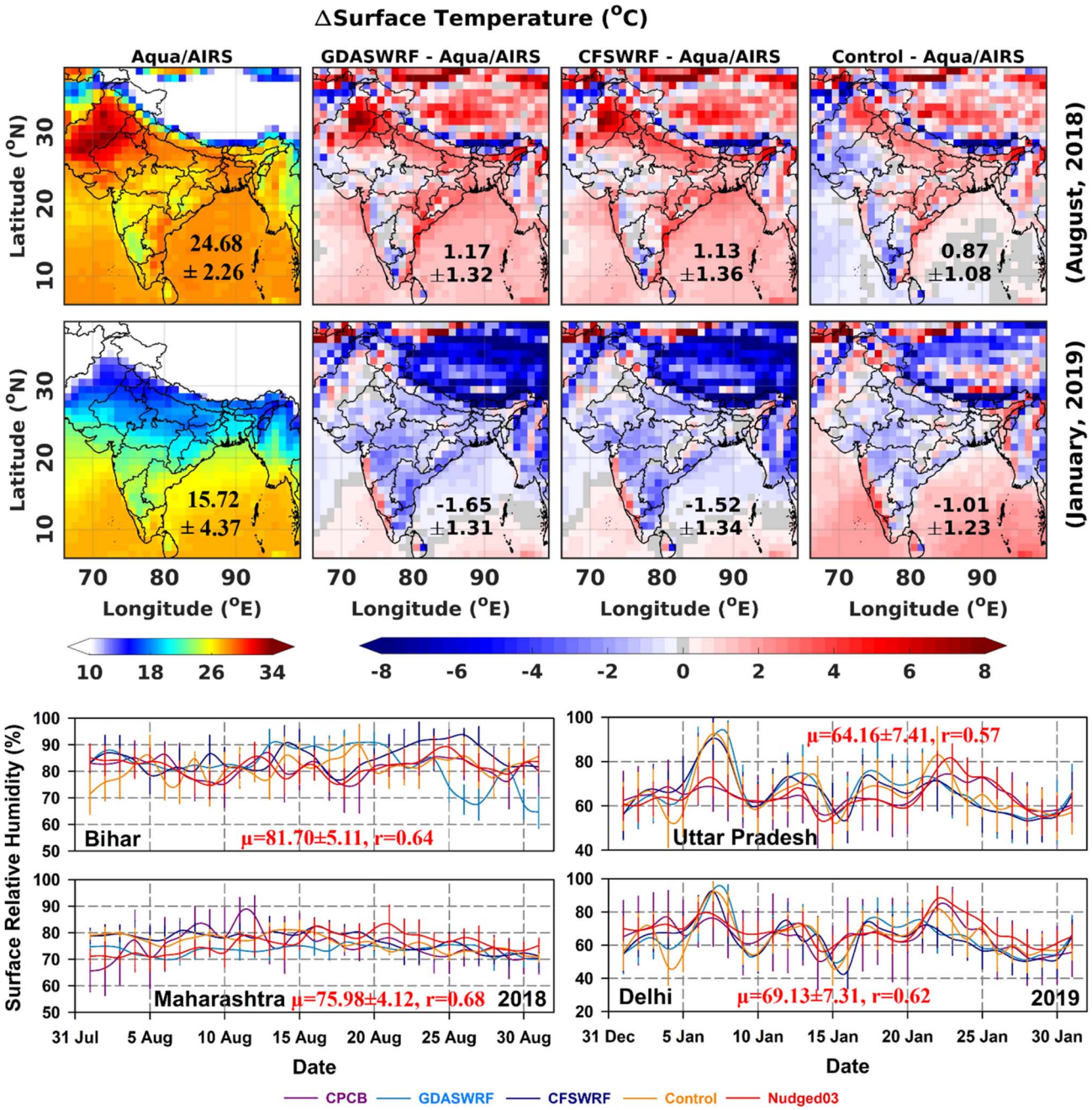

Fig. 1 Spatial distribution of surface temperature observed by Aqua/AIRS over the Indian region, and the difference in simulated temperature from various WRF-Chem meteorology sensitivity simulations (top panel). Inset values report the domain-mean ± spatial standard deviation ($\mu \pm 1\sigma$) over the Indian landmass, along with Mean Bias Error (MBE) relative to Aqua/AIRS. Daily mean time series comparison of surface relative humidity for various Indian states from different WRF-Chem meteorology and grid nudging sensitivity simulations with CPCB measurements and MERRA2 (bottom panel). Temporal mean ± SD ($\mu \pm 1\sigma$) and correlation coefficient (r) of Nudged03 indicated within the panel.

with complex Volatile Organic Compounds (VOC) and Nitrogen Oxides (NO_x) mixtures. The Model for OZone And Related chemical Tracers – Model for Simulating Aerosol Interactions and Chemistry (MOZART-MOSAIC) mechanism[43] includes more comprehensive representations of ozone–NO_x–VOC chemistry, aerosol thermodynamics, heterogeneous reactions, and the Volatility Basis Set (VBS) approach for secondary organic aerosol formation.[44] This richer framework enables improved simulation of secondary pollutant formation, making it more suitable for regions such as northern India where wintertime stagnation and complex emissions contribute to severe pollution episodes.

Chemical mechanism performance was evaluated using pollutant concentrations from CPCB monitoring stations, MERRA2 and satellite observations for O_3, SO_4 and NO_x, as shown in Fig. 3, S8, S12, S13 and Table S5.

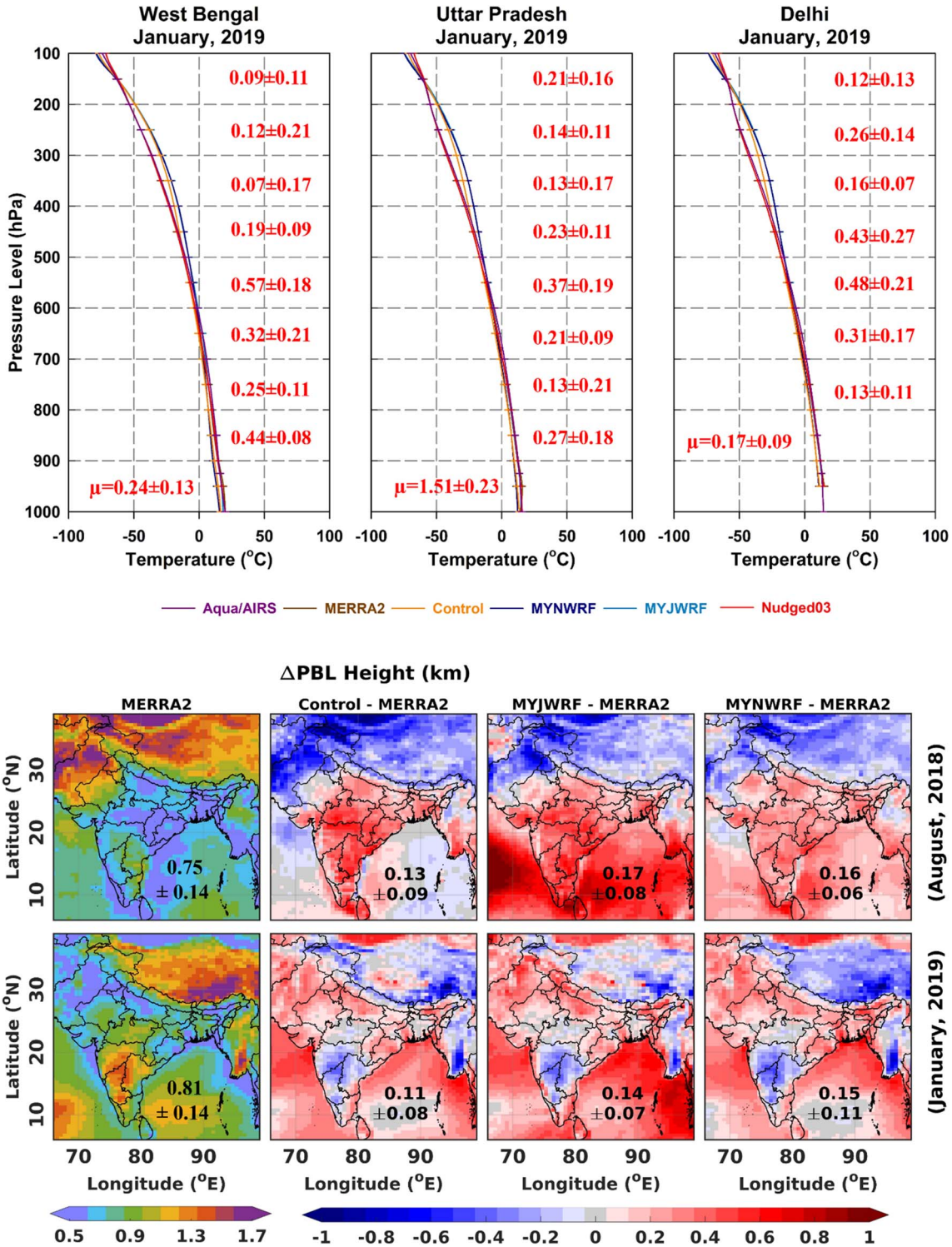

Fig. 2 Vertical temperature profile comparison between MERRA2 reanalysis, Aqua/AIRS observed with different WRF-Chem PBL and grid nudging sensitivity simulations over the entire Indian region (top panel). Profiles are accompanied by domain-mean ± SD statistics over India. Spatial distribution of PBL height by MERRA2 reanalysis over the Indian region, and the difference in simulated PBL height from various WRF-Chem PBL sensitivity simulations (bottom panel). Inset values show domain-mean ± spatial SD ($\mu \pm 1\sigma$), Mean Bias Error (MBE) and Mean Absolute Error (MAE) of Nudged03 computed over the Indian landmass.

