https://doi.org/10.5194/egusphere-2026-1987
Preprint. Discussion started: 4 June 2026
© Author(s) 2026. CC BY 4.0 License.

Twin Eyes in the Sky: Deep Learning-Based AOD Enhancement Using GOES-East and GOES-West

Andrea Porcheddu[1], Ville Kolehmainen[1], Timo Lähivaara[1], Robert Levy[2], Yingxi Rona Shi[2,3], Hai Zhang[4], and Antti Lipponen[5]

[1]Department of Technical Physics, University of Eastern Finland, Kuopio, Finland
[2]NASA Goddard Space Flight Center, Greenbelt, MD, USA
[3]University of Maryland Baltimore County, Baltimore, MD, USA
[4]Science and Technology Corporation at NOAA, College Park, MD, USA
[5]Finnish Meteorological Institute, Atmospheric Research Centre of Eastern Finland, Kuopio, Finland

Correspondence: Antti Lipponen (antti.lipponen@fmi.fi)

Abstract. High spatio-temporal resolution aerosol monitoring is critical to understand and mitigate air pollution and climate change. In this context, geostationary satellite instruments can be extremely beneficial, allowing fine-grained temporal characterization of aerosols over large regions. In this study, we combine data from the geostationary instruments Advanced Baseline Imager (ABI) on-board GOES-East and GOES-West, using Deep Learning methods to post-process NASA Dark Target ABI AOD and NOAA ABI AOD products and improve their accuracy and spatial resolution. We deploy a Transformer Encoder architecture, and compare it to a Multi Layer Perceptron (MLP) architecture predicting at single time step, showing how exploiting the temporal patterns in geostationary daily observations leads to improved accuracy and generalization in the post-process correction. Additionally, we show that further improvement can be obtained combining multi-view angles from different (though very similar) geostationary satellites. Our region of interest is the Contiguous United States (CONUS) in the years 2020-2022.

1 Introduction

Atmospheric aerosols, also referred as Particulate Matter (PM), are liquid and solid particles suspended in the atmosphere. Monitoring these particles with high spatial-temporal resolution is critical to understand their impact on air quality studies (Thangavel et al., 2022) and climate change (Haywood and Boucher, 2000). While remote sensing instruments from ground stations allow very accurate and frequent measurements (e.g. Aerosol Robotic Network (Holben et al., 1998)), satellite instruments promise to fill spatial gaps, widening the monitoring spatial coverage. Different satellite instruments have different spatial and temporal resolution: in particular, satellite instruments have typically spatial resolution ranging from tens of meters to kilometers. Sun-synchronous orbiting satellites (such as Aqua and Terra) observe the earth from different positions at different times of the day, offering temporal resolution of the order of one day. On the other hand, satellites on geo-synchronous orbit (such as GOES and Himawari) permit to observe the same location multiple times per day, enabling sub-hourly measurements (Laszlo et al., 2008). Aerosol Optical Depth (AOD) estimates the aerosol loading on the vertical column and is one of the

https://doi.org/10.5194/egusphere-2026-1987
Preprint. Discussion started: 4 June 2026
© Author(s) 2026. CC BY 4.0 License.

main quantities used to characterize aerosols from satellites. Classical satellite AOD retrieval algorithms (such as NASA Dark Target (Kaufman et al., 1997; Remer et al., 2005; Levy et al., 2007) and NOAA Enterprise (Laszlo et al., 2008; Kondragunta et al., 2020)) are based on physical radiative transfer models and estimate the aerosol loading in the air column after excluding the molecular scattering, gas absorption and estimated surface contributions. They are commonly based on a Look Up Table (LUT) approach and essentially match the radiances measured by satellite instruments to pre-calculated values corresponding to specific aerosol conditions (e.g. AOD, aerosol size distribution, observation geometry). This approach is computationally cheaper than running a full Radiative Transfer Model (RTM) for each retrieval, but comes with compromises on resolution and accuracy, considering also the uncertainty related to the approximated physical model behind.

In the recent years, Machine Learning (ML) and Deep Neural Networks (DNNs) were proven to be powerful tools for multiple tasks (such as regression and classification) and applications. Several studies applied ML to Earth Observation (EO), and in particular to satellite AOD retrievals. For example, assuming that land type data contains information regarding surface reflectance and aerosol type, Lary et al. (2009) developed ML models to bias-correct MODIS AOD. Similarly, Lanzaco et al. (2016) utilized meteorological variables to bias-correct satellite AOD at AERONET stations locations in South America. Nguyen et al. (2012) showed how ML models could be used to downscale existing satellite AOD products to a higher spatial resolution. Utilizing the radiances measured by a satellite instrument as input to a ML model, one could directly try to predict AOD using ground stations AOD as target. This direct approach would require, for example, the network to separate the aerosol signal from molecular scattering and absorptions, as well as clouds, surface and other signals. Lipponen et al. (2021, 2022) revealed that utilizing radiances and satellite AOD product variables as input to a ML model can improve AOD accuracy compared to a fully learned approach (direct AOD retrieval from radiances). Adding auxiliary variables with high spatial resolution, Taskinen et al. (2022) confirmed that this difference in accuracy holds also in AOD downscaling applications. From the air quality studies point of view, Just et al. (2018) indicated that applying a ML bias-correction can significantly improve the correlation between AOD and surface PM, suggesting its relevance in AOD to surface PM conversion. Zhang et al. (2025) utilized GEMS geostationary satellite NO_2 data to forecast NO_2 concentrations over eastern China, deploying a deep learning model to handle spatio-temporal patterns simultaneously. Sayeed et al. (2025) used a DNN model to correct bias of PM2.5 estimates using geographically weighted regression (GWR) from combined GOES-East and GOES-West AOD retrievals (Zhang and Kondragunta, 2021; Zhang et al., 2022).

In this study, we combine GOES-16 (GOES-East) and GOES-17 (formerly GOES-West) Advanced Baseline Imager (ABI) data to post-process correct and downscale NOAA ABI and NASA Dark Target AOD products. Instead of selecting separate regions of validity for GOES-16 and GOES-17 data, we combine the data on a common grid, and we show that simultaneously using features from both satellites improves the post-process correction results compared to correcting each instrument separately. To achieve our goal, we deploy a Transformer architecture and show that exploiting the temporal patterns in the data leads to increased accuracy. AERONET ground stations AOD is utilized to train and test the model. Our region of interest is the Contiguous United States (CONUS) from January 2020 to December 2022.

https://doi.org/10.5194/egusphere-2026-1987
Preprint. Discussion started: 4 June 2026
© Author(s) 2026. CC BY 4.0 License.

2 Data

The different data sources presented in this section have been projected on a common grid with cell size 500 m x 500 m by the Albers Equal Area projection. The temporal collocation is done within a window of 2 minutes and 30 seconds respect to NOAA GOES-16 ABI Level 1b. AERONET stations are collocated with the closest pixel on the grid, within the same window of 2 minutes and 30 seconds respect to NOAA GOES-16 ABI Level 1b. Part of the NOAA GOES ABI Level 1b and 2 data retrieval have been carried out using the GOES-2-go Python package (Blaylock, 2023). Both training and testing of the models is done at single pixel (i.e. we don't use spatial averaging). A complete list of variables is presented in Appendix A.

2.1 GOES-R ABI data products

Hosting twin suites of sensors, GOES-16 and GOES-17 are two geostationary satellites, that were coincidentally operational as GOES-East (longitude 75.2° West) and GOES-West (137.3° West) during our study period. From each GOES platform the Advanced Baseline Imager (ABI) provided atmospheric measurements measuring radiances in 16 spectral bands. In this study, we consider only the reflective bands (approximate center wavelengths 0.47, 0.64, 0.86, 1.37, 1.6, 2.2 μm). The Full Disk views offer radiance measurements with 10 minutes temporal resolution and different spatial resolutions depending on the spectral band (1 km, 500 m, 1 km, 2 km, 1 km, 2 km) (GOES-R Calibration Working Group and GOES-R Series Program, 2017).

The NASA Dark Target GOES-R ABI Level 2 AOD product provides AOD at 4 different wavelengths, with 10 km spatial resolution and 10 minutes temporal resolution (NASA, 2019a, b). The product is available for both land and ocean, and takes advantage of bright aerosols that contrast against a dark surface to separate the aerosol and surface signals (Kaufman et al., 1997; Remer et al., 2005; Levy et al., 2007). When surfaces are too bright (e.g. sun glint over ocean, or desert and snow over land) retrievals are not attempted. The algorithm uses a LUT approach and different combinations of aerosol fine modes and coarse modes in order to match satellite TOA radiances. Together with AOD retrievals, the product provides other information such as aerosol fine and coarse mode fraction, surface reflectances, cloud fraction, and data quality flag.

The NOAA GOES-R ABI AOD product provides AOD at 550 nm over land and ocean, with 2 km spatial resolution and 10 minutes temporal resolution (GOES-R Algorithm Working Group and GOES-R Series Program, 2017). The retrieval is performed using different algorithms for land and ocean (Laszlo et al., 2008; Kondragunta et al., 2020). In particular, the land retrieval algorithm can be further split into two algorithms. Retrievals over dark surfaces are estimated using a Dark Target approach (Kaufman et al., 1997; Remer et al., 2005; Levy et al., 2007). Retrievals over bright snow-free land surfaces are estimated applying a static 0.1°x0.1° spectral surface reflectance ratio dataset to characterize the surface contribution (Laszlo and Liu, 2020). However, the bright surface retrieval was not implemented for the version in the period of this work. The product includes data quality information providing an assessment of the quality of the algorithm retrievals, which has three categories: high, medium and low. It is recommended to use high quality retrieval for quantitative use. However, we use all the qualities in this work to show that we are able to improve the accuracies of medium and low quality AOD retrievals using our post-process correction approach.

https://doi.org/10.5194/egusphere-2026-1987
Preprint. Discussion started: 4 June 2026
© Author(s) 2026. CC BY 4.0 License.

The complementary views of GOES-16 and GOES-17 are combined on a common grid with cell size equal to the highest resolution of the different spectral bands (500 m), using nearest-neighbour interpolation and the Albers Equal Area projection. We utilize ABI Level 1b radiances, together with Level 2 AOD (from both the NASA DT and NOAA Enterprise products) as inputs for our models. In the data pre-processing, ABI TOA radiances are converted to TOA reflectances before being used as input in our model. Level 2 AOD is used as input to calculate the post-process corrections, and in the output to calculate the final post-process corrected AOD retrieval. Quality flags from Level 2 AOD products are hot-encoded and fed as input to our models. For NASA DT dataset, AOD at 550 nm is used for the post-process correction.

2.2 AERONET AOD

The Aerosol Robotic Network (AERONET) (https://aeronet.gsfc.nasa.gov/) is a global network of Sun photometers and is considered the gold standard for ground-based AOD retrievals (Giles et al., 2019; Lenoble et al., 2013). Ground stations belonging to the network can observe radiation at different wavelengths, with all instruments sharing the 1020, 870, 675 and 440 nm channels. AERONET AOD uncertainty varies from ± 0.01 to ± 0.02, with spectral dependence and highest error in the ultraviolet domain (Eck et al., 1999). In this work, AERONET AOD has been used as the ground truth in post-process correction of GOES ABI AOD. AERONET AOD at 500 nm has been transformed using the Ångstrom law to obtain AOD at 550 nm: this 550 nm AOD is used as target to train and test our models. Level 2.0 AERONET AOD has been used, meaning the AOD data is cloud cleared and quality assured with calibration applied. An AERONET station in Tucson (Arizona) has been removed to avoid a bias in the hyperparameter search.

2.3 Geometry and time variables

We calculated View Zenith Angle (VZA), View Azimuth Angle (VAA), Sun Zenith Angle (SZA) and Sun Azimuth Angle (SAA), utilizing the Astropy Python library (Astropy Collaboration et al., 2013, 2018, 2022). Using this library we could calculate the angles on each pixel of our grid (500m x 500m pixel size). To do so, the position (expressed as latitude and longitude) on earth, the datetime, and the elevation (ASTERDEM) information have been used. We encoded the day of the year (doy) using two cyclical variables:

$$\text{doy\_cos} = \cos\left(2\pi \cdot \frac{\text{doy}}{365}\right), \quad \text{doy\_sin} = \sin\left(2\pi \cdot \frac{\text{doy}}{365}\right) \tag{1}$$

After converting the time of the day from UTC to local time, we encoded it as:

$$\text{time\_of\_day} = \frac{\text{hour} + \frac{\text{minute}}{60} + \frac{\text{second}}{3600}}{24} \tag{2}$$

The spatial geometry and temporal information are used as inputs in our model.

https://doi.org/10.5194/egusphere-2026-1987
Preprint. Discussion started: 4 June 2026
© Author(s) 2026. CC BY 4.0 License.

2.4 ASTER DEM

The Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) is an instrument onboard the Terra satellite. The ASTER DEM (Digital Elevation Model) provides earth surface elevation data at the spatial resolution of approximately 30 m (Fujisada et al., 2011, 2012; NASA/METI/AIST/Japan Spacesystems, and US/Japan ASTER Science Team, 2019). In this work, ASTER DEM elevation data have been used as input to calculate observation geometry variables (together with datetime, longitude and latitude), such as viewing angles and sun angles. Elevation data is not directly used as input in our models.

