

***Interactive comment on “Estimation of ground level particulate matter concentrations through the synergistic use of satellite observations and process-based models over South Korea” by Seohui Park et al.***

**Anonymous Referee #3**

Received and published: 7 November 2018

This study proposed improved random forest model to predict ground surface PM concentrations over South Korea. Multiple satellite-derived products and model-based meteorological parameters were used as input variables. The results showed that the improved model is effective in predicting high PM concentrations compared to previous research. The manuscript is overall complete, well written and ready for publishing after the following comments are addressed.

1. section 2.2, please explain why you chose those variables as explanatory indicators.

C1

2. The authors adopted oversampling and under-sampling strategies to alleviate the biased estimation problem. "Input variables in the adjacent pixels of high concentration samples were extracted using 3 x 3 or 5 x 5 windows with the corresponding target variables (i.e., PM<sub>2.5</sub> and PM<sub>10</sub>) randomly perturbed within 5% of the focus pixel concentrations. " Will this perturbation introduce uncertainty? How do you chose appropriate window size?

3. In page 12 line 24, "However, the RF-based models developed in our study has proved to be effective for modelling high ground-level PM concentrations." Could you explain why the RF-based models in this study is more effective than previous studies? Is that because sampling strategies used in your study? If so, could you compare the model performances with and without your sampling strategies?

4. Could you explain the accuracy of MODIS-derived AOD and GOCI-derived AOD? This may help explain why GOCI-AOD-based models outperformed MODIS-AOD-based models.

---

Interactive comment on Atmos. Chem. Phys. Discuss., <https://doi.org/10.5194/acp-2018-647>, 2018.

C2