2.2.5 Emission inventories. Emission inventories control the spatial and temporal distribution of pollutants entering the atmosphere and remain among the largest sources of uncertainty in chemical transport modelling.[45] Their accuracy is especially critical for India, where emissions vary sharply across dense urban centres, industrial clusters, rural areas reliant on

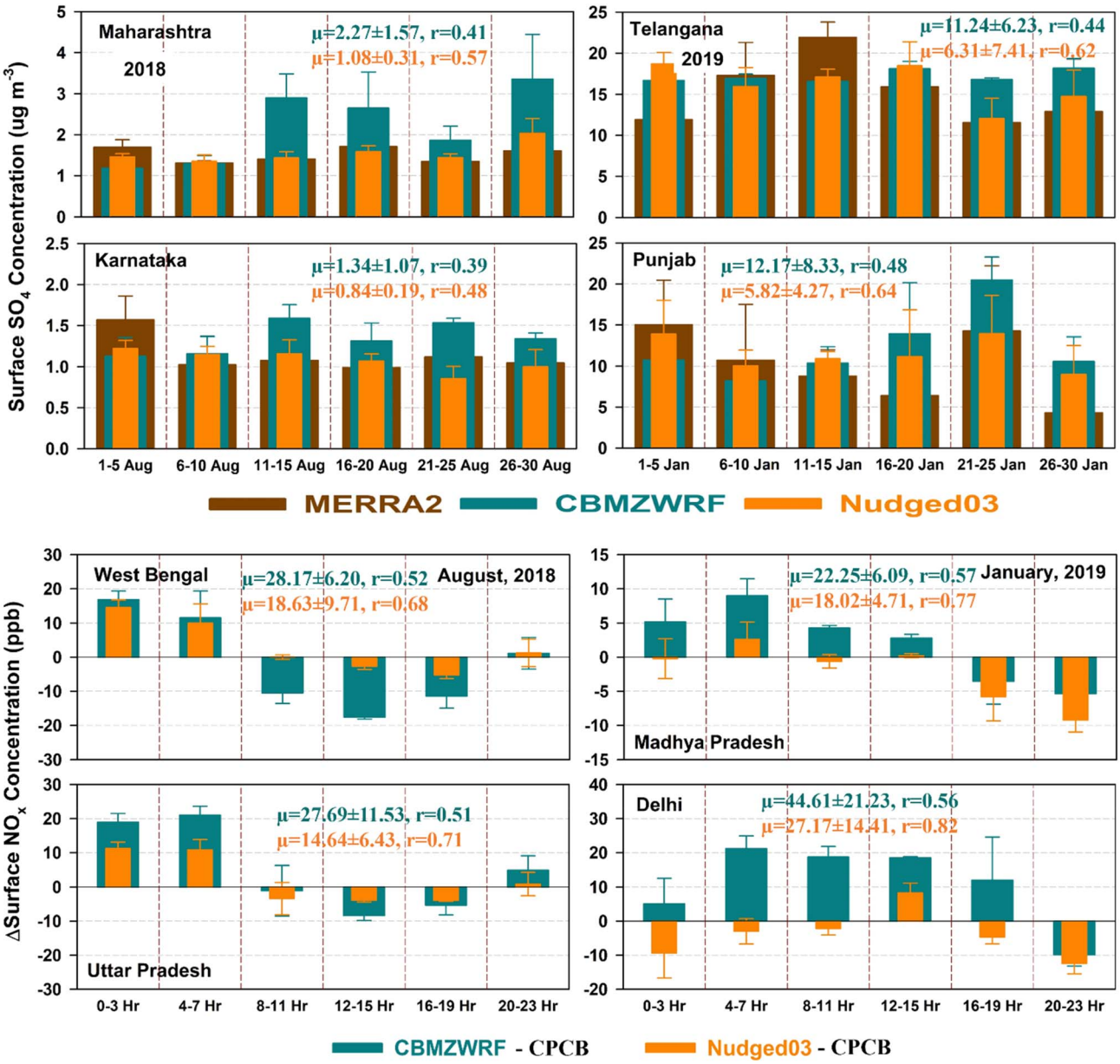

Fig. 3 Bar charts showing 5 days mean time series comparison of surface SO_4 distribution for various Indian states from different WRF-Chem chemical mechanism sensitivity simulations with MERRA2 reanalysis (top panel). Bars represent mean ± SD over the averaging window. Bar charts illustrating the diurnal variation in the difference of 4 hourly mean surface NO_x distributions for various Indian states, derived from different WRF-Chem chemical mechanism sensitivity simulations compared with CPCB measurements (bottom panel). Temporal mean ± SD ($\mu \pm 1\sigma$), Mean Absolute Error (MAE) and correlation coefficient (r) for CBMZWRF and Nudged03 are indicated respectively within the panel to facilitate quantitative comparison.

biofuel combustion, and extensive biomass-burning regions. To evaluate how these uncertainties propagate into air-quality simulations, five widely used emission inventories were examined during both the monsoon and winter simulation periods.

The Coupled Model Intercomparison Project Phase 6 (CMIP6) inventory from Community Emissions Data System,[46] available at a coarser 50 km resolution, is optimized for climate modelling applications rather than urban air-quality forecasting. Its coarse spatial representation tends to smooth strong emission gradients, which can lead to systematic underestimation of pollutant concentrations in densely populated or industrially active regions. Emissions Database for Global Atmospheric Research, Version 5 (EDGARv5),[47] with a resolution of 10 km, supplies global emission coverage and updated sectoral contributions, but its reliance on generalized global activity data limits its ability to capture India's fine-scale urban and industrial emission structures. Emissions Database for Global Atmospheric Research, Version 6.1 (EDGARv6.1),[48] also provided at 10 km resolution, incorporates more recent emission factors and activity data, yet still lacks the spatial

