Responses to RC2

High-resolution with full spatial coverage AOD dataset is important for air pollution-related studies, especially for bright surfaces. The authors have done a lot of work to generate a 250 m AOD dataset in arid and semi-arid areas in northwestern China. However, my biggest concern is the spatial resolution of current developed dataset since it is not clear how to generate the 250 m AOD dataset with two much coarse-resolution MAIAC AOD (1 km) and MxD08 (1 degrees) as main predictions although there are some high-resolution auxiliary data. It sounds incredible and I don't find any downscaling approach or descriptions in the paper.

Response: Many thanks for reviewing our manuscript and providing us with constructive feedback. We are sorry for not having expressed the downscaling method clearly in the manuscript, and this has been added in the revision of the relevant (Page 13/Lines:278-286). Actually, the 250 m AOD is generated based on the MAIAC AOD and auxiliary environmental data, and the MxD08 AOD is only used to validate the FEC AOD accuracy and temporal consistency. About downscaling approach (Section 2.6), we adopt the ensemble machine learning methods (bagging trees ensemble).

The idea behind the downscaling method is to establish either a statistical correlation or a physically-based model between coarse-scale image/product and fine-scale auxiliary variables. But a common premise for downscaling methods is that high-resolution auxiliary variables are indispensable. For AOD, elevation (DEM) and land cover (LUCC) data, which are usually at sub-kilometers resolution and have a great impact on AOD distribution, can be regarded as potential high-resolution auxiliary variables. This ensures the feasibility of downscaling AOD products. Besides, high-resolution LUCC and DEM data are easily accessible worldwide, which contributes to the large-scale implementation of the high-resolution AOD mapping based on the downscaling method[1]. The basic idea for downscaling AOD with bagging trees ensemble machine learning (ML) models is to train the relationships between MAIAC AOD and the auxiliary environmental variables at coarse resolution (1 km) using ML algorithms. We then apply the trained relationships to generate a high-resolution FEC AOD product at a fine resolution (250 m). This idea of downscaling has been developed more maturely and is widely used[1-3], and it is based on a complex mathematical feature that is capable of mining the characteristics of different environmental auxiliary variables on the representation of AOD. At the same time, we have studied your comments and responded to them point by point carefully as described below.

[1] Duveiller, G., Filipponi, F., Walther, S., Köhler, P., Frankenberg, C., Guanter, L., and Cescatti, A.: A spatially downscaled sun-induced fluorescence global product for enhanced monitoring of vegetation productivity, Earth Syst. Sci. Data, 12, 1101–1116, https://doi.org/10.5194/essd-12-1101-2020, 2020.

[2] Yang, Q., Yuan, Q., Li, T., and Yue, L.: Mapping $PM_{2.5}$ concentration at high resolution using a cascade random forest based downscaling model: Evaluation and application, Journal of Cleaner Production, 277, 123887, https://doi.org/10.1016/j.jclepro.2020.123887, 2020.

[3] Ma, Z., Shi, Z., Zhou, Y., Xu, J., Yu, W., and Yang, Y.: A spatial data mining algorithm for downscaling TMPA 3B43 V7 data over the Qinghai–Tibet Plateau with the effects of systematic anomalies removed, Remote Sensing of Environment, 200, 378-395, https://doi.org/10.1016/j.rse.2017.08.023, 2017.

Below are some other specific comments:

1. The authors are suggested to summarize previous published studies focusing on multi source AOD dataset fusion or AOD gap filling using different models to rich the Introduction since a lot of related work have been done.

Response: Thank you for your careful reading and valuable suggestions. Actually, we have summarized the common methods for multi-source AOD dataset fusion or AOD gap filling in the introduction (Pages 3-4/Lines:79-94). Based on your suggestion, we have further highlighted downscaling methodology, including fusion and gap filling. In addition, we have also added some specific case references and further described the implementation of the downscaling method used in this study (Pages 3-4/Lines:85-94).

2. Section 2.3: MxD08 AOD product is too coarse in the spatial resolution (1 degrees) to be used for comparison in such a small study region. I suggest using the MxD04 product with a high resolution 3 or 10 km.

Response: Thank you for your careful reading and precise advice. According to your suggestion, we select the MxD04L2 10 km data to replace the MxD08 AOD in revision, because we have found experimentally that MOD04 _3k is generated based on the dark target method, and not applicable in this study area (most of the area is almost no data). Responding modifications are given in the text (Page 7/Lines:164-178).

