The dataset of AOD, fAOD and cAOD over Europe has application value for environment analysis. The machine learning method was used to produce daily AODs. The manuscript should be revised before considering publication.

Response: Thank you for taking the time to review our study and provide your useful feedback. We appreciate your interest in our work and are happy to hear your thoughts and address any concerns you may have regarding the dataset we have presented. We believe that our dataset has the potential to be a valuable resource for the scientific community and look forward to discussing it with you further. Thank you again for your thoughtful review.

General comments:

1 The spatial and temporal resolution of all input and output data for the machine learning should be listed.

Response: We appreciate your valuable feedback regarding the description of the spatial and temporal resolution of the data used for machine learning. We have made the necessary revisions to the description of the variable inputs from ERA5 and ERA5_land (Line 138-160):

"2.5 ERA5 reanalysis for atmospheric meteorological data

Previous studies (Huang et al., 2007; Zhou and Savijärvi, 2014; Tai et al., 2010; Gui et al., 2019; Yan et al., 2022) have analysed the associations between weather conditions and the concentration of fine- and coarse-mode aerosols. For example, high-pressure events, characterised by atmospheric stability and low winds, retain the smaller particles, which is seen with higher-than-normal fine-mode aerosol levels (Tai et al., 2010; Gui et al., 2019). Moreover, rainfall washes out the particles from the lower part of the troposphere, especially the largest particles. There are other pathways by which aerosols can also affect weather conditions, for example by reflecting and absorbing the incoming UV radiation (Zhou and Savijärvi, 2014), or by changing the conditions for the condensation of water in the cloud(Huang et al., 2007). To account for the impact of meteorological factors on aerosols, we collected data from several atmospheric variables, such as boundary layer height, downward UV radiation, cloud cover, and precipitation, from the ECMWF ERA-5 reanalysis. This dataset provides data since 1950 at a resolution of 0.25º x 0.25º. More information about the resolution and data source for each meteorological variable can be found in Table S2.

2.6 ERA5 land reanalysis for surface data

Apart from atmospheric meteorological data, the surface data also has important impacts on aerosol. As forests contribute to a large extent to particle removal, previous studies found the deposition velocity of ultrafine particles is generally more sensitive to leaf area index than leaf area density (Lin et al., 2018; Huang et al., 2015). Also, the dry deposition of particles is affected by properties of the vegetation elements (such as leaves and branches) and soil types (Grönholm et al., 2009). Thus, we found the significant contributions of leaf

area index high vegetation, leaf area index low vegetation and soil types to aerosol. Higher Leaf area index high vegetation means more evergreen trees, deciduous trees or forest, while Higher Leaf area index low vegetation represents more crops and mixed farming, grass or shrubs. For bare ground or places with no leaves, both of them will be close to zero. The soil types describe how coarse the soil is, representing the water holding ability of soil. Coarser soil generally has lower water holding ability. Additionally, land surface information is essential for surface reflectance, which further affects the quality of satellite data included in reanalysis data. Most surface-related variables, including some near-ground meteorological data, are provided by ERA5-Land at a resolution of 0.1° x 0.1°."

We have also added resolution information in Table S2.