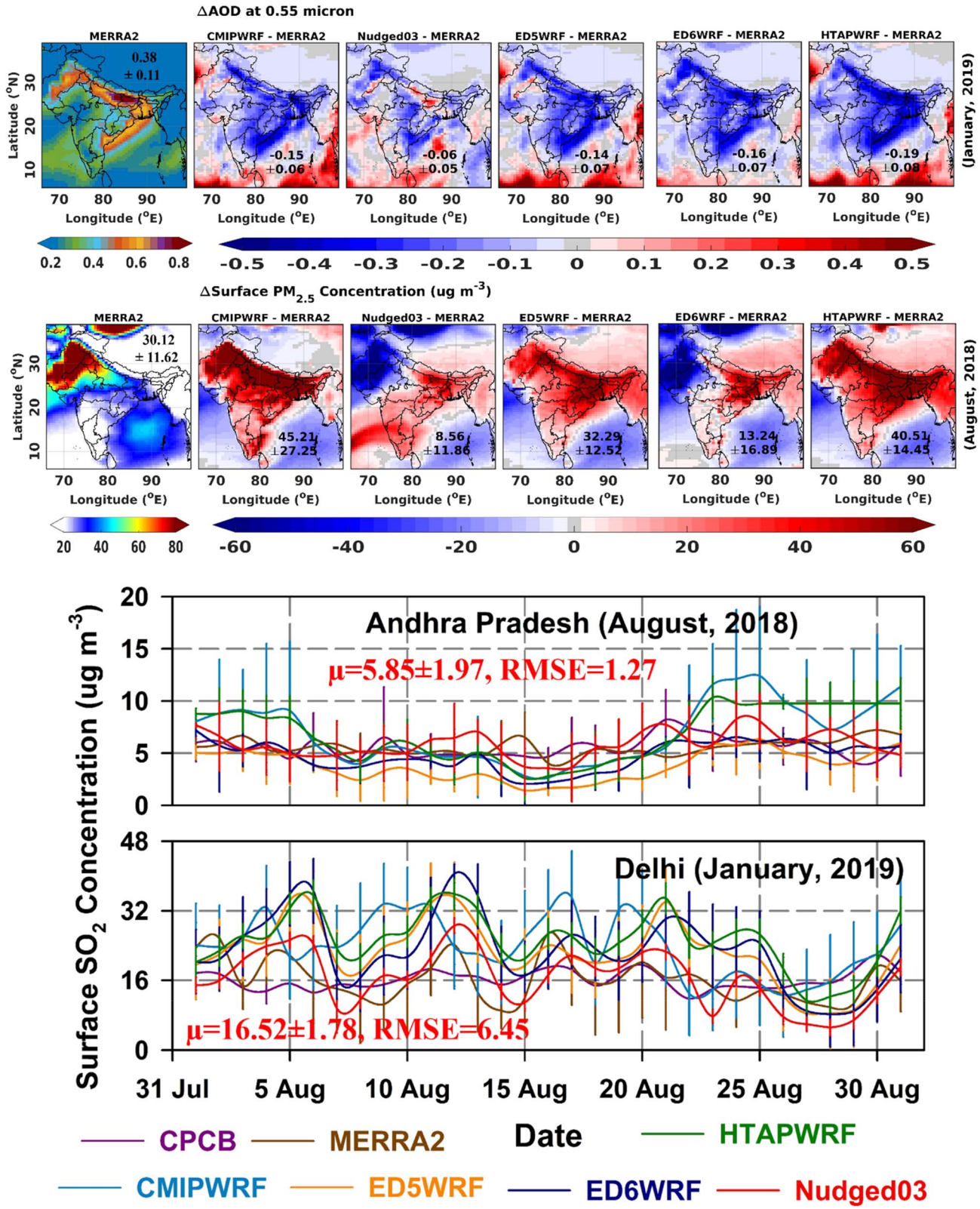

Fig. 4 Spatial distribution of AOD from MERRA2 reanalysis and the difference in simulated AOD from various WRF-Chem emission inventory sensitivity simulations over the Indian region (top panel). Inset values show domain-mean ± spatial SD ($\mu \pm 1\sigma$) and MBE over the Indian landmass. Spatial distribution of surface $PM_{2.5}$ concentration by MERRA2 over the Indian region, and the difference in simulated surface $PM_{2.5}$ from various WRF-Chem emission inventory sensitivity simulations (middle panel), with domain-mean ± SD ($\mu \pm 1\sigma$) and MBE reported within the panel. Daily mean time series comparison of surface SO_2 distribution for various Indian states from different WRF-Chem emission inventory sensitivity simulations with CPCB measurements and MERRA2 reanalysis (bottom panel). Temporal mean ± SD ($\mu \pm 1\sigma$) and RMSE ($\mu g\ m^{-3}$) of Nudged03 are indicated.

detail required to characterize the intense but localized sources common in Indian metropolitan regions. Emissions Database for Global Atmospheric Research – Hemispheric Transport of Air Pollution, Version 3 (EDGAR-HTAPv3)[49] provides updated global anthropogenic emissions intended for hemispheric transport assessments but remains limited in its ability to represent the complex emission heterogeneity characteristic of India.

The Speciated Multi-pOllutant Generator – COordinated Aerosol and oLution SOurce Emissions for Chemical Evaluation (SMoG-COALESCE) inventory[50] is a recent high-resolution, India-specific dataset (5 km) that provides detailed sectoral emissions and offers the spatial granularity needed to resolve intra-urban gradients and localized emission hotspots. Anthropogenic emissions of $PM_{2.5}$ and its precursors are primarily taken from the Speciated MultipOllutant Generator-India (SMoG-India, 2019),[51] which provides sectorally resolved emissions for residential combustion, transportation, industry, energy production, agriculture, waste burning, and several other categories. This comprehensive sectoral breakdown makes SMoG-COALESCE particularly suited for simulating the diverse and spatially heterogeneous emission landscape of India.

Emission inventories that were directly available for the study period were used without modification. For inventories not available for the exact simulation months (August 2018 and January 2019), the most recent year available in the respective database was adopted as a representative proxy. Specifically, when emissions for August 2018 or January 2019 were available, those datasets were used directly; otherwise, the latest available August and January emissions from that inventory were used to represent the corresponding simulation months. This approach ensures temporal consistency while relying on the most up-to-date emission information available for each inventory. The capability of each inventory to reproduce observed spatial and temporal pollution patterns, including regional gradients and urban pollution hotspots across the Indo-Gangetic Plain and major metropolitan centres, was systematically evaluated using Fig. 4 and S9, S12, S13. The monthly total and sectoral emission comparisons over India from different emission inventories, used in the WRF-Chem emission inventory sensitivity simulations for major primary anthropogenic species during August and January, have been analyzed and are presented in Fig. S14 of the SI. The figure also illustrates the relative contribution of different emission sectors to the total emissions over the Indian region for the corresponding species and months. These comparisons reveal substantial differences in how inventories represent sectoral contributions and spatial variability, which ultimately influence the accuracy of WRF-Chem simulations under contrasting seasonal meteorological conditions.

2.2.6 Grid nudging. Grid nudging was employed to constrain the model toward large-scale meteorological fields, thereby improving atmospheric realism without suppressing local-scale variability. This approach, commonly referred to as Four-Dimensional Data Assimilation (FDDA), enhances the accuracy of temperature, moisture, and wind fields, particularly in extended simulations.[52,53] Nudging was applied above the PBL to preserve near-surface turbulence and vertical mixing processes that strongly influence pollutant dispersion.[54] Three nudging strengths, 0.0003 s^{-1}, 0.0005 s^{-1}, and 0.001 s^{-1}, were evaluated, with details summarized in Tables S1 and S5.

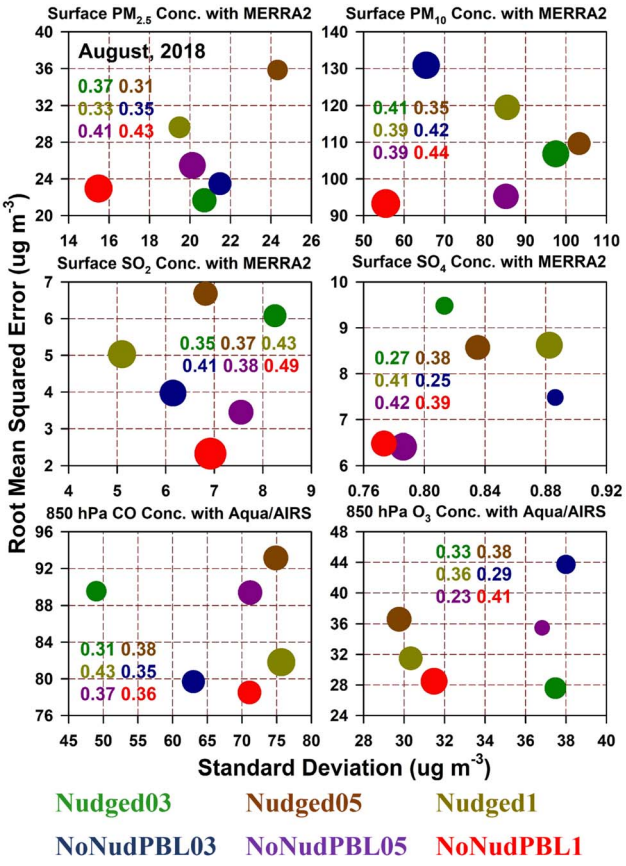

Fig. 5 Bubble plot representation of statistical comparisons for various surface and 850 hPa pollutants over the Indian region against MERRA2 and Aqua/AIRS for different WRF-Chem nudging coefficient sensitivity simulations for August 2018. Bubble size represents RMSE, colour indicates various experiments, and inset values show coefficients of determination (r^2).

The effects of nudging on meteorological fields were examined using surface variables (Fig. 5 and S10), surface temperature and wind speed (Fig. S3), and vertical temperature profiles (Fig. 2). Particular attention was given to winter simulations, where accurate representation of temperature inversions and stable boundary layers is essential for realistic pollutant accumulation.

2.3 Validation datasets

Model evaluation utilized a diverse suite of observational datasets spanning surface, satellite, and reanalysis products to ensure robust multi-scale validation. Ground-based measurements of $PM_{2.5}$, PM_{10}, NO_x, SO_2, and O_3 were obtained from the publicly available Central Pollution Control Board (CPCB) air quality data repository (https://airquality.cpcb.gov.in/ccr/#/caaqm-dashboard-all/caaqm-landing/caaqm-data-repository), providing high temporal resolution observations of pollutant

concentrations across major Indian cities. To address known CPCB data quality issues,[55,56] we applied strict quality control by removing outliers and eliminating constant or repetitive values, ensuring reliable daily chemistry and meteorological observations throughout the study period.[57,58] Satellite observations from Atmospheric Infrared Sounder (AIRS) instrument aboard NASA's Aqua satellite (Aqua/AIRS) and meteorological reanalysis from Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA2) (https://giovanni.gsfc.nasa.gov/giovanni/) were used to validate surface temperature, humidity, wind speed, and vertical structure. Rainfall was assessed using high-quality India Meteorological Department (IMD) datasets (https://imdpune.gov.in/cmpg/Griddata/Rainfall_25_NetCDF.html).[59] These datasets collectively enabled evaluation of the model's capability to capture meteorological conditions, boundary-layer behaviour, and pollutant distributions across India's diverse climate regimes.[30] For spatial panels, we summarize each map using the domain-mean and spatial standard deviation ($\mu \pm 1\sigma$) computed over the Indian landmass (consistent with the land-mask used in the analysis), and these values are reported in the corresponding figure panels/captions. For temporal panels, we report the time-mean and temporal standard deviation ($\mu \pm 1\sigma$) over the evaluation period, also indicated in the relevant figures. In addition, model skill is quantified using mean bias error (MBE), mean absolute error (MAE), root-mean-square error (RMSE), correlation coefficient (r), and coefficient of determination (r^2), computed consistently for each sensitivity experiment against the corresponding observational/reanalysis dataset. Six main-figure evaluations summarize model performance, complemented by extensive diagnostics in the SI.