3 Methods

We use a Multi Layer Perceptron (MLP) architecture as our baseline. The MLP architectures learns and predicts the post-process correction at single time steps, without having information from the rest of the daily geostationary time series. We compare this baseline to a Transformer architecture, which is capable of exploiting the temporal correlations in the time series to post-process correct the original AOD products. The two architectures are described in Appendix C. As represented in Fig.1, when considering the combination of GOES-16 and GOES-17 data, the model uses data from both satellites as input, and returns 2 outputs corresponding to the respective post-process corrections (instead of 1 output as for the single satellite models). The deep learning pipeline has been implemented using the PyTorch framework (Paszke et al., 2019).

3.0.1 Spatial-temporal Cross-Validation

To mitigate spatial overfitting in model training and assess spatial generalization in model testing, we clustered the AERONET ground stations using the DBScan algorithm (Ester et al., 1996), considering a radius of 25 km (the distance is calculated from latitudes and longitudes using the Haversine formula) and the minimum samples parameter equal to one. This ensures that 2 ground stations closer than 25 km are grouped in the same graph/cluster (at this time, altitude differences have not been considered in the clustering and could be object of future work). Furthermore, the region of interest, the CONUS, is divided into 16 non-overlapping regions. We define 5 folds, and separately for each region, we assign the clusters falling into the region to these 5 folds, following a distance-based sampling (i.e. the closer two clusters are, the less likely they are to be sampled into the same fold). If the region contains less than 5 clusters, the region is merged with another region before sampling the cluster to generate the folds. Since the collocation with AERONET AOD results in different collections of data samples for NOAA AOD and NASA Dark Target respectively, this procedure has been applied separately for the two datasets giving different spatial cross-validation (CV) splits (the results obtained for the two datasets are therefore not directly comparable and comparing them is not the objective of this study). The result of this procedure can be visualized in Fig.2 when considering the collocation with NOAA AOD. The 5 spatial folds are used to generate 5 combinations of training (3 folds), validation (1 fold) and test (1 fold) datasets.

https://doi.org/10.5194/egusphere-2026-1987
Preprint. Discussion started: 4 June 2026
© Author(s) 2026. CC BY 4.0 License.

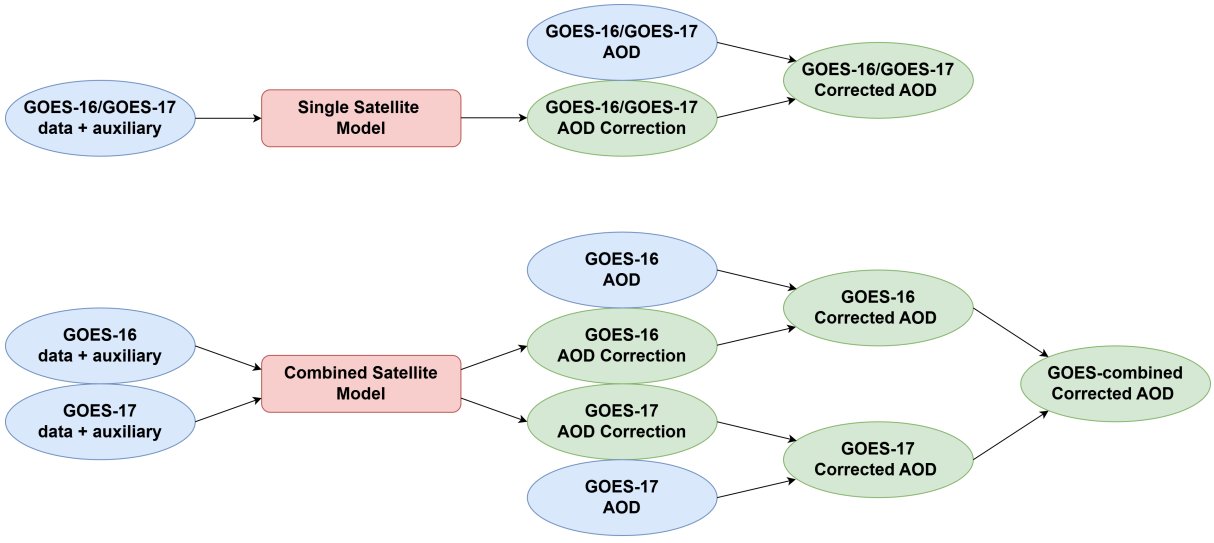

Figure 1. Schematic of models using single satellite data and a combination of GOES-16 and GOES-17 satellite data. Single satellite models consider only one satellite both in input and output: the output is one correction to a specific satellite product (for GOES-16 or GOES-17). The model combining GOES-16 and GOES-17 data, takes input data from both satellites and has two outputs, corresponding to two different corrections for the two products from GOES-16 and GOES-17. The corrected products are finally combined by averaging. For both models, the final prediction is clipped below, ensuring the minimum post-processed AOD value is non-negative.

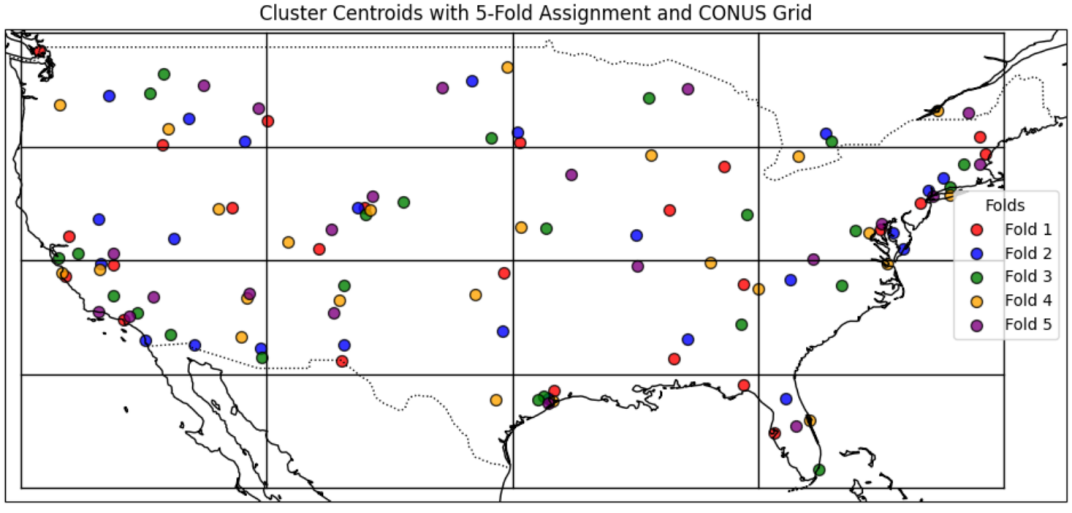

Figure 2. AERONET ground stations clusters across the CONUS region collocated with the NOAA ABI AOD dataset. The clusters area randomly assigned to spatial folds in our Cross-Validation approach based on their respective distance.

To mitigate temporal overfitting in model training and assess temporal generalization in model testing, we consider 6 different temporal splits, using one year for training, one year for validation, and one year for testing in each split. Combining spatial and temporal splits, we finally end up with 5 x 6 = 30 spatio-temporal CV splits. The splits are represented in Fig.A1.

https://doi.org/10.5194/egusphere-2026-1987
Preprint. Discussion started: 4 June 2026
© Author(s) 2026. CC BY 4.0 License.

3.0.2 Loss Function

To estimate the training and validation error, we utilize a Mean Square Error (MSE) loss function. When training the MLP model, every input corresponds to available target data. This is not always the case with the transformer model. In a time series, we can have time steps with available satellite AOD but missing AERONET AOD. Instead of filtering out useful information from the satellite data, we keep all the available input data and mask the target data when it is not available. In this way, only the available ground stations data is considered when estimating the training and validation errors, while all the available input data is used for the prediction (i.e. the model predicts satellite AOD also when AERONET target data is partially not available in the training).

3.0.3 Model training

For each deep learning architecture considered (MLP and transformer), we used the temporal CV split 1, 3, and 5 to perform hyperparameter tuning. The hyperparameter tuning is performed separately for each spatial CV split. Hyperparameters chosen on temporal CV split 1 are used on temporal CV split 1 and 2. Analogously hyperparameters chosen on temporal CV split 3 are used on temporal CV split 3 and 4 (and so on). In this way we avoid using a test set in the hypeparameters choice in any split. The hyperparameters were tuned using a uniformly random selection: for every model and CV split, 50 random hyperparameters combinations have been tried out, always selecting the combination resulting in the best validation error. The hyperparameters tuned and their respective exploration space are represented in table C1 and C2. The choice of tuning the hyperparameters separately on each spatial split can be seen as simulating real world conditions: when training the model for deployment, our cross-validation procedure should justify the entire methodology, and not just specific hyperparameters. The early stopping technique (Goodfellow et al., 2016; Bishop and Bishop, 2024) is used to stop the training when the validation loss has not decreased further after 30 epochs (i.e. patience = 30). The parameters with the best validation loss are selected. To train the model parameters we utilized the Adam optimizer (Kingma and Ba, 2017; Bishop and Bishop, 2024).

Specifically for the transformer models, we utilized a learning rate schedule combining warm-up and decay (Bishop and Bishop, 2024). Warm-up is performed for 10 epochs with a linear schedule, while the following decay is implemented with a cosine schedule (Loshchilov and Hutter, 2017).

Since the transformer uses time series as input, and time series can naturally have missing data, we apply temporal data augmentation to improve the robustness of the model. In our implementation, half of the training batch is randomly undersampled temporally (i.e. time series steps are removed from inputs and targets) at each training step.

After hyperparameter tuning, the hyperparameters selected are used to train new models on all the temporal CV splits.

4 Results

Fig.3 shows our main result for both NOAA ABI AOD and NASA DT ABI AOD: combining ABI data from GOES-16 and GOES-17 improves over using a single satellite instrument, and the best result is obtained exploiting the temporal correlation

https://doi.org/10.5194/egusphere-2026-1987
Preprint. Discussion started: 4 June 2026
© Author(s) 2026. CC BY 4.0 License.

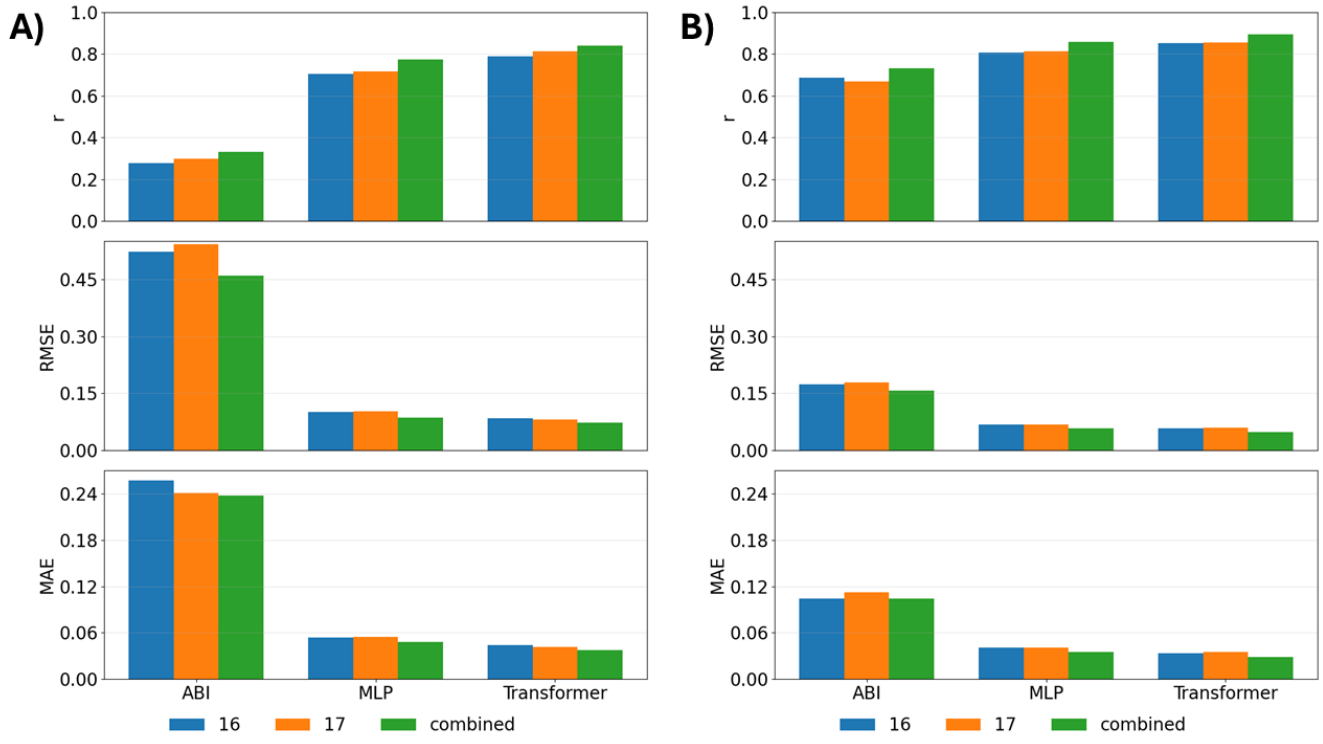

Figure 3. Comparison of original ABI AOD, its post-process correction using MLP and transformer models. Panel A) represents the results for NOAA ABI AOD, while B) represents the results for NASA DT ABI AOD. All data quality flags are included. The metrics have been calculated respect to AERONET AOD ground stations data: they are obtained averaging the metrics obtained on different spatio-temporal CV splits test sets. Different colors represent metrics from different model configurations: AOD obtained from GOES-16 (blue), AOD obtained from GOES-17 (orange) and AOD obtained from the combination of the GOES-16 and -17 (green).

in the geostationary data (in this case using a transformer architecture). The metrics presented here are the Pearson correlation coefficient (r), the Root Mean Square Error (RMSE) and the Mean Absolute Error (MAE). They are calculated separately on each spatio-temporal CV split and then averaged to obtain the final summary metrics displayed in Fig.3. The metrics calculated for ABI AOD in the "combined" configuration are obtained after averaging directly AOD from GOES-16 and GOES-17. All the data quality flags are included, which explains the poor performance before the correction. Although Fig. 3 shows qualitatively similar improvements for both NOAA and NASA datasets, it is important to notice that both the spatial resolution and number of collocations with AERONET AOD is different for the two datasets: NASA DT product is a 10-km aggregation with pixel selection built in, while NOAA's product is a 2-km pixel level retrieval. The two datasets are therefore not directly comparable. Although for NASA DT the relative improvement is smaller than for the NOAA dataset, the general result still holds.