Table S2. The list of variables used in this study

SHORT NAME	SOURCE	RESOLUTION	LONG NAME	UNIT
U10	ERA5_land	0.1°	10m u component of wind	m s**-1
V10	ERA5_land	0.1°	10m v component of wind	m s**-1
RH	ERA5_land	0.1°	Surface relatively humidity	(0 - 100)
LAI_HV	ERA5_land	0.1°	leaf area index high vegetation	m**2 m**-2
LAI_LV	ERA5_land	0.1°	leaf area index low vegetation	m**2 m**-2
MSDWSWRF	ERA5_land	0.1°	Surface solar radiation downwards	J m**-2
ASN	ERA5_land	0.1°	snow albedo	(0 - 1)
SP	ERA5_land	0.1°	surface pressure	Pa
TE	ERA5_land	0.1°	total evaporation	m
D2M	ERA5	0.25°	2m dewpoint temperature	K
T2M	ERA5	0.25°	2m temperature	K
BLD	ERA5	0.25°	boundary layer dissipation	J m**-2
BLH	ERA5	0.25°	boundary layer height	m
HCC	ERA5	0.25°	high cloud cover	(0 - 1)
TCC	ERA5	0.25°	total cloud cover	(0 - 1)
LCC	ERA5	0.25°	low cloud cover	(0 - 1)
SLT	ERA5	0.25°	soil type	1-7, higher is finer soil with stronger ability contains water
MCC	ERA5	0.25°	medium cloud cover	(0 - 1)
TCO3	ERA5	0.25°	total column ozone	J m**-2
TP	ERA5	0.25°	total precipitation	m
ALUVD	ERA5	0.25°	uv visible albedo for diffuse radiation	(0 - 1)
ALUVP	ERA5	0.25°	uv visible albedo for direct radiation	(0 - 1)
YEAR	Time	\	Year	\
DOW	Time	\	day of week	\
DOY	Time	\	day of year	\
LAT	Spatial	\	latitude	\
LON	Spatial	\	longtitude	\
NE	Minimum directional distance	0.1°	minimum distance to nearest sites in North-east direction	m
SE	Minimum	0.1°	minimum distance to nearest	m

SHORT NAME	SOURCE	RESOLUTION	LONG NAME	UNIT
	directional distance		sites in South-east direction	
SW	Minimum directional distance	0.1°	minimum distance to nearest sites in South-west direction	m
NW	Minimum directional distance	0.1°	minimum distance to nearest sites in North-west direction	m
CAMS_BC AOD550	CAMSRA	0.75° x 0.75°	black carbon aerosol optical depth 550nm	\
CAMS_DU AOD550	CAMSRA	0.75° x 0.75°	dust aerosol optical depth 550nm	\
CAMS_OMAOD550	CAMSRA	0.75° x 0.75°	organic matter aerosol optical depth 550nm	\
CAMS_SS AOD550	CAMSRA	0.75° x 0.75°	sea salt aerosol optical depth 550nm	\
CAMS_SU AOD550	CAMSRA	0.75° x 0.75°	sulphate aerosol optical depth 550nm	\
MERRA_AOD	MERRA-2	0.625°×0.5°	MERRA2 aerosol optical depth 550nm	\

Due to the different resolutions of each data, the method of spatio-temporal matching should be clarified.

To address your concern about the method of spatio-temporal matching, we have provided additional details and clarification (Line 175-180) in the revised version:

"In order to address the different spatial resolutions, we employed bilinear resampling to standardize all gridded data to a horizontal resolution of 0.1° x 0.1° (equivalent to approximately 9 km at mid-latitudes). Subsequently, we extracted the corresponding values at the grid cell where the AERONET sites located. Regarding the temporal resolution, we computed daily averages for each product. This involved utilizing all available data points for a specific day to calculate the average. For example, we used hourly data from MERRA-2, ERA5, and ERA5-land, while CAMSRA data was available at a 3-hourly resolution. The AERONET data, however, was obtained at a daily frequency."

Thank you for bringing this to our attention, and we hope that the revised description provides a clearer understanding of the spatial and temporal matching methodology employed in our study.

2 As the satellite AOD was given up, I think all the inputs are reanalysis data. So the temporal resolution of AOD, fAOD and cAOD is not necessary daily. Then, which one or some certain times in one day were selected to produce daily AOD, fAOD and cAOD? And Why?

Response: Thank you for your feedback and the question regarding the temporal resolution of AOD, fAOD, and cAOD in our study. In our study, we purposely chose to use daily averages for the AOD, fAOD, and cAOD products.

The decision to use daily averages was made based on the intended future application of this AOD product, which is to estimate ground-level PM2.5 and PM10 on a daily level.

Additionally, short-term health impact assessment studies typically focus on air pollutant exposure at daily scales, as it aligns with the daily scale of health data used in such studies.

To obtain the daily averages for each product, we generally took all available data for a given day and calculated the average. For example, we used hourly data for MERRA-2, ERA5, and ERA5-land, 3-hourly data for CAMSRA, and daily data for AERONET to obtain the daily averages for each product, we generally took all available dataset on that day to calculate it, for example, using hourly data for MERRA-2, ERA5 and ERA5 -land, 3-hourly data for CAMSRA and daily data for AERONET.