3 Results and discussion

The WRF-Chem sensitivity simulations conducted for August 2018 (monsoon) and January 2019 (winter) provide a comprehensive framework for evaluating model performance across India's contrasting meteorological regimes. August is dominated by deep convection, strong moisture transport, high humidity, and vigorous wet deposition, whereas January is characterized by stagnant conditions, shallow planetary boundary layers (PBLs), frequent temperature inversions, and weak winds that promote severe pollutant accumulation. These stark meteorological contrasts exert strong controls on transport, dispersion, deposition, aerosol hygroscopicity, and secondary pollutant formation. By systematically evaluating boundary conditions, physical parameterizations, chemical mechanisms, emission inventories, and grid-nudging strengths, and by comparing simulations against CPCB surface measurements, MERRA2 reanalysis, Aqua/AIRS satellite retrievals, and IMD rainfall, we identify an optimized WRF-Chem configuration capable of realistically capturing meteorological and air-quality processes across India. Statistical metrics including Mean Absolute Error (MAE), Mean Bias Error (MBE), Root Mean Square Error (RMSE), correlation coefficient (r), and coefficient of determination (r^2) were used to quantify model skill.[28]

3.1 Meteorological sensitivity simulations

Meteorological initial and boundary conditions had a dominant influence on model performance across both seasons. Simulations driven by ERA5 consistently outperformed those using NCEP GDAS or NCEP CFSv2 for key variables including surface temperature, relative humidity, cloud fields, and near-surface winds. As shown in Fig. 1, ERA5 reduced surface temperature biases by up to 2 °C relative to GDAS and CFSv2, particularly across northern and central India. Comparison with Aqua/AIRS surface temperature showed that ERA5 exhibits the best performance over the Indian landmass during August, with an RMSE of 2.09 °C and a correlation coefficient of 0.97. In contrast, GDAS and CFSv2 showed higher RMSE values of 2.63 °C and 2.81 °C, with lower correlation coefficients of 0.91 and 0.89, respectively. This improvement reflects ERA5's finer spatial and temporal resolution, which enables more accurate representation of mesoscale processes such as monsoonal convection, moisture convergence, and weak synoptic disturbances. These improvements are further illustrated in Fig. S2, which highlights enhanced spatial temperature distributions important for capturing land–atmosphere interactions.

Wind speed and wind direction, which are critical determinants of horizontal pollutant transport and dispersion, also exhibited strong sensitivity to the choice of meteorological initial and boundary dataset. Winter simulations forced with GDAS and CFSv2 tended to underestimate weak-wind and stagnation events in the Indo-Gangetic Plain (IGP), whereas ERA5-driven simulations more accurately captured the magnitude, timing, and diurnal cycle of low-level winds. Fig. S3 demonstrates ERA5's ability to resolve diurnal wind variability, while Fig. S4 shows improved spatial patterns of surface wind speed, both of which are crucial for simulating pollutant accumulation and dispersion during winter.[29]

Comparison with MERRA2 850 hPa wind speed showed that ERA5 exhibits the best performance over the Indian landmass during January, with an RMSE of 1.72 m s^{-1} and a correlation coefficient of 0.61. In contrast, GDAS and CFSv2 showed higher RMSE values of 1.81 m s^{-1} and 1.89 m s^{-1}, with lower correlation coefficients of 0.58 and 0.55, respectively. ERA5 also provided a markedly better representation of relative humidity and moisture fields, which influence cloud formation, fog and haze occurrence, and wet deposition. In August, the improved moisture structure led to more realistic rainfall patterns, as evidenced by Fig. S2 and S6 (accumulated rainfall), enabling a more accurate simulation of aerosol wet scavenging processes. In January, realistic humidity fields improved the representation of fog, haze, and aerosol hygroscopic growth, all of which have strong implications for visibility and particulate matter concentrations. A detailed statistical analysis of ERA5 performance for various meteorological variables has been summarized in Table S2.

The combined improvements in temperature, wind, and humidity fields under ERA5 translated directly into enhanced air-quality simulations. More realistic dispersion and boundary-layer structure led to improved $PM_{2.5}$, NO_x, and CO concentrations. These results confirm ERA5 as the preferred

meteorological initial and boundary dataset for WRF-Chem simulations over India, especially in regions affected by complex interactions among synoptic forcing, monsoonal dynamics, inversions, and fog formation. Overall, the use of ERA5 forcing reduced surface temperature RMSE by approximately 20–25% relative to GDAS and CFSv2 during August 2018, while increasing correlation coefficients by up to 0.08. During January 2019, ERA5 improved 850 hPa wind speed skill by reducing RMSE by ~0.1–0.2 m s^{-1} and increasing correlation by ~0.05–0.06. These quantitative improvements highlight the strong dependence of air-quality simulations on the accuracy of large-scale meteorological forcing.

3.2 PBL and microphysics scheme sensitivity

The PBL and microphysics schemes are critical components of atmospheric modelling because they govern vertical mixing, pollutant dispersion, cloud formation, precipitation, and wet deposition. Together, these processes strongly influence near-surface pollutant concentrations and therefore the skill of air-quality simulations over India.

Boundary-layer dynamics exert particularly strong control over pollutant accumulation by determining the depth of vertical mixing and the extent of surface pollutant dilution. Fig. 2 shows that the YSU PBL scheme produced the most realistic PBL heights during both monsoon and winter seasons. During January 2019, the ability of YSU to capture shallow nocturnal PBLs and strong inversion layers was essential for simulating wintertime pollution buildup over the IGP. In contrast, MYJ and MYNN3 systematically overestimated PBL height, resulting in excessive vertical mixing and underestimation of $PM_{2.5}$, NO_x, and CO concentrations across the IGP and other polluted urban regions.

Evaluation against MERRA2 PBL height indicates that the YSU scheme performs better over the Indian landmass during January, with an RMSE of 0.18 km and a correlation coefficient of 0.49. In comparison, MYJ and MYNN3 show higher RMSE values (0.21 km each) and lower correlations (0.38 and 0.44, respectively). Consistent behaviour is found for August when compared with Aqua/AIRS 850 hPa temperature, where YSU again shows superior performance (RMSE: 2.58 °C; correlation: 0.58), while MYJ and MYNN3 exhibit larger errors (3.29 °C and 3.19 °C) and weaker correlations (0.52 and 0.51, respectively). These findings are consistent with earlier studies demonstrating that YSU effectively represents diurnal boundary-layer dynamics and nonlocal mixing under stable conditions over India.[15,34] Spatial distributions of $PM_{2.5}$ (Fig. 4), together with vertical structure diagnostics (Fig. S5), further highlight the YSU scheme's ability to capture pollutant accumulation under stable winter conditions and to improve simulations of surface $PM_{2.5}$, NO_x, and O_3.

During August, the YSU scheme again performed best by realistically representing deep convective mixing and the diurnal evolution of the boundary layer under monsoon conditions. Accurate simulation of daytime PBL growth is essential for representing pollutant lofting into the free troposphere, while nighttime collapse strongly influences surface accumulation and nocturnal chemistry. The robustness of YSU across both monsoon and winter therefore highlights its suitability for simulating boundary-layer processes and pollutant dispersion in regions characterized by strong seasonal contrasts in stability and pollution levels. A detailed statistical analysis of YSU PBL scheme performance for various meteorological variables has been summarized in Table S2.

Microphysics schemes introduced additional sensitivity, particularly during the monsoon when rainfall dominates pollutant removal. Microphysics governs cloud processes, precipitation formation, and wet deposition, making it central to both meteorological and air-quality performance.[33] Fig. S6 compares accumulated rainfall for the Lin, Thompson, and Morrison schemes. The Lin scheme[38] produced the most realistic rainfall patterns, especially over central and southern India where convective precipitation dominates during the monsoon season (with IMD RMSE: 12.16 mm and correlation: 0.37). Accurate representation of both stratiform and convective rainfall structures is critical because rainfall intensity, frequency, and spatial distribution determine the efficiency and timing of wet deposition, the primary mechanism for pollutant removal during this period.