2.3 MODIS MxD04L2 data

MYD04L2 and MOD04L2 are the level 2 atmospheric aerosol products from Aqua and Terra respectively, where spatial and temporal resolutions are 10 km × 10 km and

1 day respectively (Chen et al, 2021b). The MxD04L2 AOD product mainly provides two algorithms, the Dark Target (DT) and Deep Blue (DB) algorithms, to retrieve global AOD distribution. Based on the MODIS Collection 6.1, we chose 550 nm combined DT and DB AOD to validate FEC AOD. It is worth noting that the Aqua and Terra launch time is different, so we can acquire MOD04L2 data from March 2000 to February 2020, but as for MYD04L2, we only acquire data from July 2002 to February 2020. All processes are realized through downloading from NOAA website (https://ladsweb.modaps.eosdis.nasa.gov/) and calculating and analyzing local computer, and main works, including geometric correction, projection conversion, image mosaics, clipping, computing daily and monthly mean of AOD, and numerical extraction, perform in MODIS Reprojection Tool (MRT) and ENVI and ArcGis software.

3. Section 2.4: Please clarify the version and level of AERONET data, and the number of the stations (also suggest adding them in Figure 1) used in the study. In addition, the author should highlight the novelty of their study and the differences compared to previous studies

Response: Thank you for your careful reading and precious advice. In revision, we have clarified the version and level of AERONET data and added them in Figure 1 (Pages 6-8/Lines:144, 186-189). In terms of the novelty and differences of this study, this is the first highest resolution AOD dataset with complete coverage of Northwest China (a typical area of scarce information or limited data applicability). Secondly, the data is generated by the current mainstream machine learning algorithms, and the performance and efficiency are reliable. Of course, based on this dataset, we have also found some new phenomena, which again validate the impact of ecological engineering on air quality in China, while raising new research questions for the southeast Taklamakan Desert. In summary, the FEC AOD effectively compensates for the deficiency and constraints of in-situ observation and satellite AOD products. Meanwhile, FEC AOD products provide a new choice for future atmosphere research in Northwest China and the ability to capture finer local information.

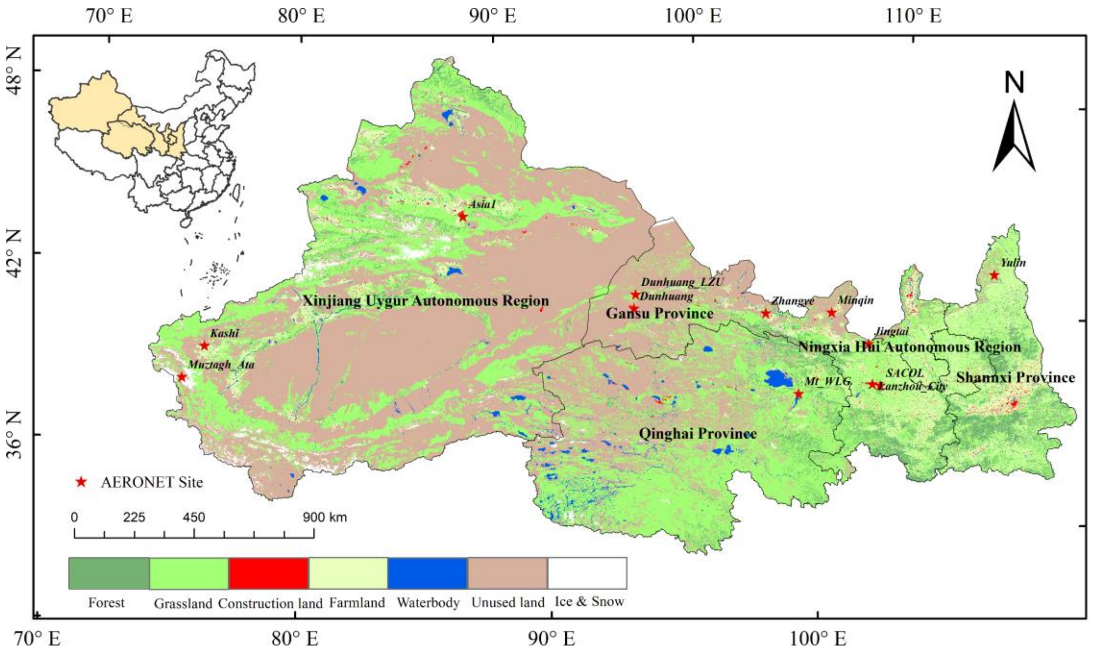

Figure 1. Study area. The figure shows typical arid and semi-arid areas, five provinces in northwest China.