We believe that selecting daily averages provides a suitable temporal resolution for our study's objectives and future applications related to air pollution exposure and health impact assessment. We appreciate your feedback and the opportunity to clarify our approach.

3 Why chose LightGBM from kinds of machine learning methods? Decision-tree based machine learning methods would adopt some fixed thresholds, which may create systematic "boundary" in the product. For example, if the latitude was included in the input data, you can see a AOD systematic boundary at a latitude line. Other parameters has the similar affects.

Response: Thank you for your feedback regarding the choice of LightGBM as the machine learning method in our study. We appreciate your concern about decision-tree based methods potentially creating systematic boundaries due to fixed thresholds.

The decision to use LightGBM was based on several factors that make it suitable for our specific application. Firstly, LightGBM is known for its high computational efficiency, making it well-suited for handling large-scale datasets like the 18-year predictions for the whole of Europe. Its faster training and prediction times on large datasets outperform other gradient frameworks, random forests, and support vector machines (SVM).

Secondly, LightGBM incorporates a gradient-based One-Side Sampling (GOSS) technique, which helps the model prioritize important data points to capture the general pattern. This prioritization, along with early stopping and regularization techniques, helps prevent overfitting and ensures good generalization performance.

Furthermore, the LightGBM framework offers a wide range of customizable APIs and parameters. This allows us to customize our own loss function and to obtain quantile predictions, which are useful for improving the model structure and estimating uncertainty in our predictions.

Regarding the concern about systematic boundaries, it is true that a single decision tree can create clear boundaries in predictions based on fixed thresholds. However, both Random Forest and LightGBM ensemble multiple decision trees, which helps to blur these

boundaries. By averaging the predictions of multiple trees or using gradient boosting to correct errors, the ensemble methods can reduce the impact of fixed thresholds and produce more flexible decision boundaries. Additionally, considering more variables and interactions can enable the model to capture more complex patterns and reduce the influence of fixed thresholds. As shown in Figure 11 of our study, we observed that no systematic "boundary" is prominent in our product. In the future, we plan to explore kernel-based machine learning methods as well to further mitigate the potential systematic boundary problem.

4 The spatial distribution, I am not sure if it means some AERONET sites data were not used in training, and only used in test? If so, that's real spatial independent validation. If not, we can not give the accuracy over locations which has no AERONET site.

Response: Thank you for your feedback and the question regarding the spatial distribution and validation of our model. We appreciate your concerns and the opportunity to clarify the validation process. We have revised the manuscript (Line 196-198) and (Line 220-227) to provide a more detailed explanation.

In our study, we conducted two validation processes to assess the performance of the model. Firstly, we randomly selected 70% of the AERONET sites as training data for the quantile LightGBM models. An additional 20% of the sites were used to optimize the model, and the remaining 10% of the sites served as completely independent test data. Table S1 in our study presents the results, showing that the R-squared values for the independent test sites are 0.72, 0.69, and 0.70 for AOD, fAOD, and cAOD, respectively.

The second validation process involved using 5-fold cross-validation, which repeated the first process multiple times to test the stability and consistency of the model configurations. Table S1 also presents the results for the cross-validating test sites, indicating R-squared values ranging from 0.68 to 0.74 for AOD, 0.65 to 0.73 for fAOD, and 0.68 to 0.74 for cAOD. These values are similar to the results obtained in the first process.

Furthermore, Table S3 in our study compares the results between randomly selecting test sites and using the top 20% of sites that are farthest from their nearest neighbors as test sites. The small differences observed between these two situations further indicate the robustness of our model, even in locations far away from AERONET sites.

In summary, our validation processes and results demonstrate that the framework of our model is robust and capable of providing accurate predictions even in locations without AERONET sites. We appreciate your feedback and the opportunity to clarify the validation procedures.

Minor comments:

1 The abbreviation should be explained at the first appearance, such as "NMB" in the supplement.

Added.

2 The section numbers are wrong in chapter 4.

Revised.