Fig.4 breaks our main result into details, showing the metrics obtained on the different spatio-temporal CV splits when considering NOAA ABI AOD. The dots correspond to averaged metrics across different temporal splits, while the shadows

https://doi.org/10.5194/egusphere-2026-1987
Preprint. Discussion started: 4 June 2026
© Author(s) 2026. CC BY 4.0 License.

represent the metrics standard deviation. Each dot is calculated separately for different spatial split, model and satellite configuration. Fig.B1 shows the same results obtained with NASA DT AOD. Both the MLP and transformer architecture improve considerably the AOD retrievals compared to the original AOD products. The transformer improvement is generally better than the MLP improvement, and considering the correlations in the geostationary time series seems to consistently improve model generalization.

The same metrics for NOAA ABI AOD are shown in a different fashion in Fig.5, where we explictly compare the results obtained by different configurations of our models (using only GOES-16 data, only GOES-17 data, and combining GOES-16 and GOES-17 data). An analogue figure for DT ABI AOD can be found in the appendix (Fig.B2). We are looking again at different metrics (r, RMSE, MAE) obtained on different spatio-temporal CV splits. For both AOD products, combining data from GOES-16 and GOES-17 improves the AOD retrievals post-process correction. The combination seems to generally improve also the generalization of the model across the spatio-temporal CV splits. Notice that these plots are considering all quality flags for the different data products: the observed variability in the metrics among different CV splits is reduced to expected levels when considering the best AOD qualities.

We can take a look at scatterplots comparing NOAA ABI AOD (and its post-process corrections) to AERONET AOD in Fig.6. These scatterplots aggregate data points from all the spatio-temporal CV splits, while the metrics are averages of metrics obtained on single CV splits (they are the same as in Fig.3). Again, all data quality flags are included.

The post-process correction generically reduces the overestimation of the original product at low AOD levels. Combining data from GOES-16 and GOES-17 seems to help further in reducing this bias, especially when looking at the results obtained with the MLP architecture. Combining data from the two satellites produces also better estimates for high AOD values. The underestimation at high AOD level common to all the post-process correction model has probably to be attributed to a mismatch in the distributions of high AERONET AOD levels on the different CV splits, and would be less important when training a single model on a larger dataset with our methodology. Fig.B3 presents the same scatterplots when considering NASA DT ABI AOD compared to AERONET AOD. The post-process correction generally attenuates the overestimation of the original product. Specifically combining GOES-16 and GOES-17 data with the transformer architecture seems to produce better results all over the AERONET AOD values range. All the models show underestimation at high AOD levels, but this may be attributed to the scarcity of such data points in our training data compared to the NOAA ABI AOD dataset.

To assess the performance of our models at different locations across the CONUS region, we calculated evaluation metrics at single stations as show in Fig.7. This figure specifically refers to RMSE metrics for the post-process correction of the NOAA ABI AOD product compared to AERONET AOD. These metrics (differently from the others shown previously) are calculated at each station merging all the data points from different temporal CV splits. The ground stations locations (represented by the dots locations on the maps) are representative of all the spatial CV splits. The general improvement of the post-process correction is evident, although some locations remain still problematic. For example, the comparison with the Railroad_Valley2 AERONET station in Nevada is characterized by a RMSE larger than 0.2 for all the models, while the MLP-based postprocessing at the Granite Island station in Michigan gives worse results than the original NOAA product. Further studies will be needed to better understand these discrepancies and understand the limitations of the approach. Larger errors on the western

https://doi.org/10.5194/egusphere-2026-1987
Preprint. Discussion started: 4 June 2026
© Author(s) 2026. CC BY 4.0 License.

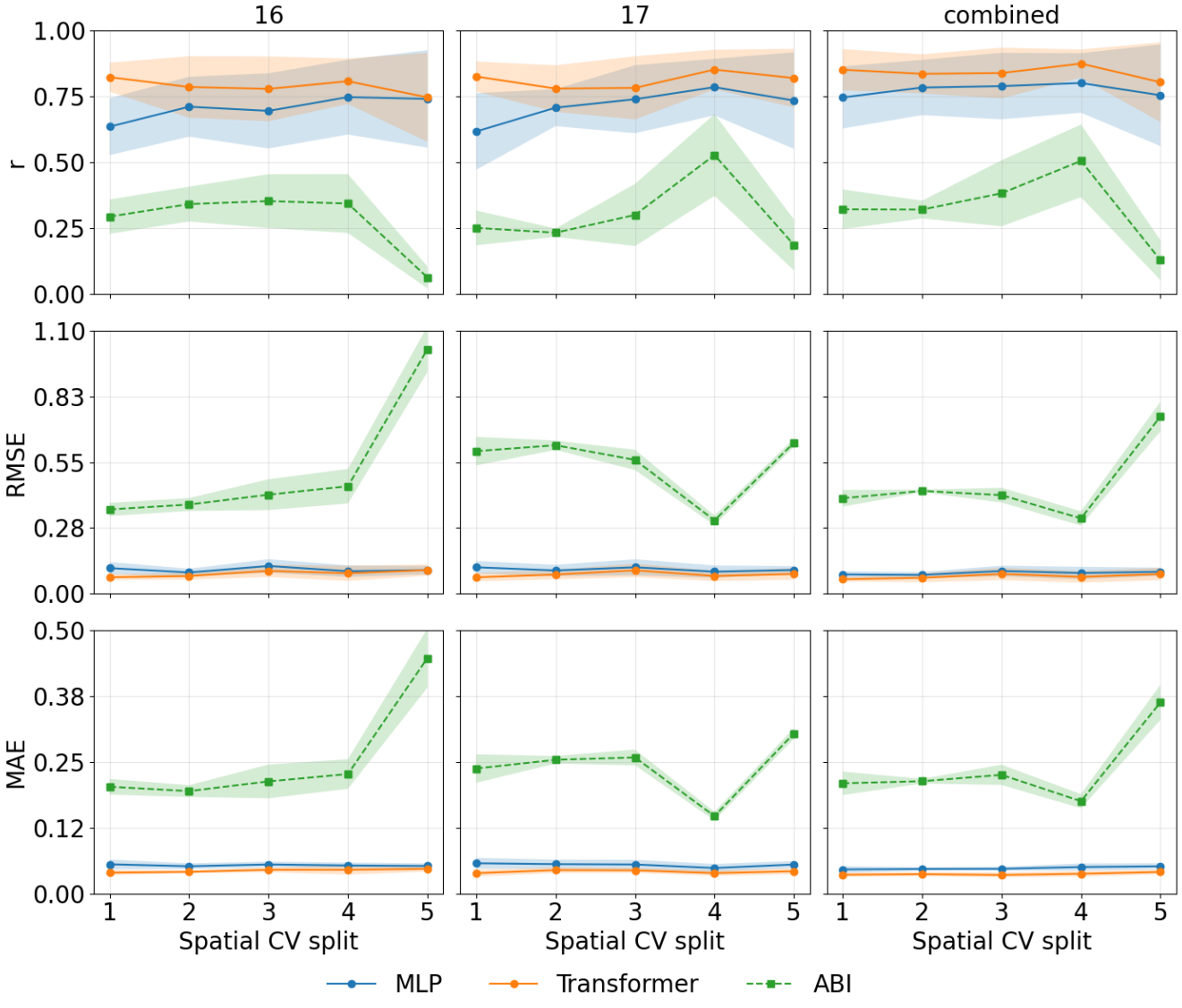

Figure 4. Comparison of metrics obtained on different spatio-temporal CV splits, for different models and satellite configurations. The metrics have been calculated in respect to AERONET measurements and NOAA ABI AOD data (and related corrections). Each row represent a different metric (correlation, RMSE, MAE), while each column represent a different satellite configuration (GOES-16, GOES-17, combined). Dots correspond to averages of metrics across temporal splits, while shadows correspond to relative standard deviations. In each panel, ABI AOD product (green), MLP AOD (blue) and Transformer AOD (orange) are compared.

225 region are attenuated by combining the satellites data, even more when considering the transformer architecture. This area is generically more problematic for AOD satellite retrievals, due to orographic and land type differences compared to the eastern region, making the surface reflectance contribution in AOD retrievals more difficult to estimate. This phenomenon can also be

https://doi.org/10.5194/egusphere-2026-1987
Preprint. Discussion started: 4 June 2026
© Author(s) 2026. CC BY 4.0 License.

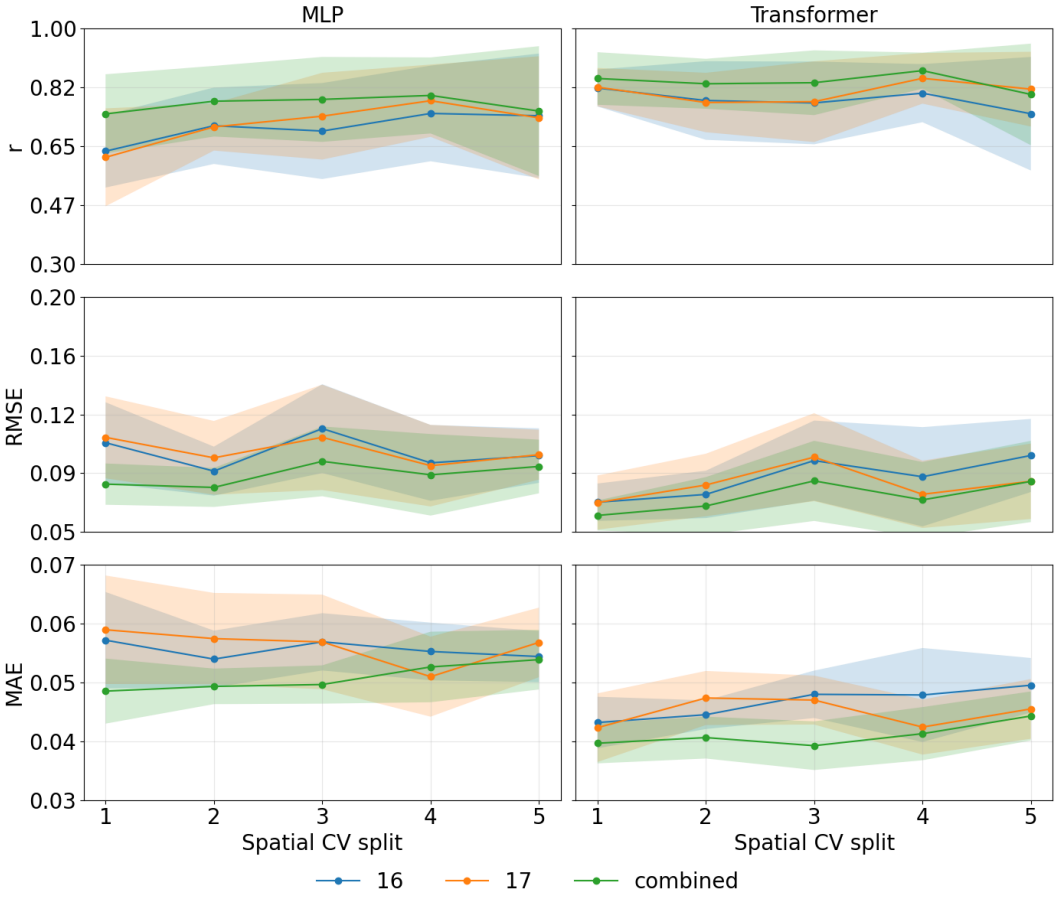

Figure 5. Comparison of metrics obtained on different spatio-temporal CV splits, for different models and satellite configurations. The metrics have been calculated in respect to AERONET measurements and NOAA ABI AOD data (and related corrections). Each row represent a different metric (correlation, RMSE, MAE), while each column represent different models used for the post-process correction (MLP, Transformer). Dots correspond to averages of metrics across temporal splits, while shadows correspond to relative standard deviations. In each panel, GOES-16 model (blue), GOES-17 model (orange) and GOES-combined model (green) are compared.

seen when looking at Fig.7 for the NASA DT ABI AOD post-process correction: the difference in error between the western and the eastern area of the maps is here particularly evident (due to the choice in the colormap scale) when looking at the original product RMSE metrics. The same considerations made for the NOAA dataset remain valid, although in this case the difference between the MLP and transformer post-process correction is less pronounced.

Fig.8 represent bias boxplots comparing satellite AOD to AERONET AOD. In this case we evaluate the bias while varying the land cover type for the NOAA ABI AOD dataset. We consider the IGBP classification (a list of the land cover types can be found in the appendix section A) and the MODIS land cover type dataset (Sulla-Menashe and Friedl, 2018; Loveland and Belward, 1997; Belward et al., 1999). The most noticeable bias in the original product seems to come from evergreen needleleaf

https://doi.org/10.5194/egusphere-2026-1987
Preprint. Discussion started: 4 June 2026
© Author(s) 2026. CC BY 4.0 License.

