Review of "Assessing the climate and air quality effects of future aerosol mitigation in India using a global climate model combined with statistical downscaling"

The manuscript explores the possibilities of using a global climate model to investigate the effects of aerosol mitigation in India. A machine learning (ML) approach using Random Forest regression is used to downscale PM2.5 concentrations over a polluted city, New Delhi with the help of measured PM2.5 concentrations. Different PM loading future scenarios are projected and compared with the uncorrected and ML-corrected model outputs. The effects of aerosol mitigation are investigated in terms of radiative effects and effective radiative forcing under the PM future scenarios. The authors claim the improvement of global-scale model output in simulating the PM2.5 concentration over a small domain and their effectiveness in estimating the radiative impacts. The study demonstrates the potential of the emerging technique of ML in improving the large-scale model output in the process of statistical downscaling. The study is relevant and unique as mentioned above, and has a significant contribution to the relevant scientific domain. However, some concerns remain significant and need to be considered before publishing.

General Comments

The manuscript focuses on two aspects. (1) Demonstration of an ML technique in improving a global-scale model to simulate the PM2.5 concentrations over a small region via statistical downscaling under different emission scenarios. (2) Estimating the radiative effects of future aerosol scenarios using the RF-corrected model simulations. The manuscript structure is difficult to follow until reaching the present 'Conclusions' section which is not a conclusion, but a nice overview/summary of the entire work. If the authors want to highlight their simulation results regarding the impact of future aerosol mitigation, more discussion is needed with proper references to the existing findings, else it remains as a technical paper demonstrating the potential of ML in statistical downscaling. Currently, the physical mechanisms for some of the simulation outcomes are not given/found, but some tentative reasons are proposed. Many studies are documented the current aerosol-impact scenario using multiple scientific techniques (insitu, remote sensing, etc.) and future projections also for the Indian region. To ascertain the second aspect of the current manuscript, the first part needs to be flawless and should be explained confidently. Many parts of the manuscript are confusing which calls for further explanations for the smooth reading. One of the highlights of the findings is that the improvement of air quality is mostly due to the reduction of OC loading. However, the negative radiative forcing is attributed to the reduction of BC emission. This is an example of confusion arising while going through the manuscript. The map of India shown in Fig. 3 is not matching with the maps published by the institutions that provided the insitu data nor with one of the authors' affiliated institution. Please correct the map as per the source or remove the political boundaries as per the journal's recommendations. Language also needs improvement. The main concerns are listed below.

Methodology:

- Why does section 2.2 stand apart from sections 2.6 and 2.7?
- The hyper parameters were adjusted using different combinations based on the best error statistics. Can you please show the performance of the validation test data?
- What do you mean by setting the depth of each tree to infinity? How can you make sure of avoiding over-fit while keeping the depth of the tree as infinity? What are the criteria for fixing the number of

trees? It is said that a default value of 100 is taken as the number of trees in the present study. Why 100?

- Why the feature importance values are normalized, by doing so what is the chance of smoothing the non-linearity of the dependence of the input variables? Isn't there any criteria to fix the number of input variables? As the authors have pointed out the input variables are mutually correlated which is obvious in the atmosphere, including all of them may lead to over-fitting. What is the authors' claim on this point?
- For the global-scale modelling, ECLIPSE V6b emission scenarios are used. How appropriate is this emission inventory for simulating PM over the Indian region or what are the criteria for selecting this inventory for this study? What is the contribution of this inventory to the high under-estimation of PM loading by the model over Delhi as shown in the manuscript?
- How the exclusion of the mineral dust component solves the issue related to the PM2.5 peaking? How authors can make sure that this exclusion won't affect the other simulation results?
- Coming to the radiative forcing calculations (section 2.8), how is the definition given to the radiative forcing is related to the conventional definitions found in the published literature? If there is any difference please highlight and justify those, else give supporting citations.
- How the effective radiative forcing is estimated?

Other comments

L1: This opening sentence is misleading. The study demonstrates the potential of the ML technique in downscaling a global-scale mode output..

L6: You mean the model output is better than the measured PM2.5 values?

L11: This is a highly impactful statement. Better to give caution to the reader by mentioning the associated large uncertainty as seen in Fig. 3(e).

L38-40: As per the sentence, the role of ACI in aerosol indirect effects is undermined, hence please modify the sentence. Also, please explain the 'local meteorological dynamics' with references.

L72-73: Cannot find in any of the given references that 'emissions from New Delhi' significantly contribute to the ATAL. Please clarify.

L94: HAM 'threats' the chemical compounds..?

L158: Why do OC emissions increase by 2030 in CLE scenario while all others show a reduction?

L267: What is '2D yearly mean value'?

L356: If that is the case, what is the significance of feature importance values?

L385: This __ somewhat?

L395-398: Confusing. The aerosol loading in MITIG_2030 is supposed to be lower than the CLE_2030, then how RF in MITIG_2030 is more negative than CLE_2030 at the Himalayan foothills. Bright background due to strong haze is expected to be more in CLE_2030.

L434-436: cannot understand. Do you mean that the CDNC burden was more in CLE_2030 scenario?

L435: Expand CDNC in the manuscript.

Conclusions: This section can be renamed as summary and conclusions by adding the significant findings of the study as bullet points.