The realistic rainfall fields simulated by the Lin scheme led to improved predictions of pollutant concentrations, particularly for SO_2 and PM_{10}, as demonstrated in Fig. 4 and S9. In contrast, the Thompson and Morrison schemes tended to overestimate rainfall in several regions, leading to excessive wet scavenging and unrealistically low concentrations of aerosols and trace gases. Furthermore, cloud fraction representation, crucial for radiative transfer, boundary-layer evolution, and photolysis rates, was most accurate in the Lin scheme (Fig. S7), enhancing the reliability of pollutant concentration predictions during monsoon conditions.[60]

A comprehensive statistical comparison of surface temperature, wind speed, relative humidity, cloud fraction, and accumulated rainfall against Aqua/AIRS observations, MERRA2 reanalysis, and IMD measurements[59] confirms that simulations using ERA5 meteorology coupled with the YSU PBL and Lin microphysics schemes achieved the best overall performance across the Indian region (Table S3). The sensitivity analysis therefore highlights the superior performance of the YSU scheme for simulating boundary-layer processes and the Lin scheme for capturing precipitation and pollutant removal mechanisms. Their combined application offers a robust framework for improving meteorological and air-pollution simulations over India, and these schemes are recommended for air-quality modelling in both winter and monsoon periods to support more accurate forecasts and policy-relevant assessments. Relative to MYJ and MYNN3, the YSU scheme reduced wintertime PBL height RMSE by ~15–20%, resulting in systematically higher near-surface pollutant concentrations that are more consistent with CPCB observations. For microphysics, the Lin scheme reduced monsoon rainfall RMSE by ~30–40% compared to Thompson and Morrison, leading to more realistic wet scavenging and avoiding excessive removal of aerosols and trace gases.

3.3 Chemical mechanism and emission inventory sensitivity

Chemical mechanisms and emission inventories played a crucial role in determining the accuracy of simulated pollutant concentrations. Fig. 3, 4 and S8, S9 show that the MOZART-MOSAIC mechanism consistently outperformed CBMZ in reproducing O_3, $PM_{2.5}$, SO_2, and CO across India. MOZART-MOSAIC includes detailed ozone–NO_x–VOC chemistry, aerosol thermodynamics, heterogeneous reactions, and VBS-based SOA formation,[41,43] which together enable a more realistic representation of secondary pollutant formation. This advantage was particularly evident during January 2019, when reduced photolysis rates increase the relative importance of multiphase and heterogeneous processes. Fig. S8 and S9 highlight the improved representation of pollutant variability across urban, rural, and industrial regions achieved with MOZART-MOSAIC compared to CBMZ.

Emission inventory choice introduced an additional major source of sensitivity. The SMoG-COALESCE inventory, with its high spatial resolution (5 km) and India-specific sectoral detail,[50] provided the most accurate representation of pollution distributions during both monsoon and winter. Comparison with MERRA2 surface $PM_{2.5}$ showed that SMoG-COALESCE exhibits better performance over the Indian landmass during January, with an RMSE of 44.46 μg m^{-3} and a correlation coefficient of 0.87. In contrast, EDGARv6.1 and EDGAR-HTAPv3 showed higher RMSE values of ~70 μg m^{-3} for both, with lower correlation coefficients of 0.80 and 0.41, respectively. Urban regions such as Delhi, Kolkata, and Mumbai showed substantial improvements in simulated pollutant concentrations when SMoG-COALESCE was used, owing to its realistic representation of industrial, transport, and residential emissions (Fig. 4). In contrast, coarser global inventories such as EDGARv5, EDGARv6.1, EDGAR-HTAPv3, and CEDS-CMIP6 failed to capture

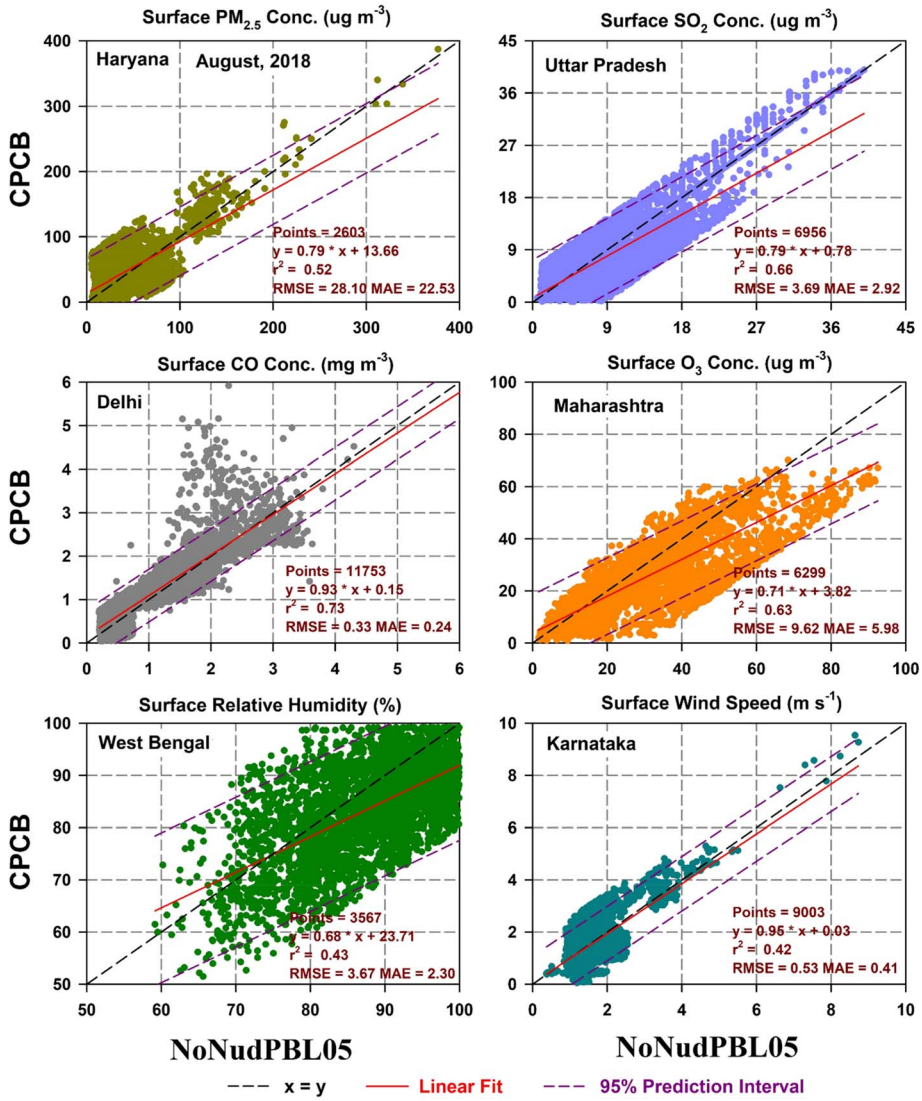

Fig. 6 A statistical comparison of preferred WRF-Chem-simulated meteorological parameters and pollutant concentrations with CPCB measurements is shown for various Indian states with available continuous monitoring data (see Fig. S1) for August 2018. Panels report state-wise regression statistics (MAE, RMSE, and r^2) indicated in parentheses, enabling direct quantitative assessment of model skill across regions.

localized emission hotspots and intra-urban gradients, leading to systematic underprediction of pollutant concentrations in densely populated regions and along industrial corridors (Fig. 4 and S9). A detailed statistical analysis of the performance of all emission inventories for various air pollutants is summarized in Table S4.

Overall, the combination of the MOZART-MOSAIC chemical mechanism with the SMoG-COALESCE emission inventory provided the most reliable predictions for air quality over India, particularly in urban and industrial regions with complex emission signatures. These results highlight the importance of using both a chemically comprehensive mechanism and a high-resolution, region-specific emission inventory for accurate air-quality simulations (Fig. 3, 4 and S8, S9, S12, S13). The use of the SMoG-COALESCE inventory reduced $PM_{2.5}$ RMSE by approximately 35–40% relative to global inventories during winter and increased correlation coefficients by up to 0.46 in some regions (Fig. S12 and S13). These differences highlight the dominant role of emission resolution and sectoral representation in controlling simulated pollution levels over India, particularly under stagnant winter conditions.

Although MOZART-MOSAIC consistently outperforms CBMZ-MOSAIC for the present configurations and evaluation periods, this result should be interpreted in a process-based context rather than as an indication of universal superiority. The improved performance of MOZART-MOSAIC during January 2019 is physically consistent with its more comprehensive treatment of ozone–NO_x–VOC chemistry, heterogeneous reactions, aerosol thermodynamics, and volatility-basis-set (VBS) secondary organic aerosol formation. These processes become increasingly important under wintertime conditions characterized by weak photolysis, shallow boundary layers, and enhanced multiphase chemistry, which are prevalent over northern India during severe pollution episodes. At the same time, it is important to recognize that apparent improvements in model skill may partially reflect compensating errors among meteorology, emissions, and chemistry. For example, uncertainties in emission magnitudes, spatial allocation, or vertical mixing can offset biases in chemical production or loss pathways, leading to improved agreement with observations without necessarily implying optimal representation of individual processes (Fig. S14). Under different regimes, such as pre-monsoon dust-dominated conditions or post-monsoon biomass-burning periods, the relative importance of mineral dust chemistry, aerosol optical properties, fire emission injection heights, and plume aging may alter the comparative performance of chemical mechanisms. Consequently, the advantages identified here for MOZART-MOSAIC should be viewed as regime-dependent, and further evaluation across additional seasons, years, and emission scenarios is required to assess broader transferability.