4. Lines 273-288: It is not clearly how to train and validate the model. Please clarify what are the inputs to the model, and what is the real/true vaule for target? What are the training samples and verification samples? Are they independent of each other?

Response: Thank you for your careful reading. In this study, we use 12 environmental covariates (1 km) as downscaling method (bagging trees ensemble algorithms) input to acquire AOD-environmental covariates (AODe) model in 1 km and utilize AODe model and 250 m environmental covariates to acquire FEC AOD. In the modeling process, we adopt the 10 cross-validation folds as a tool for verification. That is, instead of dividing the modeling set and validation set, we repeat the cross-validation. The 10 cross-validation folds take turns using 9 of them as training data to train the model and 1 as validation data to assess model performance and calculate the average validation error over all folds. This method gives a good estimate of the predictive accuracy of the final model trained using the full data set and can effectively avoid overfitting. As a result, we have no real/true value and use the pseudo instead. This again illustrates the necessity and urgency of producing high-resolution AOD datasets in this study area, which is a data-scarce area with only a few observational data. In addition, conventional methods of dividing modeling sets and validation based on truth values usually only capture the accuracy of predictions, but the 10 cross-validation folds through repeated cross-validation not only can get the accuracy but also can get the uncertainty of the prediction, for instance, knowing where the data uncertainty is higher, and then this is where we need to strengthen the observation or build stations in the future.

5. Figure 4: I don't see much differences compared with 1km AOD, and I think the authors need show the advantages of 250m data set, e.g., may zoom in the image by looking at the AOD distributions at urban areas.

Response: Thank you for your careful reading. Section 3.1 is to validate FEC AOD accuracy based on in-situ and satellite. Figure 4 belongs to the part of comparison with satellite AOD products. So we can find out the spatial consistency from Figure 4, and more differences between FEC AOD and MAIAC 1 km AOD, which we have also considered in the original manuscript (Pages 16-17/Lines:342-345), ie. Figure S2 and

S3. We select two typical cities (Urumqi and Lanzhou) in NW China, and randomly zoom in the districts (Shaybak and Chengguan districts) respectively, and we can find the FEC AOD has a strong potential to describe local AOD features or fine AOD distribution compared with MAIAC AOD.