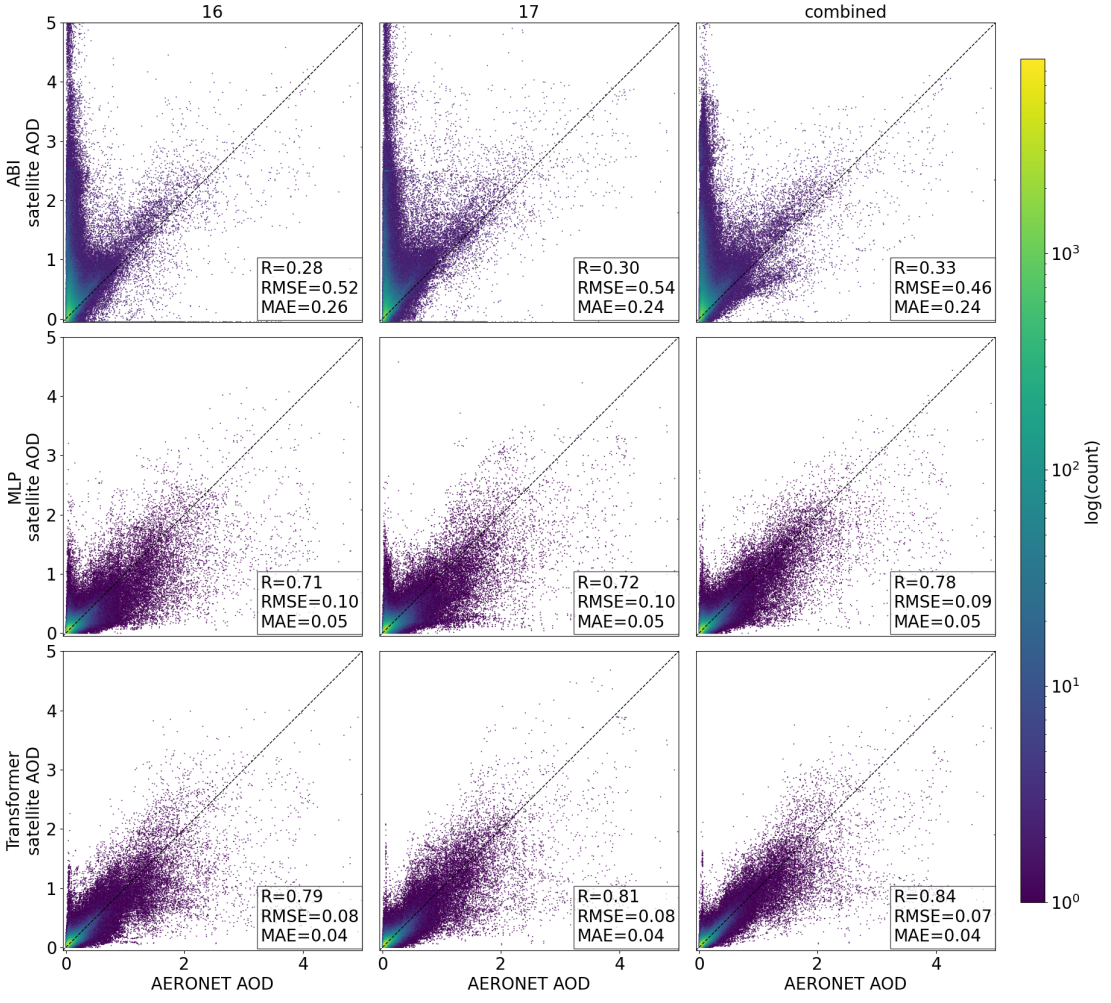

Figure 6. Scatterplots comparing NOAA ABI AOD (and its post-process corrections using MLP and transformer) and AERONET ground stations data. On the rows we distinguish the original NOAA ABI AOD product and its post-process corrections, while on the columns we distinguish the different configurations considered (using only GOES-16 data, only GOES-17 data, or a combination of the two). While the scatterplots contain points from all the CV splits test sets (N = 1546018), the metrics are calculated averaging the metrics obtained on the CV splits test sets (same metrics as in fig.3).

https://doi.org/10.5194/egusphere-2026-1987
Preprint. Discussion started: 4 June 2026
© Author(s) 2026. CC BY 4.0 License.

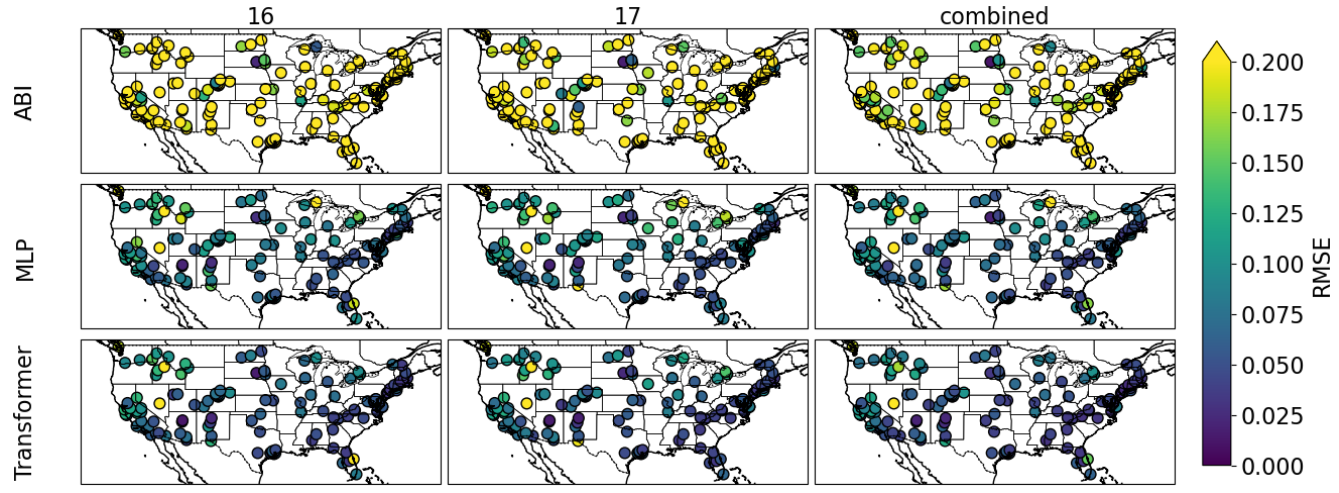

Figure 7. Comparison between NOAA ABI AOD (and its post-process corrections using MLP and transformer) and AERONET ground stations data, at single ground stations over the region of interest. The color of the dots represent the RMSE metric at each station location: these metrics have been calculated directly on all temporal CV splits test sets (not averaging metrics on different temporal splits test sets). Stations from all spatial CV splits test sets are included. Rows represent the different models (NOAA ABI AOD and its post-process corrections using MLP and transformer), while columns represent different configurations (model using only GOES-16 data, only GOES-17 data, or a combination of the two)

.

forests (land type 1), barren (land type 16), and water bodies (land type 17). The bias is greatly reduced by the post-process corrections, especially when considering the combination of satellite instruments with a Transformer architecture. Fig.B5 in the appendix shows an analogous figure for the post-processing of NASA DT AOD. In this case most of the bias seems to come from open shrublands (land type 7), urban and built-up (land type 13), and barren (land type 16). Again, the post-process correction combining the satellites with a Transformer architecture noticeably improves the result. Please notice that the y-axis scale for the two figures 8 and B5 is different, and one has to take into account that all original quality flags are considered in these plots. Further, only the minimal filtering of AERONET stations stated in the data section is used (e.g. stations near water bodies have not been removed). While there is a noticeable improvement for both datasets, some pattern in the bias related to the landtype is still visible. We tried to use the landtype information in the input, facing challenges and not being able to improve our result. Possible future work could take this into account.

To conclude, we can take a look at Fig.9 and check the behaviour of our models respect to the Aerosol_Cloud_Fraction_Land from the NASA DT GOES-16 ABI AOD dataset. Our post-processing models seem to be able to correct the original product when non-negligible cloud contamination is detected. A similar result is found when considering the corresponding variable from the GOES-17 dataset. While this result may seem counterintuitive, one has to take into account that the satellite data has been collocated with quality assured cloud-filtered AERONET data. While a more thorough analysis is needed to understand

https://doi.org/10.5194/egusphere-2026-1987
Preprint. Discussion started: 4 June 2026
© Author(s) 2026. CC BY 4.0 License.

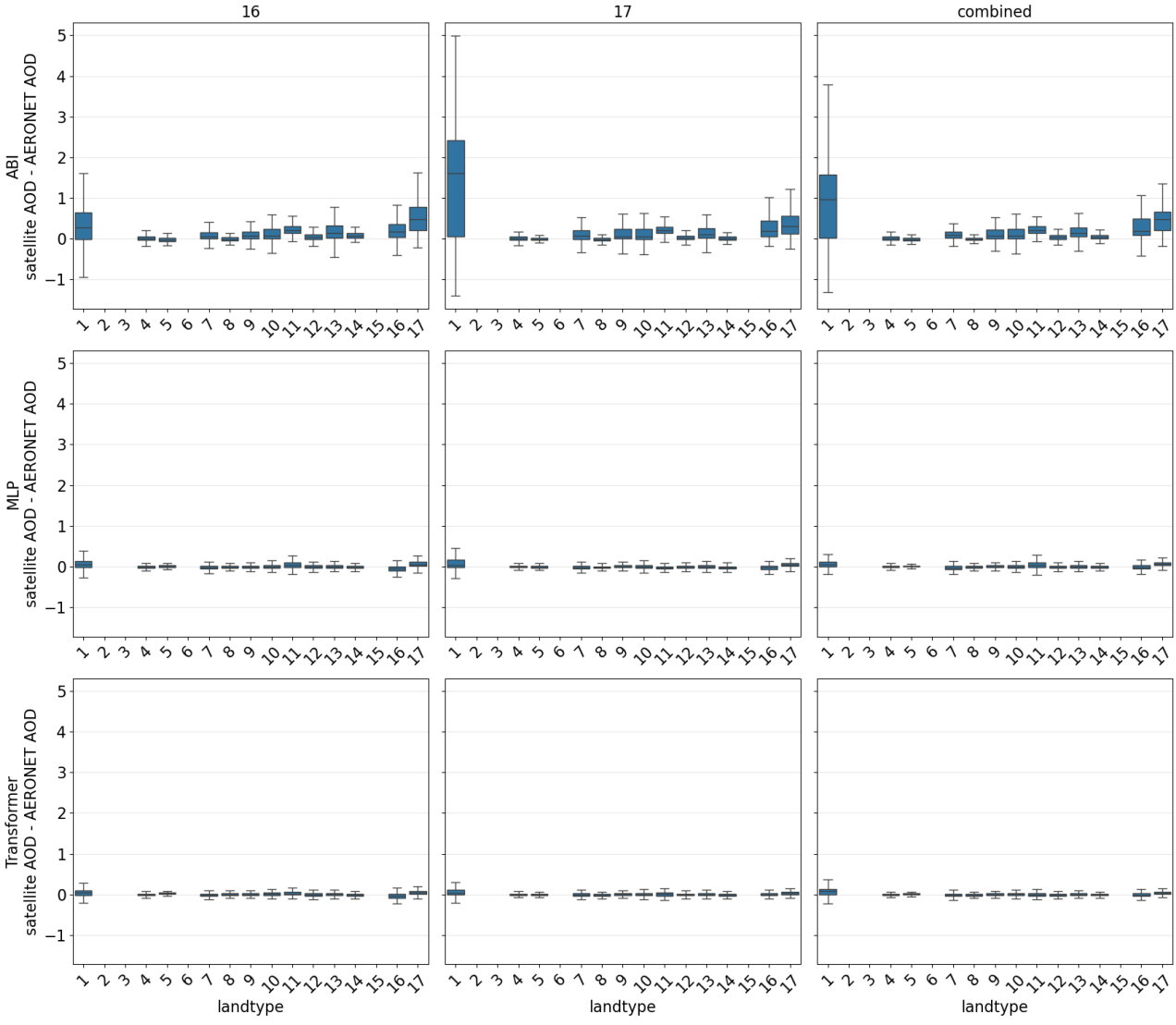

Figure 8. Boxplots representing the bias of satellite AOD compared to AERONET ground stations AOD at different land cover types. Different rows represent different satellite AOD models (NOAA ABI AOD, its post-process corrections using MLP, and its post-process corrections using MLP transformer). The columns represent different configurations (AOD obtained using only GOES-16 data, only GOES-17 data, and the combination).

and reliably predict the behaviour of our models near clouds, this result seems to strengthen our previous findings, suggesting that our methodology can significantly enhance low quality AOD retrievals.

https://doi.org/10.5194/egusphere-2026-1987
Preprint. Discussion started: 4 June 2026
© Author(s) 2026. CC BY 4.0 License.