3.4 Grid nudging sensitivity

Grid nudging was applied in WRF-Chem to improve the realism of large-scale meteorological fields by gently constraining the model toward reanalysis data while preserving local variability essential for accurate representation of pollutant transport and dispersion. Fig. 5, S10 and Table S5 illustrate the influence of different nudging strengths on surface meteorological and chemical variables, demonstrating that a nudging coefficient of 0.0005 s^{-1} provided the most effective balance between correcting large-scale biases and maintaining mesoscale dynamical fidelity. This optimal strength aligns with earlier recommendations for regional air-quality modelling, which emphasize that nudging should correct synoptic features without suppressing local circulations.[52,53]

Excessive nudging at 0.001 s^{-1} weakened boundary-layer turbulence, sea-breeze circulations, and urban heat-island effects, all of which are crucial for pollutant dispersion. Conversely, weaker nudging at 0.0003 s^{-1} was insufficient to correct synoptic-scale wind and temperature biases, allowing errors in the reanalysis boundary conditions to propagate through the domain (Table S5).

Fig. 1 and 2, reinforced by vertical structure diagnostics in Fig. S3 and Table S5, show that moderate nudging (0.0005 s^{-1}) significantly improved the representation of vertical temperature profiles during January 2019. The enhanced simulation of temperature inversions and boundary-layer stability strengthened the model's ability to capture conditions that trap pollutants near the surface over northern India. Improved inversion strength was accompanied by better reproduction of low-level wind speed and direction, which play a critical role in determining wintertime $PM_{2.5}$ and NO_x accumulation (Table S5). In addition, moderate nudging improved moisture and relative humidity fields, enhancing the simulation of fog, haze, cloud microphysics, and wet deposition processes features especially important during strongly stratified winter conditions.

These meteorological improvements translated directly into enhanced pollutant simulations. Fig. 5 and 6 demonstrate that the 0.0005 s^{-1} nudging configuration produced substantially more accurate $PM_{2.5}$ and NO_x concentrations, particularly during stagnant cold-season episodes when small errors in stability and wind fields can lead to large discrepancies in pollutant predictions (Table S5). Thus, the analyses clearly indicate that moderate grid nudging at 0.0005 s^{-1}, applied above the PBL, provides the optimal configuration for improving both meteorological and air-quality predictions over India. It corrects large-scale biases without eroding mesoscale variability essential for pollutant dispersion. These findings further support its inclusion in the preferred WRF-Chem configuration identified in this study, alongside ERA5 initial and boundary conditions, the YSU PBL scheme, the Lin microphysics scheme, the MOZART-MOSAIC chemical mechanism, and the SMoG-COALESCE emission inventory (see Fig. S8 and S9).

Finally, when evaluated as part of the optimized model setup, this nudging configuration contributed substantially to improved agreement with CPCB ground observations across major Indian states (Fig. 6). Coefficients of determination ranged from 0.41 for surface $PM_{2.5}$ in Rajasthan (January 2019) to 0.73 for surface CO in Delhi (August 2018), while meteorological parameters such as wind speed and relative humidity showed r^2 values of 0.42 (Karnataka, August 2018) and 0.67

(Andhra Pradesh, January 2019), respectively (Fig. S11). These results highlight the importance of grid nudging in enhancing the fidelity of meteorological and chemical simulations over India's complex atmospheric environment and emphasise its value in generating reliable, policy-relevant air-quality forecasts.

Therefore, the sensitivity experiments demonstrate that differences in model configuration can lead to changes of 20–50% in key performance metrics, emphasizing that physically consistent choices of meteorological forcing, PBL and microphysics schemes, chemical mechanisms, emission inventories, and nudging strategies are essential for reliable air-quality simulations over India.

We note that the present evaluation does not explicitly include pre-monsoon dust-dominated conditions or post-monsoon biomass-burning episodes. Consequently, while the optimized configuration is robust across monsoon and winter extremes, additional uncertainties may arise in seasons where mineral dust emissions, aerosol mixing state and optical properties, and biomass-burning emission timing/injection heights play a dominant role. Extending the optimized configuration to these regimes is a logical next step to further refine season-specific aerosol and emission process representations.

4 Summary and conclusions

This study presents a comprehensive, multi-component sensitivity analysis of WRF-Chem (v4.2.2) to assess its ability to simulate meteorology and air quality over India under two sharply contrasting seasonal regimes, peak monsoon conditions (August 2018) and peak winter conditions (January 2019). By systematically perturbing meteorological initial and boundary conditions, planetary boundary layer (PBL) schemes, microphysics schemes, chemical mechanisms, anthropogenic emission inventories, and grid-nudging strengths, while holding all other settings fixed in each experiment, we identify an optimal model configuration that significantly improves the realism of both meteorological fields and pollutant distributions. The deliberate use of monsoon and winter periods provides a stringent test of WRF-Chem's coupled meteorology–chemistry behaviour across India's broad range of dynamical, thermodynamic, and removal conditions.

Across all experiments, the selection of meteorological initial and boundary conditions emerged as the dominant control on model performance, establishing the baseline realism upon which chemistry and aerosol processes operate. Simulations driven by ERA5 consistently outperformed those forced by NCEP GDAS and NCEP CFSv2 for key variables, including near surface temperature, relative humidity, wind fields, cloud characteristics, and vertical thermodynamic structure. ERA5's superior resolution and improved representation of mesoscale circulation and moisture transport enhanced the simulation of monsoonal convection and rainfall-related processes, while also better capturing wintertime stagnation, inversion strength, and shallow boundary-layer evolution. These meteorological improvements translated directly into more realistic pollutant dispersion, accumulation, and removal, underscoring that credible air-quality predictions over India require meteorological forcing that robustly represents both convective and stable-regime dynamics.

Sensitivity to boundary layer physics further demonstrated that accurate representation of vertical mixing and stability is essential for reproducing observed pollution levels. Among the tested schemes, the YSU PBL scheme provided the most realistic PBL structure in both seasons. Its ability to represent shallow nocturnal PBLs and strong inversions during winter improved the simulation of near-surface pollutant build-up over northern India, while its performance under convective conditions supported more realistic daytime mixing during the monsoon. Microphysics choices introduced additional sensitivity, particularly during August when wet scavenging is the principal removal pathway for aerosols and soluble gases. The Lin microphysics scheme produced the most realistic rainfall distribution and cloud characteristics among the tested options, contributing to improved representation of monsoon-season wet deposition and associated pollutant variability. Together, these results highlight that physically consistent boundary-layer and cloud-precipitation processes are prerequisite for reliable air-quality simulation across India's seasonal extremes.

Chemical mechanisms and emission inventories introduced substantial additional uncertainty by controlling secondary formation pathways and the magnitude and structure of emitted precursors. Within the tested configurations, MOZART-MOSAIC performed better than CBMZ-MOSAIC for multiple pollutants, consistent with its more comprehensive treatment of ozone–NO_x–VOC chemistry, aerosol thermodynamics, and secondary pollutant formation processes that are especially relevant under wintertime stagnant conditions. Emission inventory choice exerted a similarly strong influence on simulated pollution magnitudes and spatial gradients. The SMoG-COALESCE inventory, with its India-specific sectoral detail and high spatial resolution, best captured observed urban-industrial hotspots and regional gradients, whereas coarser global inventories (*e.g.*, EDGAR and CEDS-CMIP6) tended to smooth localized emissions and systematically underpredict concentrations over densely populated and industrial corridors, particularly across the Indo-Gangetic Plain. The combined results reinforce that both chemically appropriate mechanism complexity and regionally realistic, high-resolution emissions are necessary to reproduce India's highly heterogeneous pollution environment.

Grid nudging experiments further emphasized the need to balance synoptic-scale realism with preservation of mesoscale and boundary-layer variability. A moderate nudging coefficient of 0.0005 s^{-1} applied above the PBL provided the best overall performance, improving large-scale circulation patterns, vertical thermal structure, and winter inversion strength without excessively damping local circulations and mixing processes that govern pollutant dispersion. Stronger nudging suppressed mesoscale variability relevant to air quality, whereas weaker nudging did not adequately constrain synoptic-scale biases. These findings support the use of moderate, above-PBL nudging as a pragmatic approach for extended regional air-quality simulations over India.

