Paper Title: Evaluation of WRF-Chem simulated meteorology and aerosols over northern India during the severe pollution episode of 2016

General review

This paper presents an evaluation of the capabilities of WRF-Chem in replicating seasonal meteorological patterns and aerosol chemistry, with a specific focus on $PM_{2.5}$ and black carbon, across the Indo-Gangetic Plain. The authors have conducted a comparative assessment, comparing the simulations to reanalysis and observational data, in order to assess its performance. The findings of this investigation indicate that the WRF-Chem model is a suitable tool for examining the interplay between aerosols and meteorology during periods of intense pollution. Additionally, the study underscores enhancements in the representation of diurnal boundary layer processes and emission estimations within the model. However, numerous studies have already been conducted to evaluate the performance of WRF-Chem in simulating meteorology and aerosols over the Indian region (Kumar et al., Jena et al., Sengupta et la., etc.). This abundance of existing research makes it challenging to identify the novelty of the current study. Therefore, I recommend a substantial revision of the manuscript and suggest the authors to emphasize on bringing out the novelty of their research before resubmission.

Major Comments:

The paper needs a clearer explanation of its scientific motivation. It's crucial to clarify why this analysis is being conducted, especially considering previous publications that have validated WRF-Chem for aerosol studies over the IGP. The authors should provide a strong rationale for their study or highlight the unique aspects that set their work apart from previous research in this area. The authors should include a comparative discussion that highlights how their results, specifically concerning meteorology and aerosol simulation, compare with or differ from existing research. The assessment of aerosol feedback on meteorology needs to be more explicit.

The authors utilized the MOSAIC 4-bin scheme for aerosol chemistry characterization. However, it is unclear whether they incorporated aqueous phase chemistry into their model. Including aqueous phase chemistry is crucial as it

replicates aerosol wet removal processes, especially related to fog/haze formation during winter. These processes significantly impact atmospheric composition and are valuable for air quality research. Unfortunately, this aspect is mostly absent in the current manuscript, with no clear mention of its inclusion. Winter aerosol chemistry in the IGP is notably affected by aqueous phase chemistry, as highlighted in Acharja et al. (2023). The absence of this process in the model introduces uncertainties, a point consistently emphasized throughout the manuscript.

The authors recognize the model's limitations in accurately representing emissions and land use information. It would be helpful to explain the measures taken to mitigate these limitations and clarify how model validation remains meaningful despite these acknowledged challenges.

The authors acknowledge the imperfections in representing emissions but haven't specified the steps taken to mitigate this uncertainty. This study utilized the 2010 EDGAR-HTAPv2.2 emissions dataset to assess air quality from September to November 2016. However, emissions in India have significantly changed over the 6-year period. Using static emissions without accounting for these changes does not offer an accurate evaluation of the model's performance. Simply acknowledging this uncertainty, a point already discussed in previous studies, doesn't add substantial value to this research. Additionally, considering the diurnal cycle in emissions is crucial. The authors should, at the very least, apply the diurnal cycle based on existing literature, rather than omitting it entirely from emissions modeling.

The authors observed that the model overestimates dust in September due to exaggerated wind and underestimated dust deposition. This issue was previously addressed by Kalenderski et al. (2013), who attempted to adjust the model for this discrepancy.

The model generates outputs hourly, and IEM-ASOS weather data for meteorological parameters and CPCB data of PM2.5 are also available at an hourly resolution. However, in the manuscript, model performance metrics are calculated based on monthly averaged modeled versus observed values, and daily mean values are compared in time series plots. This approach does not accurately reflect the model's performance and can be misleading. To assess the actual model

performance, the authors might consider providing performance statistics based on hourly datasets.

In October, there is substantial biomass burning activity in Punjab and Haryana States, impacting air quality in rural and urban areas downwind of the IGP. The FINN emission inventory notably underestimates these fire emissions (Jena et al., 2015). However, the manuscript almost entirely overlooks the discussion and analysis of this significant event.

The claim that "WRF-Chem accurately represents afternoon meteorology and reasonably reproduces wind patterns" needs elaboration. It's crucial to explain how these factors influence the daily fluctuations in PM2.5 and BC concentrations.

The comparison of model-generated PM2.5 and BC concentrations with MERRA-2 global reanalysis data, using the GOCART scheme, raises concerns. Utilizing observational data for validation would enhance the reliability of the results. Reconsidering this approach is advisable for improved credibility.

The modeled PM2.5 composition predominantly consists of nitrate aerosol. However, during winter in Delhi, chloride significantly contributes to a substantial portion of PM2.5 composition (Ali et al., 2019; Pawar et al., 2023). Does your model setup incorporate chloride chemistry, and is chloride emission included in your inventory?

Recommendation: The authors are encouraged to revise the paper to clarify the scientific objectives of their study. It is essential to differentiate their work from existing literature on the topic. To achieve this, they should thoroughly review previous studies and identify gaps in the current state of knowledge. One potential aspect to explore further could be the vertical distribution of aerosols and their intricate interactions with meteorological conditions during peak pollution seasons. By addressing such gaps and specifying their research focus, the authors can revise their paper with a well-defined scientific objective that contributes valuable insights to the existing body of research.

Minor Comments:

I have concerns about the model setup. Is there a spin-up period given to the model run, and if so, how long is it? The manuscript lacks this information. The authors mentioned the application of nudging but did not specify whether it is applied in the Planetary Boundary Layer (PBL) or across the entire atmosphere. Additionally, the type of nudging and its method are not clear.

Furthermore, the calculation method for Aerosol Optical Depth (AOD) at 550 nm in WRF-Chem needs clarification. The evaluation statistics of AOD with MODIS data were generated using monthly mean values, while MODIS AOD data are available at daily resolution. To assess true model performance, it is crucial to compare daily mean MODIS AOD with daily modeled AOD.

The manuscript lacks proper scientific justification for the underestimation or overestimation of meteorological parameters, PM2.5, and its composition. This aspect needs to be supported with sound scientific reasoning.

Specifically, in line no. 67, a comma is needed after ‘globally.’ In line no. 82, the term "End of October" is somewhat vague. It is recommended that the authors specify the exact starting date of the event. Although this information is provided later in the manuscript, including it here would enhance clarity.