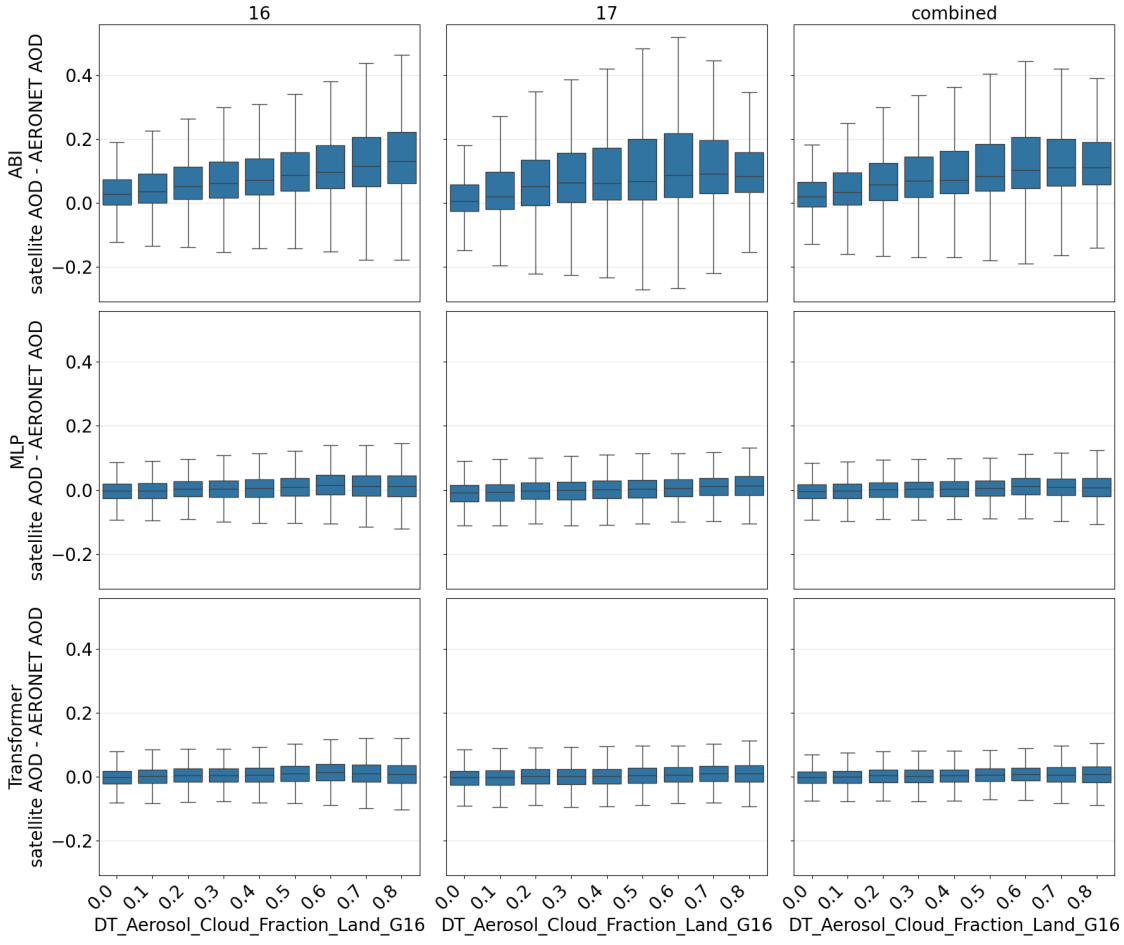

Figure 9. Boxplots representing the bias of satellite AOD compared to AERONET ground stations AOD at different bins of the variable Aerosol_Cloud_Fraction_Land from the NASA DT ABI AOD GOES-16 dataset. Different rows represent different satellite AOD models (NASA DT ABI AOD, its post-process corrections using MLP, and its post-process corrections using MLP transformer). The columns represent different configurations (AOD obtained using only GOES-16 data, only GOES-17 data, and the combination).

5 Conclusions

In this work, we studied the post-process correction of geostationary satellite AOD products using deep learning models. In particular, we compared modelling the corrections at single time-steps (using a MLP architecture) to modelling the corrections as a time-series (using a transformer architecture). Furthermore, we analyzed the difference in modelling such corrections using single satellite data (from GOES-16 or GOES-17) and using a combination of satellites with different position and viewing angles. We evaluated our methodology on two different datasets, considering the post-process correction of both NOAA ABI AOD and NASA DT ABI AOD products, and comparing our final result to AERONET AOD ground stations data. Our results

https://doi.org/10.5194/egusphere-2026-1987
Preprint. Discussion started: 4 June 2026
© Author(s) 2026. CC BY 4.0 License.

are promising: as expected, both modelling temporal correlations and combining the two satellites improve the AOD retrieval corrections, resulting in better accuracy and generalization. Qualitatively the improvement is evident for both the datasets considered, although relatively less important for the NASA DT dataset.

Our models combining GOES-16 and GOES-17 data could probably be improved further. While using data from both satellites as inputs (Fig.1), our models have two different outputs (one for each satellite) and the loss function is minimized pushing these two outputs to match AERONET AOD separately. The final result is then obtained by averaging into a single output. This average is not directly part of the loss function. Future studies could focus on a different model, specifically designed to do so, and potentially improving our results. Another limitation of this study is in the collocation: our datasets considers only cases when both GOES-16 and GOES-17 data are available. In such case, one could use the models trained using single satellite data for post-process correcting the retrieval. Other models capable of handling missing data from one of the instruments could be considered in the future. How to combine geostationary instruments and low-orbiting satellite instruments (such as MODIS) using deep learning will also be a topic of interest for possible extensions of this study.

In conclusion, given our promising results and the potential for future work, we believe the methodology presented in this study to be an effective approach for correcting satellite AOD retrievals and supporting related applications.

https://doi.org/10.5194/egusphere-2026-1987
Preprint. Discussion started: 4 June 2026
© Author(s) 2026. CC BY 4.0 License.

Code and data availability. All the datasets utilized in this study are open access and available online. Data from NOAA's GOES-R series satellite (level 1b radiances and NOAA GOES-R ABI AOD) is available on Amazon S3. All the NASA data (Dark Target GOES-R ABI Level 2 AOD, ASTER DEM) can be downloaded using the NASA Earthdata Search website at https://www.earthdata.nasa.gov/. Data and code will be available from the authors on reasonable request.

Appendix A: Lists of variables used from datasets

NOAA GOES-R Series ABI Level 1b Radiances

The following variables are used for both GOES-16 and GOES-17 ABI products:

- Rad (band 1, 0.47 μm) [$\mathrm{Wm}^{-2}\mu m^{-1}sr^{-1}$]
- kappa0 0.47 μm (band 1, 0.47 μm)[$(\mathrm{Wm}^{-2}\mu m^{-1})^{-1}$]
- Rad (band 2, 0.64 μm) [$\mathrm{Wm}^{-2}\mu m^{-1}sr^{-1}$]
- kappa0 (band 2, 0.64 μm) [$(\mathrm{Wm}^{-2}\mu m^{-1})^{-1}$]
- Rad (band 3, 0.86 μm) [$\mathrm{Wm}^{-2}\mu m^{-1}sr^{-1}$]
- kappa0 (band 3, 0.86 μm) [$(\mathrm{Wm}^{-2}\mu m^{-1})^{-1}$]
- Rad (band 4, 1.37 μm) [$\mathrm{Wm}^{-2}\mu m^{-1}sr^{-1}$]
- kappa0 (band 4, 1.37 μm) [$(\mathrm{Wm}^{-2}\mu m^{-1})^{-1}$]
- Rad (band 5, 1.6 μm) [$\mathrm{Wm}^{-2}\mu m^{-1}sr^{-1}$]
- kappa0 (band 5, 1.6 μm) [$(\mathrm{Wm}^{-2}\mu m^{-1})^{-1}$]
- Rad (band 6, 2.2 μm) [$\mathrm{Wm}^{-2}\mu m^{-1}sr^{-1}$]
- kappa0 2.2 (band 6, 2.2 μm) [$(\mathrm{Wm}^{-2}\mu m^{-1})^{-1}$]

Radiances have ben converted to TOA reflectances multiplying by the respective kappa0 factor, and dividing by the cosine of the solar zenith angle (sza).

NOAA GOES-R Series ABI Level 2 AOD

The following variables are used for both GOES-16 and GOES-17 ABI products:

- AOD
- DQF (Data Quality Flag)

https://doi.org/10.5194/egusphere-2026-1987
Preprint. Discussion started: 4 June 2026
© Author(s) 2026. CC BY 4.0 License.

NASA Dark Target GOES-R ABI Level 2 AOD

The following variables are used for both GOES-16 and GOES-17 ABI products:

- DT_Corrected_Optical_Depth_Land_2
- DT_Land_Ocean_Quality_Flag

AERONET AOD

- AERONET_AOD_500nm
- AERONET_440-870_Angstrom_Exponent

AERONET AOD 500 nm is transformed to AOD 550 nm by plugging the Ångstrom exponent into the Ångstrom law. The transformed AOD is used as target for training and evaluation each model.

Input variables lists

Tables A1 represent the input variables used in our models. The list of input for single satellite model depends on the satellite instrument (GOES-16 or GOES-17) and the Level 2 AOD dataset (NOAA or NASA). A model considering only one satellite would use variables listed for a specific satellite and a specific Level 2 AOD dataset. A model combining the two satellites would use variables for single satellite model (left table) for each satellite, and the auxiliary input variables list. Notice that AOD DQF variable are hot-encoded, therefore corresponds to multiple input features in practice.

IGBP land cover types

IGBP classification contains the following land cover types:

- **1**: Evergreen needleleaf forests
- **2**: Evergreen broadleaf forests
- **3**: Deciduous needleleaf forests
- **4**: Deciduous broadleaf forests
- **5**: Mixed forests
- **6**: Closed shrublands
- **7**: Open shrublands
- **8**: Woody savannas

https://doi.org/10.5194/egusphere-2026-1987
Preprint. Discussion started: 4 June 2026
© Author(s) 2026. CC BY 4.0 License.

- **9**: Savannas
- **10**: Grasslands
- **11**: Permanent wetlands
- **12**: Croplands
- **13**: Urban and built-up
- **14**: Cropland/natural
- **15**: Snow and ice
- **16**: Barren
- **17**: Water bodies

List of input variables for single satellite model
TOA Reflectance (Band 1, 0.47 μm)
TOA Reflectance (Band 2, 0.64 μm)
TOA Reflectance (Band 3, 0.86 μm)
TOA Reflectance (Band 4, 1.37 μm)
TOA Reflectance (Band 5, 1.6 μm)
TOA Reflectance (Band 6, 2.2 μm)
AOD
AOD DQF*
View Zenith Angle (VZA)
View Azimuth Angle (VAA)

List of auxiliary input variables
day_of_year_cos
day_of_year_sin
time_of_day
Solar Zenith Angle (SZA)
Solar Azimuth Angle (SAA)

Table A1. Input variable lists (*Multiple input variables represent different one-hot encoded DQF values.)

https://doi.org/10.5194/egusphere-2026-1987
Preprint. Discussion started: 4 June 2026
© Author(s) 2026. CC BY 4.0 License.

Spatio-temporal Cross-Validation representation

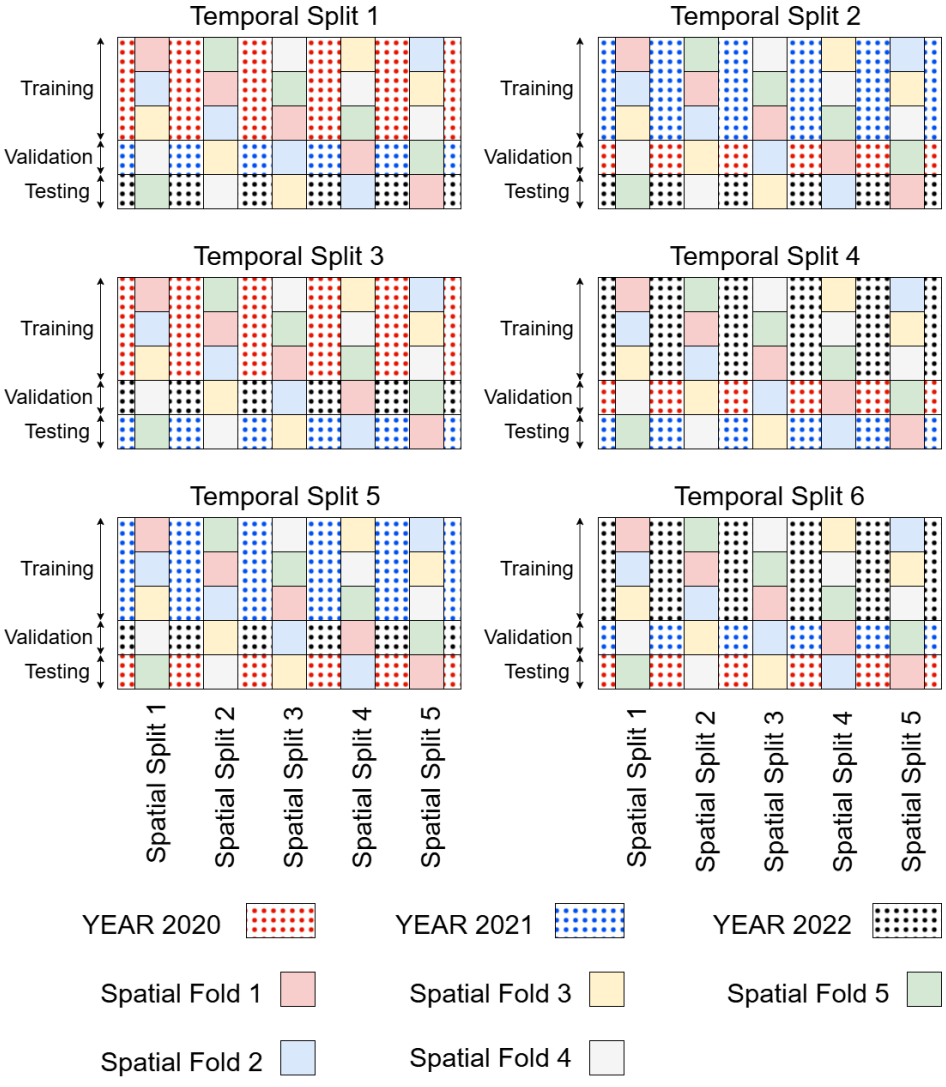

Figure A1. Representation of the spatio-temporal CV splits. Each temporal split consider the same 5 spatial splits. We have 6 temporal splits, for a total of 30 spatio-temporal CV splits. Each spatial fold corresponds to a different group of AERONET ground stations, while each temporal fold corresponds to a different year (2020, 2021 or 2022). Each spatio-temporal split considers a different combination of AERONET ground stations (spatial folds) and year (temporal folds) to train the model (the validation set is used to monitor and stop the training when validation loss stops decreasing), while using the remaining ground stations and year (not used in training) to test the model and assess its performance.

https://doi.org/10.5194/egusphere-2026-1987
Preprint. Discussion started: 4 June 2026
© Author(s) 2026. CC BY 4.0 License.