Based on the combined evaluation, the preferred WRF-Chem configuration for India in this study consists of ERA5 meteorological initial and boundary conditions, the YSU PBL scheme, Lin microphysics, the MOZART-MOSAIC chemical mechanism, the SMoG-COALESCE anthropogenic emission inventory, and grid nudging ($0.0005\ s^{-1}$) applied above the boundary layer. This configuration improves the simulation of near-surface meteorology, pollutant concentrations, and their spatiotemporal patterns across diverse Indian regions under both monsoon and winter regimes. Agreement with CPCB observations and satellite/reanalysis products confirms that the optimized setup provides a robust foundation for applications requiring physically consistent meteorology–chemistry coupling, including air-quality forecasting, retrospective assessments, and scenario-based analyses relevant to mitigation planning.

At the same time, the conclusions of this work should be interpreted with appropriate caution. The optimized configuration demonstrates robustness within the evaluated months and tested sensitivity matrix, but it should not be interpreted as universally optimal across all seasons, years, or emission scenarios. In regional air-quality modelling, improved agreement with observations can arise partly from compensating errors among meteorological forcing, emissions, and chemistry. Accordingly, the stronger performance of MOZART-MOSAIC relative to CBMZ-MOSAIC here should be viewed as regime-dependent and linked to the specific seasonal conditions examined, rather than as unconditional superiority. A further limitation is that the sensitivity experiments were conducted for peak monsoon (August) and peak winter (January) only, and therefore do not explicitly evaluate pre-monsoon dust-dominated conditions or post-monsoon biomass-burning episodes. Although these processes are represented in WRF-Chem, key uncertainties such as dust emission parameterizations, aerosol optical properties, and biomass-burning emission magnitude, temporal allocation, and plume injection height may influence model skill in those regimes and therefore warrant targeted evaluation. Extending the optimized configuration to the pre- and post-monsoon periods, as well as testing its transferability across additional years and emission scenarios, constitutes an important next step toward refining season-specific process representations and strengthening the generality of the framework developed here.

Author contributions

DG conceived the study. IN carried out model simulations and did the main analysis. All authors contributed to the interpretation of the results and finalizing the manuscript.

Conflicts of 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.

Data availability

All data that support the findings of this study are included within the article (and any supplementary information (SI)). The model code used for the simulations can be accessed from https://www2.acom.ucar.edu/wrf-chem, the data used for validation included Aqua/AIRS, MERRA-2, IMD, and CPCB datasets, available from https://giovanni.gsfc.nasa.gov/giovanni/, https://imdpune.gov.in/cmpg/Griddata/Rainfall_25_NetCDF.html, https://airquality.cpcb.gov.in/ccr/#/caaqm-dashboard-all/caaqm-landing/caaqm-data-repository, for data processing and visualization, we used MATLAB software, available at https://www.mathworks.com/products/matlab.html. Supplementary information is available. See DOI: https://doi.org/10.1039/d5ea00167f.

Acknowledgements

The First author acknowledges IIT Delhi for the institute assistantship. The authors thank the IIT Delhi HPC facility for computational resources. This work is partially supported by the Ministry of Environment, Forest and Climate Change (MoEF&CC), Government of India, under the NCAP-COALESCE project (Grant 14/10/2014-CC (Vol. II)). The views expressed in this document are solely those of the authors and do not necessarily reflect those of the Ministry. The authors gratefully acknowledge Prof. Chandra Venkataraman (IIT Bombay) for coordinating the COALESCE consortium. We would like to acknowledge the MERRA-2 science team for the aerosol datasets. We are also grateful to the ECMWF science team for providing ERA-5 cloud and meteorological variables. We acknowledge the reviewers for their suggestions, which helped to improve the manuscript.

References

1 A. Pandey, P. Sadavarte, A. B. Rao and C. Venkataraman, Trends in multi-pollutant emissions from a technology-linked inventory for India: II. Residential, agricultural and informal industry sectors, *Atmos. Environ.*, 2014, **99**, 341–352.
2 S. D. Ghude, R. Kumar, C. Jena, S. Debnath, R. G. Kulkarni, S. Alessandrini, M. Biswas, S. Kulkrani, P. Pithani, S. Kelkar, V. Sajjan, D. M. Chate, V. K. Soni, S. Singh, R. S. Nanjundiah, M. Rajeevan, *et al.*, Evaluation of PM2.5 Forecast using Chemical Data Assimilation in the WRF-Chem Model: A Novel Initiative Under the Ministry of Earth Sciences Air Quality Early Warning System for Delhi, India, *Curr. Sci.*, 2020, **118**(11), 1803–1815.
3 A. E. Cohen, S. M. Cavallo, M. C. Coniglio and H. E. Brooks, A Review of Planetary Boundary Layer Parameterization Schemes and Their Sensitivity in Simulating Southeastern U.S. Cold Season Severe Weather Environments, *Weather Forecast.*, 2015, **30**, 591–612.
4 M. Mohan and S. Bhati, Analysis of WRF Model Performance over Subtropical Region of Delhi, India, *Adv. Meteorol.*, 2011, **2011**, 621235.

5 B. Mahapatra, M. Walia and N. Saggurti, Extreme weather events induced deaths in India 2001–2014: Trends and differentials by region, sex and age group, *Weather Clim. Extrem.*, 2018, **21**, 110–116.
6 D. Ganguly, A. Jayaraman, T. A. Rajesh and H. Gadhavi, Wintertime aerosol properties during foggy and nonfoggy days over urban center Delhi and their implications for shortwave radiative forcing, *J. Geophys. Res.:Atmos.*, 2006, **111**, D15217.
7 W. C. Skamarock, *et al.*, *A Description of the Advanced Research WRF Version 3*, NCAR Tech. Note NCAR/TN-475+STR, Mesoscale Microscale Meteorol. Div, Natl. Cent. Atmos. Res, Boulder, 2008, vol. 475, p. 1.
8 R. Shrivastava, S. K. Dash, R. B. Oza and D. N. Sharma, Evaluation of Parameterization Schemes in the WRF Model for Estimation of Mixing Height, *International Journal of Atmospheric Sciences*, 2014, **2014**, 451578.
9 P. Gunwani and M. Mohan, Sensitivity of WRF model estimates to various PBL parameterizations in different climatic zones over India, *Atmos. Res.*, 2017, **194**, 43–65.
10 S. Ghoshal Chowdhury, D. Ganguly, A. W. Khan and S. Dey, Aerosol-Cloud Interactions and Their Role in Modulating Lightning Activity: Evidence From Extreme Events Over India, *J. Geophys. Res.:Atmos.*, 2025, **130**, e2024JD043251.
11 M. Rajeevan, P. Rohini, K. Niranjan Kumar, J. Srinivasan and C. K. Unnikrishnan, A study of vertical cloud structure of the Indian summer monsoon using CloudSat data, *Clim. Dyn.*, 2013, **40**, 637–650.
12 A. R. Ragi, M. Sharan and Z. S. Haddad, Objective Detection of Indian Summer Monsoon Onset Using QuikSCAT Seawinds Scatterometer, *IEEE Transactions on Geoscience and Remote Sensing*, 2017, **55**, 3466–3474.
13 D. J. Stensrud, *Parameterization Schemes: Keys to Understanding Numerical Weather Prediction Models*, Cambridge University Press, 2011, vol. 9780521865.
14 A. Yerramilli, *et al.*, Simulation of surface ozone pollution in the Central Gulf Coast region during summer synoptic condition using WRF/Chem air quality model, *Atmos. Pollut. Res.*, 2012, **3**, 55–71.
15 K. B. R. R. Hariprasad, *et al.*, Numerical simulation and intercomparison of boundary layer structure with different PBL schemes in WRF using experimental observations at a tropical site, *Atmos. Res.*, 2014, **145–146**, 27–44.
16 A. Anand K A, D. Ganguly, I. Nandi, M. Kajino and S. Dey, Disentangling the Separate and Combined Effects of Aerosol-Radiation and Aerosol-Photolysis Interactions on Air Quality Over India, *J. Geophys. Res.:Atmos.*, 2025, **130**, e2025JD045316.
17 M. S. Sikder and F. Hossain, Improving operational flood forecasting in monsoon climates with bias-corrected quantitative forecasting of precipitation, *Int. J. River Basin Manag.*, 2019, **17**, 411–421.
18 M. Gupta and M. Mohan, Validation of WRF/Chem model and sensitivity of chemical mechanisms to ozone simulation over megacity Delhi, *Atmos. Environ.*, 2015, **122**, 220–229.
19 P. Liu, *et al.*, Differences between downscaling with spectral and grid nudging using WRF, *Atmos. Chem. Phys.*, 2012, **12**, 3601–3610.
20 P. Mukhopadhyay, S. Taraphdar, B. N. Goswami and K. Krishnakumar, Indian Summer Monsoon Precipitation Climatology in a High-Resolution Regional Climate Model: Impacts of Convective Parameterization on Systematic Biases, *Weather Forecast.*, 2010, **25**, 369–387.
21 C. V. Srinivas, *et al.*, Simulation of the Indian summer monsoon regional climate using advanced research WRF model, *Int. J. Climatol.*, 2013, **33**, 1195–1210.
22 B. Schell, I. J. Ackermann, H. Hass, F. S. Binkowski and A. Ebel, Modeling the formation of secondary organic aerosol within a comprehensive air quality model system, *J. Geophys. Res.:Atmos.*, 2001, **106**, 28275–28293.
23 R. A. Zaveri, R. C. Easter, J. D. Fast and L. K. Peters, Model for Simulating Aerosol Interactions and Chemistry (MOSAIC), *J. Geophys. Res.:Atmos.*, 2008, **113**, D13204.
24 A. Balzarini, *et al.*, WRF-Chem model sensitivity to chemical mechanisms choice in reconstructing aerosol optical properties, *Atmos. Environ.*, 2015, **115**, 604–619.
25 M. Mohan and M. Gupta, Sensitivity of PBL parameterizations on PM10 and ozone simulation using chemical transport model WRF-Chem over a sub-tropical urban airshed in India, *Atmos. Environ.*, 2018, **185**, 53–63.
26 G. a. Grell, *et al.*, Fully coupled 'online' chemistry within the WRF model, *Atmos. Environ.*, 2005, **39**, 6957–6975.
27 S. H. Kulkarni, *et al.*, How Much Does Large-Scale Crop Residue Burning Affect the Air Quality in Delhi?, *Environ. Sci. Technol.*, 2020, **54**, 4790–4799.
28 A. Sharma, S. Srivastava, D. Mitra and R. P. Singh, Spatiotemporal distribution of air pollutants during a heat wave–induced forest fire event in Uttarakhand, *Environ. Sci. Pollut. Res.*, 2023, **30**, 110133–110160.
29 N. Anchev, B. Jakimovski, V. Spiridonov and G. Velinov, Temperature Dependent Initial Chemical Conditions for WRF-Chem Air Pollution Simulation Model, in *ICT Innovations 2020. Machine Learning and Applications*, ed. V. Dimitrova and I. Dimitrovski, Springer International Publishing, 2020, pp. 1–14.
30 G. R. Govardhan, R. S. Nanjundiah, S. K. Satheesh, K. K. Moorthy and V. R. Kotamarthi, Performance of WRF-Chem over Indian region: Comparison with measurements, *J. Earth Syst. Sci.*, 2015, **124**, 875–896.
31 E. Kalnay, *et al.*, The NCEP/NCAR 40-Year Reanalysis Project, *Bull. Am. Meteorol. Soc.*, 1996, **77**, 437–472.
32 S. Saha, *et al.*, *NCEP Climate Forecast System Version 2 (CFSv2) 6-hourly Products*, 2011.
33 A. Sathyanadh, T. V. Prabha, B. Balaji, E. A. Resmi and A. Karipot, Evaluation of WRF PBL parameterization schemes against direct observations during a dry event over the Ganges valley, *Atmos. Res.*, 2017, **193**, 125–141.
34 S.-Y. Hong, Y. Noh and J. Dudhia, A New Vertical Diffusion Package with an Explicit Treatment of Entrainment Processes, *Mon. Weather Rev.*, 2006, **134**, 2318–2341.
35 Z. I. Janjić, The Step-Mountain Eta Coordinate Model: Further Developments of the Convection, Viscous Sublayer,

