Reviewer #1

The authors estimated the ground-level PM2.5 pollution across China from Himawari-8 data by applying a fast space-time Light Gradient Boosting Machine. As an application study, authors investigated the diurnal variations of ground-level PM2.5 pollution. It was demonstrated that proposed application method is effective for accurate estimation of surface PM2.5.
Response: We appreciate the time and effort you have spent reviewing this manuscript; we have carefully revised it and provide responses to all the questions raised.

it is better to add relative reference (Zhang et al., 2020; Gui et al., 2020) in the introduction section and to describe the characteristics of this study.
Response: We have added these references and described the characteristics of our study in the revised Introduction.

It is better to extended the Section 2.2 into two parts, the first one citing the LightGBM model and its expansibility in PM applications, at the same time the authors should show what's the advantages of LightGBM in terms of other similar machine leaning models.
Response: We have split Section 2.2 into two parts, i.e., "2.2.1 LightGBM model" and "2.2.2 Model development". We have introduced the LightGBM model, its expansibility, and advantages in PM estimations in section 2.2.1 as follows:

"The LightGBM model, a newly developed tree-based machine-learning approach, which was introduced in 2017 (Ke et al., 2017). Using the gradient boosting framework to construct the decision tree, this approach can tackle both regression and classification tasks, and as such can be expanded for PM applications. It can also tackle the main challenge faced in traditional machine-learning approachesnamely, computational complexities, which are very time-consuming. LightGBM is a fast, distributed, and highly efficient method that reduces the number of data samples (*M*) and features (*N*).
…
In addition to the main technologies mentioned above, there are other features of the optimization, such as the leaf-wise tree growth strategy with depth restriction (Shi, 2007), histogram difference acceleration, sequential access gradient, and the support of category feature and parallel learning."

The second part address the modifications that the authors added or modified to the current model, detailing why these modifications are necessary and what kind of inspirations could be taken by readers to remote sensing or more broadly applications.
Response: We have clarified in the second part the modifications made and why we made them as follows:

"It is well known that air pollution has spatiotemporal heterogeneity, leading to large differences in $PM_{2.5}$ concentrations in both time and space. Such characteristics have always been ignored in most traditional statistical regression and artificial intelligence methods.

Studies have shown that including spatiotemporal information has led to improved $PM_{2.5}$ estimates using remote sensing techniques (Z. Li et al., 2017; Wei et al., 2019a, 2020). Therefore, we have introduced a new approach to integrate spatiotemporal information into the LightGBM model. The new model developed here is called the STLG model."

It is better to provide references for all models in Table 3.
Response: Done per your suggestion.

Which version of AOD is used from Himawari-8?
Response: The latest Himawari-8 Version 2 AOD product was employed. This has been clarified in the revised manuscript.

The caption of Figure 1 is not clear enough.
Response: We have expanded the caption of Figure 1 to make it clearer to the readers.

What would be the possible reason for the large gap in upper diagram in Figure 5, by the way, all the labels should be described in the caption.
Response: The large gap in the number of data samples is mainly caused by different degrees of cloud contamination in the satellite aerosol products for different days. We have clarified this in the revision. We have also added more descriptions of each panel in the caption for this figure.

Relative reference:
Zhang, T., He, W., Zheng, H., Cui, Y., Song, H., & Fu, S. (2020). Satellite-based ground PM2. 5 estimation using a gradient boosting decision tree. Chemosphere, 128801.
Gui, K., Che, H., Zeng, Z., Wang, Y., Zhai, S., Wang, Z., ... & Zhang, X. (2020). Construction of a virtual PM2. 5 observation network in China based on high-density surface meteorological observations using the Extreme Gradient Boosting model. Environment International, 141, 105801.