Appendix B: NASA ABI DT AOD post-process correction results

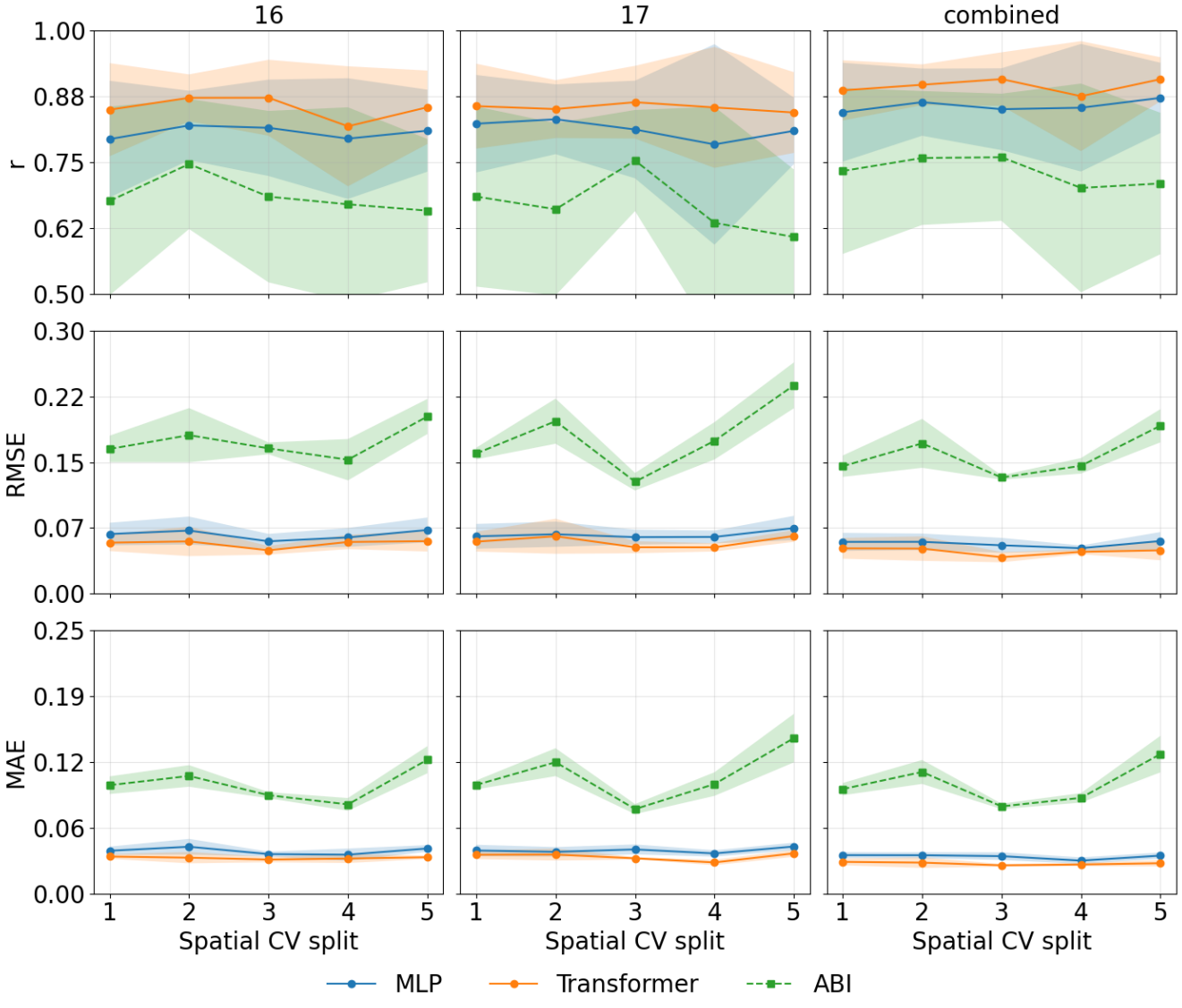

Figure B1. Comparison of metrics obtained on different spatio-temporal CV splits, for different models and satellite configurations. The metrics have been calculated in respect to AERONET measurements and NASA DT ABI AOD data (and related corrections). Each row represent a different metric (correlation, RMSE, MAE), while each column represent a different satellite configuration (GOES-16, GOES-17, combined). Dots correspond to averages of metrics across temporal splits, while shadows correspond to relative standard deviations. In each panel, ABI AOD product (green), MLP AOD (blue) and Transformer AOD (orange) are compared.

https://doi.org/10.5194/egusphere-2026-1987
Preprint. Discussion started: 4 June 2026
© Author(s) 2026. CC BY 4.0 License.

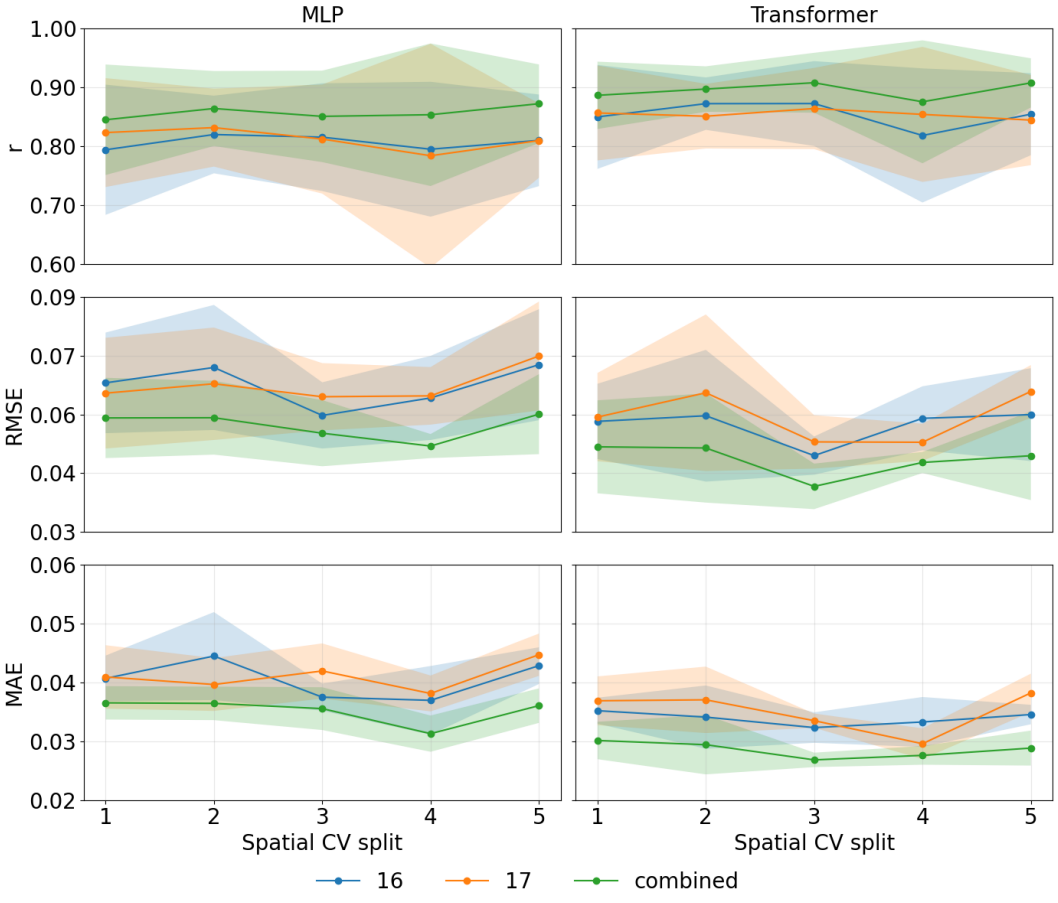

Figure B2. Comparison of metrics obtained on different spatio-temporal CV splits, for different models and satellite configurations. The metrics have been calculated in respect to AERONET measurements and NASA DT ABI AOD data (and related corrections). Each row represent a different metric (correlation, RMSE, MAE), while each column represent different models used for the post-process correction (MLP, Transformer). Dots correspond to averages of metrics across temporal splits, while shadows correspond to relative standard deviations. In each panel, GOES-16 model (blue), GOES-17 model (orange) and GOES-combined model (green) are compared.

https://doi.org/10.5194/egusphere-2026-1987
Preprint. Discussion started: 4 June 2026
© Author(s) 2026. CC BY 4.0 License.

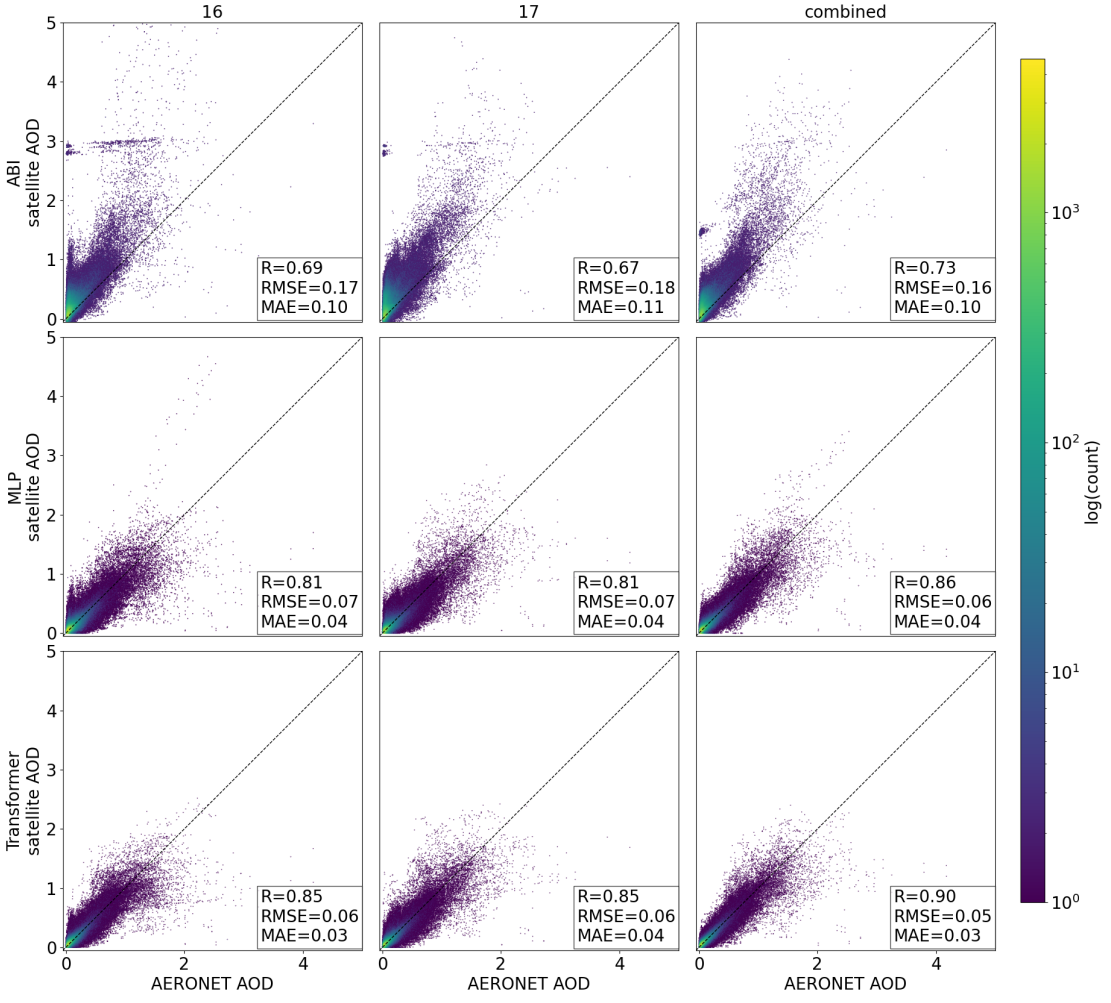

Figure B3. Scatterplots comparing NASA DT ABI AOD (and its post-process corrections using MLP and transformer) and AERONET ground stations data. On the rows we distinguish the original NASA DT ABI AOD product and its post-process corrections, while on the columns we distinguish the different configurations considered (using only GOES-16 data, only GOES-17 data, or a combination of the two). While the scatterplots contain points from all the CV splits test sets (N = 1220488), the metrics are calculated averaging the metrics obtained on the CV splits test sets (same metrics as in fig.3).

https://doi.org/10.5194/egusphere-2026-1987
Preprint. Discussion started: 4 June 2026
© Author(s) 2026. CC BY 4.0 License.