and Turbulence Closure Schemes, *Mon. Weather Rev.*, 1994, **122**, 927–945.
36 M. Nakanishi and H. Niino, An Improved Mellor–Yamada Level-3 Model: Its Numerical Stability and Application to a Regional Prediction of Advection Fog, *Boundary-Layer Meteorol.*, 2006, **119**, 397–407.
37 P. Pithani, *et al.*, WRF model sensitivity to choice of PBL and microphysics parameterization for an advection fog event at Barkachha, rural site in the Indo-Gangetic basin, India, *Theor. Appl. Climatol.*, 2019, **136**, 1099–1113.
38 S.-H. Chen and W.-Y. Sun, A One-dimensional Time Dependent Cloud Model, *Journal of the Meteorological Society of Japan. Ser. II*, 2002, **80**, 99–118.
39 G. Thompson, P. R. Field, R. M. Rasmussen and W. D. Hall, Explicit Forecasts of Winter Precipitation Using an Improved Bulk Microphysics Scheme. Part II: Implementation of a New Snow Parameterization, *Mon. Weather Rev.*, 2008, **136**, 5095–5115.
40 H. Morrison, G. Thompson and V. Tatarskii, Impact of Cloud Microphysics on the Development of Trailing Stratiform Precipitation in a Simulated Squall Line: Comparison of One- and Two-Moment Schemes, *Mon. Weather Rev.*, 2009, **137**, 991–1007.
41 C. Knote, *et al.*, Influence of the choice of gas-phase mechanism on predictions of key gaseous pollutants during the AQMEII phase-2 intercomparison, *Atmos. Environ.*, 2015, **115**, 553–568.
42 R. A. Zaveri and L. K. Peters, A new lumped structure photochemical mechanism for large-scale applications, *J. Geophys. Res.:Atmos.*, 1999, **104**, 30387–30415.
43 L. K. Emmons, *et al.*, Description and evaluation of the Model for Ozone and Related chemical Tracers , version 4 (MOZART-4), *Geosci. Model Dev.*, 2010, **3**, 43–67.
44 C. Knote, *et al.*, Simulation of semi-explicit mechanisms of SOA formation from glyoxal in aerosol in a 3-D model, *Atmos. Chem. Phys.*, 2014, **14**, 6213–6239.
45 K. Tibrewal, *et al.*, Reconciliation of energy use disparities in brick production in India, *Nat. Sustain.*, 2023, **6**, 1248–1257.
46 R. M. Hoesly, *et al.*, Historical (1750–2014) anthropogenic emissions of reactive gases and aerosols from the Community Emissions Data System (CEDS), *Geosci. Model Dev.*, 2018, **11**, 369–408.
47 M. Crippa, *et al.*, High resolution temporal profiles in the Emissions Database for Global Atmospheric Research, *Sci. Data*, 2020, **7**, 121.
48 F. Monforti Ferrario, *et al.*, *EDGAR v6.1 Global Air Pollutant Emissions*, 2022, https://data.europa.eu/89h/df521e05-6a3b-461c-965a-b703fb62313e.
49 M. Crippa, *et al.*, The HTAP_v3 emission mosaic: merging regional and global monthly emissions (2000–2018) to support air quality modelling and policies, *Earth Syst. Sci. Data*, 2023, **15**, 2667–2694.
50 C. Venkataraman, *et al.*, Drivers of PM2.5 Episodes and Exceedance in India: A Synthesis From the COALESCE Network, *J. Geophys. Res.:Atmos.*, 2024, **129**, e2024JD040834.
51 R. M. Lal, *et al.*, Transboundary emission contribution to PM2.5 concentrations in Indian cities, *Environ. Res. Lett.*, 2025, **20**, 54008.
52 J. H. Bowden, T. L. Otte, C. G. Nolte and M. J. Otte, Examining Interior Grid Nudging Techniques Using Two-Way Nesting in the WRF Model for Regional Climate Modeling, *J. Clim.*, 2012, **25**, 2805–2823.
53 G. Zittis, A. Bruggeman, P. Hadjinicolaou, C. Camera and J. Lelieveld, Effects of Meteorology Nudging in Regional Hydroclimatic Simulations of the Eastern Mediterranean, *Atmosphere*, 2018, **9**(12), 470.
54 P. Lucas-Picher, J. Cattiaux, A. Bougie and R. Laprise, How does large-scale nudging in a regional climate model contribute to improving the simulation of weather regimes and seasonal extremes over North America?, *Clim. Dyn.*, 2016, **46**, 929–948.
55 D. Sharma and D. Mauzerall, Analysis of Air Pollution Data in India between 2015 and 2019, *Aerosol Air Qual. Res.*, 2022, **22**, 210204.
56 K. Vohra, *et al.*, Urgent issues regarding real-time air quality monitoring data in India: Unveiling solutions and implications for policy and health, *Atmos. Environ.:X*, 2025, **25**, 100308.
57 V. Singh, *et al.*, Diurnal and temporal changes in air pollution during COVID-19 strict lockdown over different regions of India, *Environ. Pollut.*, 2020, **266**, 115368.
58 M. Zhou, Y. Xie, C. Wang, L. Shen and D. L. Mauzerall, Impacts of current and climate induced changes in atmospheric stagnation on Indian surface PM2.5 pollution, *Nat. Commun.*, 2024, **15**, 7448.
59 D. S. Pai, *et al.*, Development of a new high spatial resolution (0.25{\textdegree} X 0.25{\textdegree}) Long period (1901-2010) daily gridded rainfall data set over India and its comparison with existing data sets over the region, *Mausam*, 2014, **65**, 1–18.
60 W. Zhang, S. A. Rutledge, W. Xu and Y. Zhang, Inner-core lightning outbreaks and convective evolution in Super Typhoon Haiyan (2013), *Atmos. Res.*, 2019, **219**, 123–139.