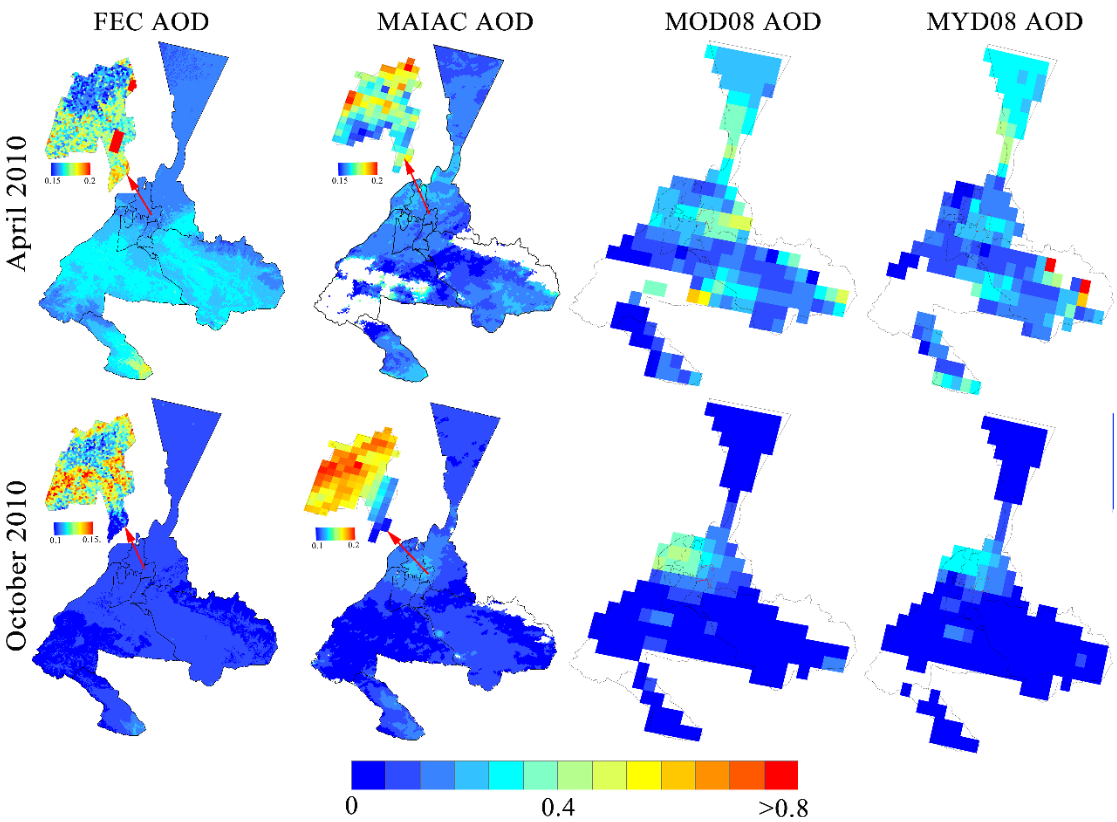

Figure S2. The spatial pattern difference of four AOD products in April 2010 and October 2011 over Urumqi.

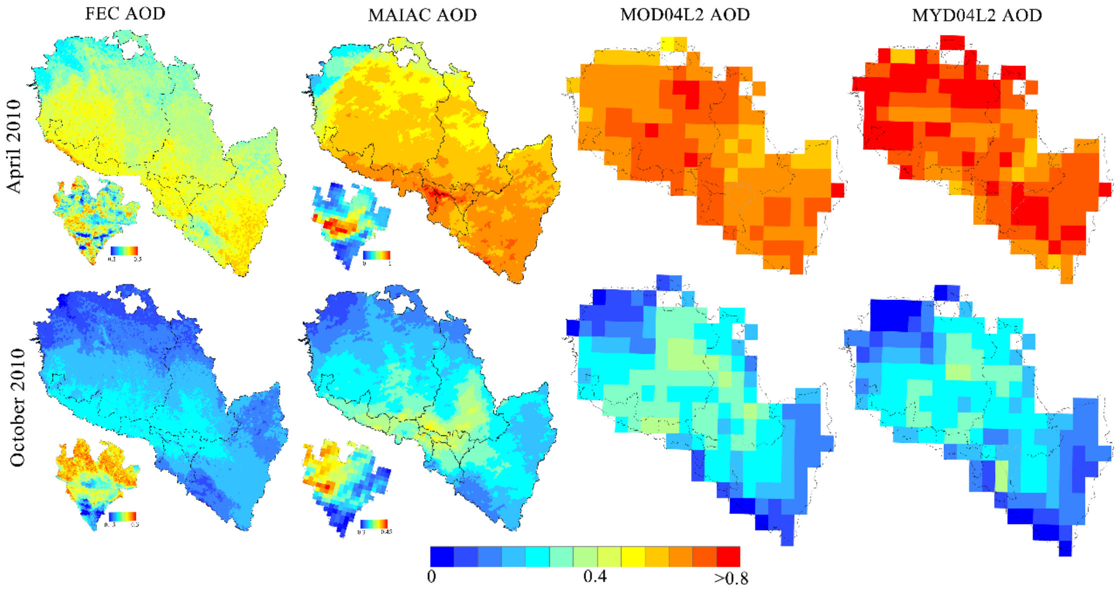

Figure S3 The spatial pattern difference of four AOD products in April 2010 and October 2011 over Lanzhou.

6. Lines 341-353: The results are pretty similar among different AOD products and difficult to distinguish the difference, and more quantitative comparison results are needed.

Response: Thank you for your careful reading. Actually, Section 3.1 theme is to validate FEC AOD accuracy based on in-situ and satellite. So we focus more on the consistency between the FEC AOD and other satellite AOD products. Surely, we also added some quantitative comparison results in the revision (Pages 17-18/Lines:361-370).

7. Sections 3.2 and 3.3: It is recommended to calculate the monthly and seasonal long-term trends and statistical significance by removing the seasonal cycles to have a look at how AOD changes throughout the study area since AOD data for nearly 20 years are available.

Response: Thank you for your valuable advice. The main objective of this study is to create a new advanced-performance, full-coverage, and high-resolution AOD dataset and validate it. In addition, we analyze the spatiotemporal pattern in Section 3.2, and the temporal variability of AOD is deeply revealed by temporal information entropy (TIE) and time-series information entropy (TSIE), so we think this has a good look at how AOD changes throughout the study area since AOD data for nearly 20 years. Of course, we also recognize your comments, which will be further explored in the next step of FEC AOD application research.