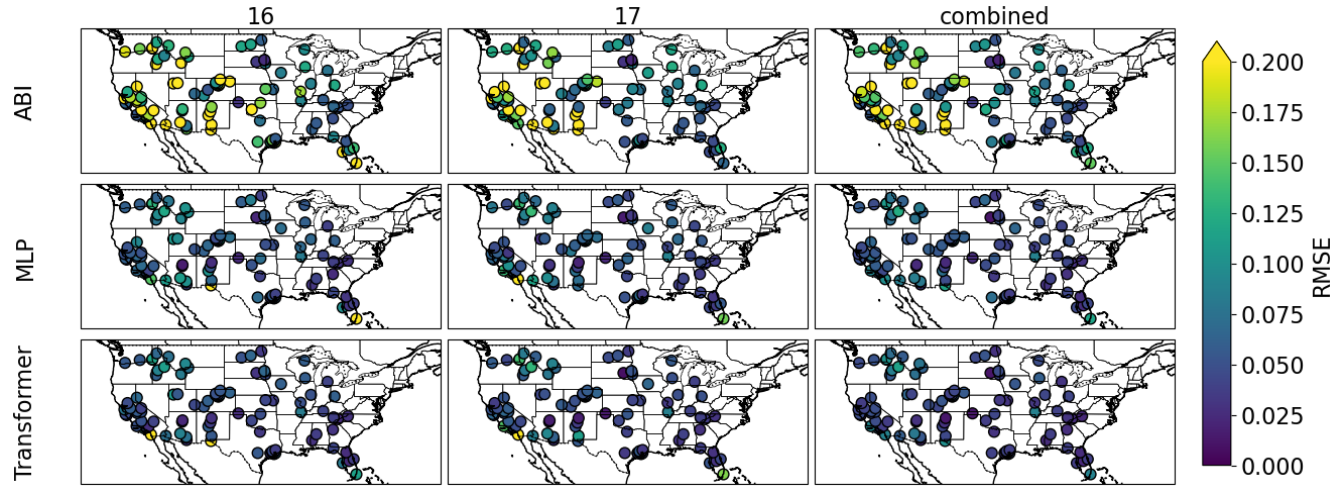

Figure B4. Comparison between NASA DT ABI AOD (and its post-process corrections using MLP and transformer) and AERONET ground stations data, at single ground stations over the region of interest. The color of the dots represent the RMSE metric at each station location: these metrics have been calculated directly on all temporal CV splits test sets (not averaging metrics on different temporal splits test sets). Stations from all spatial CV splits test sets are included. Rows represent the different models (NASA DT AOD and its post-process corrections using MLP and transformer), while columns represent different configurations (model using only GOES-16 data, only GOES-17 data, or a combination of the two)

.

https://doi.org/10.5194/egusphere-2026-1987
Preprint. Discussion started: 4 June 2026
© Author(s) 2026. CC BY 4.0 License.

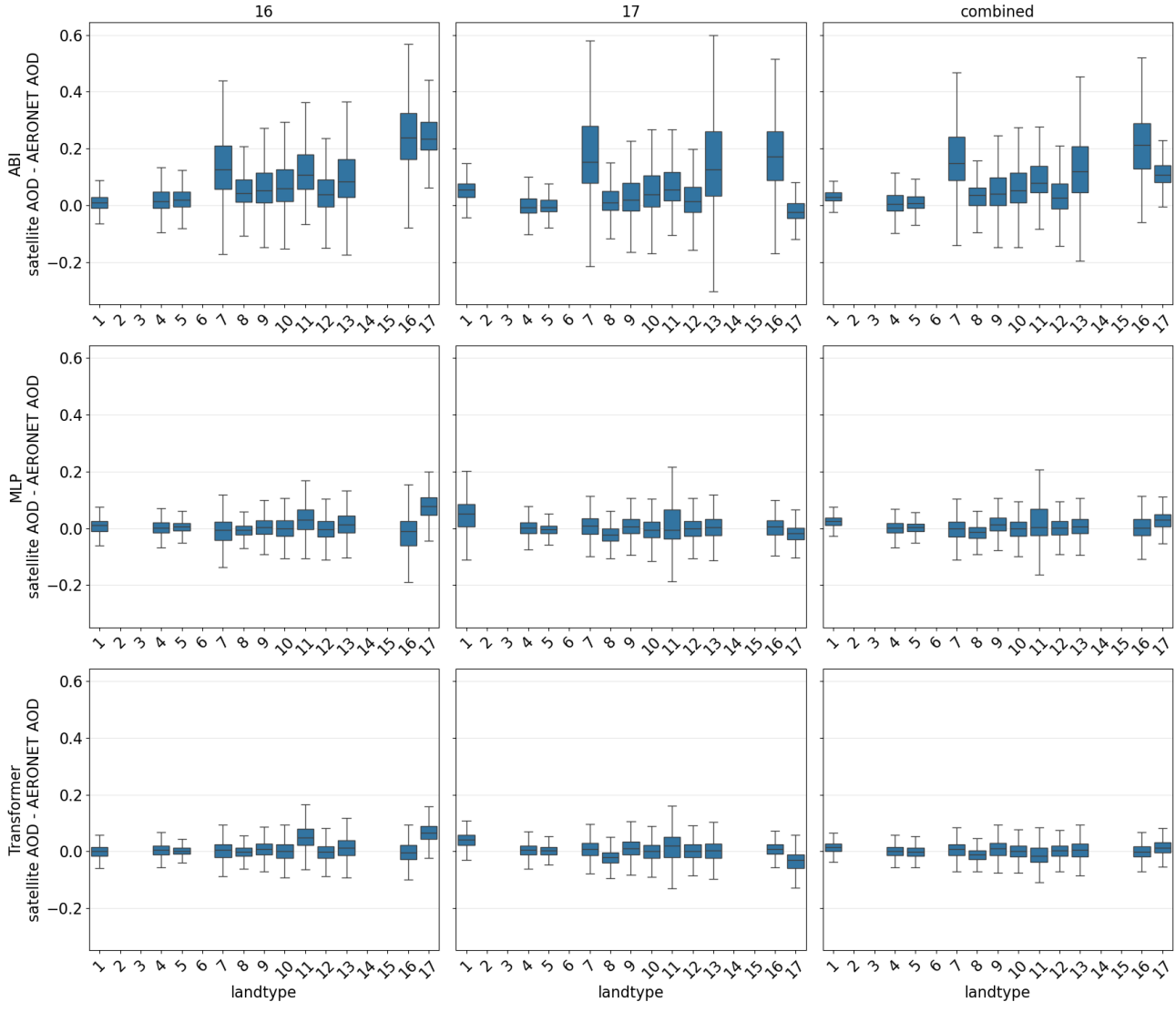

Figure B5. Boxplots representing the bias of satellite AOD compared to AERONET ground stations AOD at different land cover types. Different rows represent different satellite AOD models (NASA DT ABI AOD, its post-process corrections using MLP, and its post-process corrections using MLP transformer). The columns represent different configurations (AOD obtained using only GOES-16 data, only GOES-17 data, and the combination).

https://doi.org/10.5194/egusphere-2026-1987
Preprint. Discussion started: 4 June 2026
© Author(s) 2026. CC BY 4.0 License.

Appendix C: Model architectures

This section outlines the Deep Learning models considered in this study. To show the advantage of Transformer architectures, we utilize a Multi Layer Perceptron (MLP) as baseline for comparison. The deep learning pipeline has been implemented using the PyTorch framework (Paszke et al., 2019).

MLP

The Multi Layer Perceptron (MLP) (Rosenblatt, 1958; Bishop and Bishop, 2024) is one of the simplest Deep Learning architectures and it can be described by iterative usage of equation C1.

$$a_j^{n+1} = \phi\left(\sum_{i=1}^{D_n} w_{ji}^{n+1} a_i^n + w_{j0}^{n+1}\right) \tag{C1}$$

Let's look at Fig.C1 as a simple example of a MLP with 2 hidden layers. The input layer (identified by n = 1) contains 2 input nodes (a_1^1 and a_1^2): these nodes represent the input variables. For each node in the following layer (the first hidden layer, identified by n = 2), we apply equation C1, using $D_0 = 2$ in place of D_n (i.e. the number of nodes in the previous. In this case, the input layer). The w_{ji}^{n+1} are parameters to be tuned during the model training: they are usually referred as weights (when $i \neq 0$) and bias terms (when $i = 0$).

In Fig.C1 the role of weights and bias terms is clarified in the red box, where they are represented as connection between nodes. The result of the inputs aggregation is then given as input to the activation function ϕ. The role of the activation function is to bring non linearity into the model, and so is generally a non linear function (except for the output layer, where the choice depends on the nature of the output). Equation C1 is applied to each node in the layer, and to the following layers.

In this work, we utilize two hidden layers as in Fig.C1. The output layer does not have activation function (i.e. the output transformation is a linear transformation). The hyperparameters tuned and their respective search space can be found in table C1.

Hidden nodes 1	32, 64, 128, 256, 512
Hidden nodes 2	32, 64, 128, 256, 512
Hidden activation function	ReLU (Jarrett et al., 2009), Leaky ReLU (Maas et al., 2013), SELU (Klambauer et al., 2017), ELU (Clevert et al., 2016)
Batch size	128, 256, 512, 1024
Learning rate	10^{-2}, 10^{-3}, 10^{-4}

Table C1. Hyperparameters tuned and their respective search space for the MLP architecture.

Transformer

Transformer models (Vaswani et al., 2023; Bishop and Bishop, 2024) are based on the attention mechanism (Bahdanau et al., 2016; Vaswani et al., 2023). The transformer model implemented in this study is essentially a transformer encoder as presented

https://doi.org/10.5194/egusphere-2026-1987
Preprint. Discussion started: 4 June 2026
© Author(s) 2026. CC BY 4.0 License.

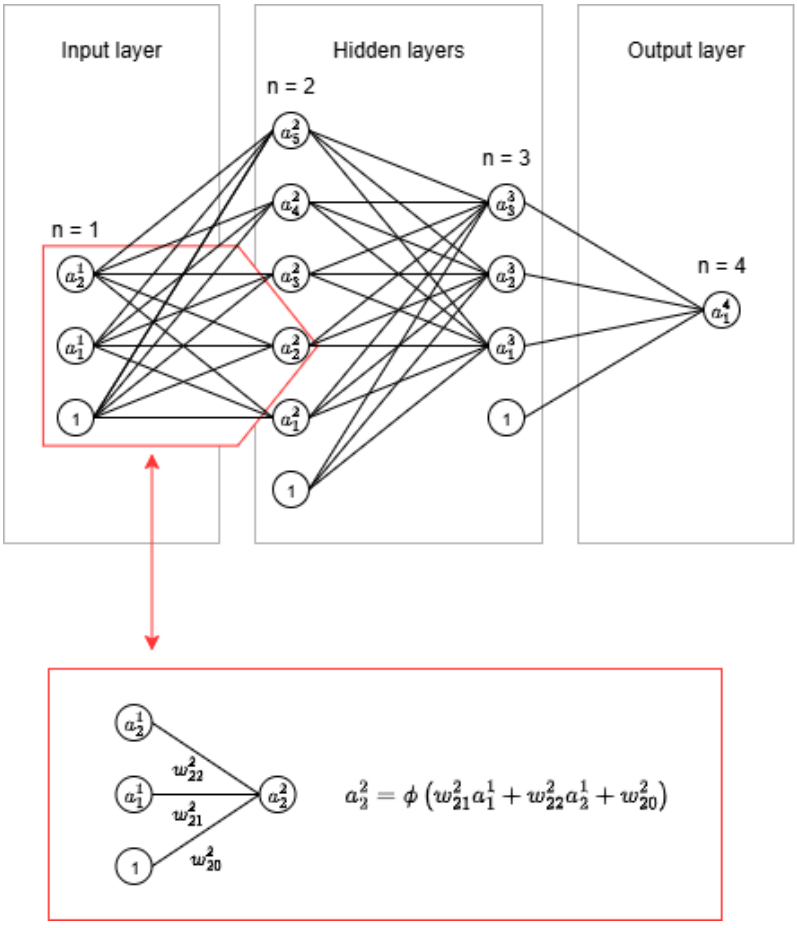

Figure C1. Representation of a MLP architecture with 2 input nodes, 2 hidden layers (5 hidden nodes in the first layer and 3 hidden nodes in the second layer), and one output node.

in Vaswani et al. (2023), and it is used to predict an AOD bias correction at each time step in a time series of AOD retrievals. Using the self-attention mechanism, the transformer model estimates relationships among the inputs at different time steps (tokens), enabling long-range interactions without the need of a recurrent (and not well parallelizible) model.

To understand the self-attention mechanism, let's look at Fig. C2 and define 3 matrices containing parameters to be learned in the training process: $\boldsymbol{W}^{(q)}$, $\boldsymbol{W}^{(k)}$, $\boldsymbol{W}^{(v)}$. $\boldsymbol{W}^{(q)}$ and $\boldsymbol{W}^{(k)}$ have size D x D (D being the number of features describing each token), while $\boldsymbol{W}^{(v)}$ has size D x D_v. Utilizing the input matrix $\boldsymbol{X}$ of size N x D (N being the number of tokens), the query matrix $\boldsymbol{Q}$, key matrix $\boldsymbol{K}$ and value matrix $\boldsymbol{V}$ can be calculated:

$$\boldsymbol{Q} = \boldsymbol{X}\boldsymbol{W}^{(q)} \qquad \boldsymbol{K} = \boldsymbol{X}\boldsymbol{W}^{(k)} \qquad \boldsymbol{V} = \boldsymbol{X}\boldsymbol{W}^{(v)} \tag{C2}$$

The self-attention mechanism is finally defined by:

https://doi.org/10.5194/egusphere-2026-1987
Preprint. Discussion started: 4 June 2026
© Author(s) 2026. CC BY 4.0 License.

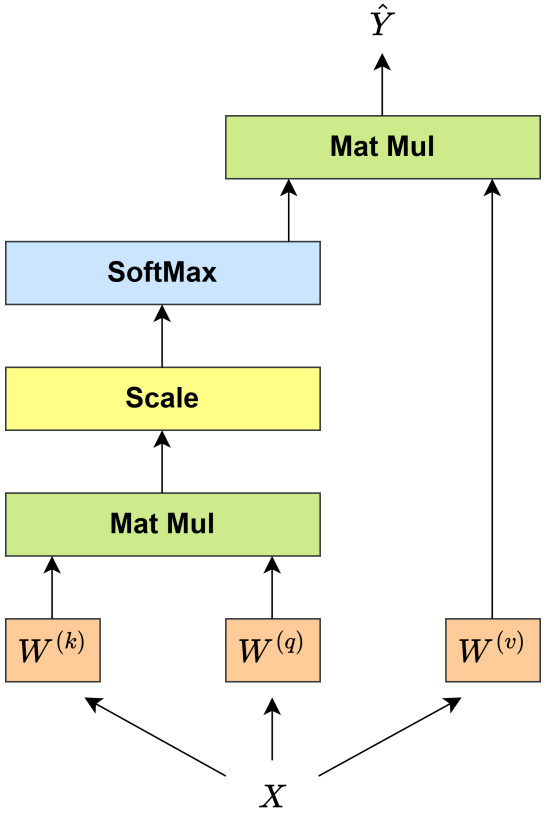

Figure C2. Self-attention mechanism.

$$\hat{\boldsymbol{Y}} = \text{softmax}\left(\frac{\boldsymbol{Q}\boldsymbol{K}^{\boldsymbol{T}}}{\sqrt{\text{D}}}\right)\boldsymbol{V} \tag{C3}$$

where:

$$\text{softmax}(x_i) = \frac{exp(x_i)}{\sum_i exp(x_i)} \tag{C4}$$

The softmax function calculates the so-called attention weights. These weights are simply used to combine different tokens in the input time series. Notice that $\boldsymbol{Q}$, $\boldsymbol{K}$ and $\boldsymbol{V}$ are just a different representation of the original input time series: $\boldsymbol{W}^{(q)}$, $\boldsymbol{W}^{(k)}$, and $\boldsymbol{W}^{(v)}$ just act as linear transformations on separate tokens, mapping them to a different representation. Notice also that $\boldsymbol{W}^{(q)}$ is generally different from $\boldsymbol{W}^{(k)}$, and consequently $\boldsymbol{Q}$ is generally different from $\boldsymbol{K}$, so the attention weights will

https://doi.org/10.5194/egusphere-2026-1987
Preprint. Discussion started: 4 June 2026
© Author(s) 2026. CC BY 4.0 License.

not be in general symmetric (i.e. the i-th step in the sequence is influenced by the j-th differently from how the j-th step is influenced by the i-th step).

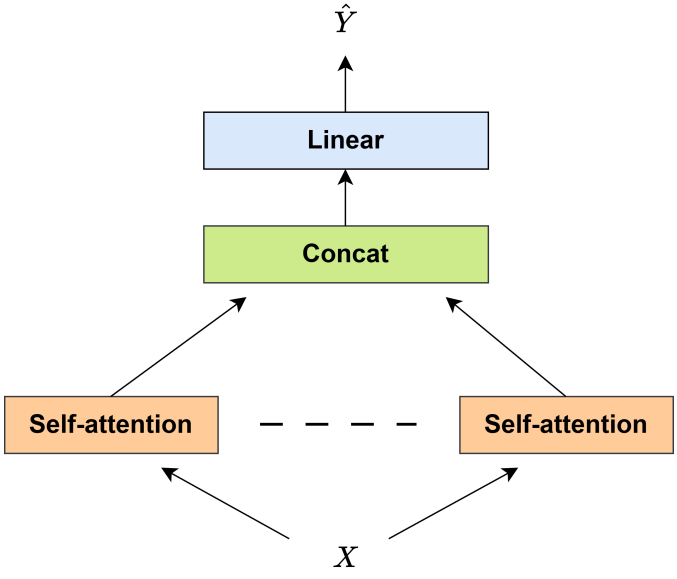

Figure C3. Multi-head self-attention.

As different filters/kernels in a Convolutional Neural Network can be used to learn and detect different patterns, in the same way one can apply the self-attention mechanism multiple times in parallel. This goes by the name of multi-head self-attention and is represented in Fig.C3. The outputs of multiple self-attention layers are concatenated and combined by a linear transformation. D_v is the same for all the attention heads and is defined as D_v = D/H (where H is the number of attention heads).

By adding non-linearity using the MLP architecture (notice that so far non-linearities are introduced only to calculate the attention weights), one can build the transformer encoder. The MLP is applied token-wise. Fig.C4 represents the architecture used in this paper, consisting of a linear embedding layer, two transformer layers, and a final linear transformation. Layer normalization (Ba et al., 2016) is applied before each multi-head self-attention. Residual/skip connections (He et al., 2015) are used between inputs and outputs of each multi-head self-attention layer.

Table C2 shows the hyperparameters tuned and their respective search space. Here D represents the dimension of the embedding. The MLP hidden dimension is set as 2 x D.

https://doi.org/10.5194/egusphere-2026-1987
Preprint. Discussion started: 4 June 2026
© Author(s) 2026. CC BY 4.0 License.

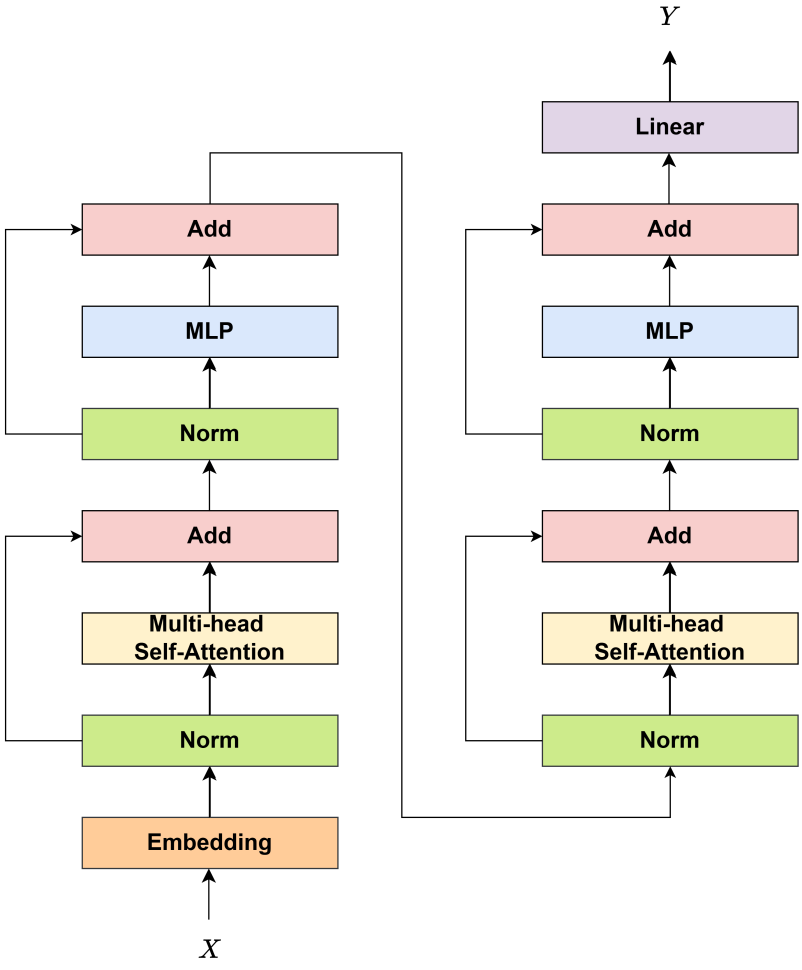

Figure C4. Transformer encoder architecture utilized in this work.

D (hidden dimension)	H*8, H*12, H*16
H (number of heads)	4, 6, 8
Batch size	128, 256, 512, 1024
Learning rate	10^{-2}, 10^{-3}, 10^{-4}

Table C2. Hyperparameters tuned and their respective search space for the transformer architecture.

Author contributions. **AP**: Formulation, implementation, data processing, analysis, writing **VK**: Formulation, analysis, writing, supervision **TL**: Formulation, analysis, writing, supervision **RL**: Analysis, writing **YS**: Data processing, analysis, writing **HZ**: Analysis, writing **AL**: Formulation, analysis, writing, supervision

https://doi.org/10.5194/egusphere-2026-1987
Preprint. Discussion started: 4 June 2026
© Author(s) 2026. CC BY 4.0 License.

Competing interests. The authors declare no competing interests.

Disclaimer. The results and conclusions are those of the author(s) and do not necessarily represent the views of NOAA or the Department of Commerce.

Acknowledgements. The research was supported by the Research Council of Finland Flagship of Advanced Mathematics for Sensing, Imaging and Modeling grant No 358944 (FAME), by the Research Council of Finland Flagship of Atmosphere and Climate Competence Center grant No 359342 (ACCC), and by the Research Council of Finland Centre of Excellence in Inverse Modelling and Imaging grants 353084. The authors thank Arttu Väisanen from the Finnish Meteorological Institute (FMI) for sharing their codes to compute satellite and sun geometry. A portion of this work used code generously provided by Brian Blaylock's GOES-2-go python package (https://github.com/blaylockbk/goes2go).

Financial support. This work was funded by the Research Council of Finland Flagship of Advanced Mathematics for Sensing, Imaging and Modeling grant No 358944 (FAME) and by the Research Council of Finland Centre of Excellence in Inverse Modelling and Imaging grants 353084.

https://doi.org/10.5194/egusphere-2026-1987
Preprint. Discussion started: 4 June 2026
© Author(s) 2026. CC BY 4.0 License.

References

Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., Greenfield, P., Droettboom, M., Bray, E., Aldcroft, T., Davis, M., Ginsburg, A., Price-Whelan, A. M., Kerzendorf, W. E., Conley, A., Crighton, N., Barbary, K., Muna, D., Ferguson, H., Grollier, F., Parikh, M. M., Nair, P. H., Unther, H. M., Deil, C., Woillez, J., Conseil, S., Kramer, R., Turner, J. E. H., Singer, L., Fox, R., Weaver, B. A., Zabalza, V., Edwards, Z. I., Azalee Bostroem, K., Burke, D. J., Casey, A. R., Crawford, S. M., Dencheva, N., Ely, J., Jenness, T., Labrie, K., Lim, P. L., Pierfederici, F., Pontzen, A., Ptak, A., Refsdal, B., Servillat, M., and Streicher, O.: Astropy: A community Python package for astronomy, , 558, A33, https://doi.org/10.1051/0004-6361/201322068, 2013.

Astropy Collaboration, Price-Whelan, A. M., Sipőcz, B. M., Günther, H. M., Lim, P. L., Crawford, S. M., Conseil, S., Shupe, D. L., Craig, M. W., Dencheva, N., Ginsburg, A., Vand erPlas, J. T., Bradley, L. D., Pérez-Suárez, D., de Val-Borro, M., Aldcroft, T. L., Cruz, K. L., Robitaille, T. P., Tollerud, E. J., Ardelean, C., Babej, T., Bach, Y. P., Bachetti, M., Bakanov, A. V., Bamford, S. P., Barentsen, G., Barmby, P., Baumbach, A., Berry, K. L., Biscani, F., Boquien, M., Bostroem, K. A., Bouma, L. G., Brammer, G. B., Bray, E. M., Breytenbach, H., Buddelmeijer, H., Burke, D. J., Calderone, G., Cano Rodríguez, J. L., Cara, M., Cardoso, J. V. M., Cheedella, S., Copin, Y., Corrales, L., Crichton, D., D'Avella, D., Deil, C., Depagne, É., Dietrich, J. P., Donath, A., Droettboom, M., Earl, N., Erben, T., Fabbro, S., Ferreira, L. A., Finethy, T., Fox, R. T., Garrison, L. H., Gibbons, S. L. J., Goldstein, D. A., Gommers, R., Greco, J. P., Greenfield, P., Groener, A. M., Grollier, F., Hagen, A., Hirst, P., Homeier, D., Horton, A. J., Hosseinzadeh, G., Hu, L., Hunkeler, J. S., Ivezić, Ž., Jain, A., Jenness, T., Kanarek, G., Kendrew, S., Kern, N. S., Kerzendorf, W. E., Khvalko, A., King, J., Kirkby, D., Kulkarni, A. M., Kumar, A., Lee, A., Lenz, D., Littlefair, S. P., Ma, Z., Macleod, D. M., Mastropietro, M., McCully, C., Montagnac, S., Morris, B. M., Mueller, M., Mumford, S. J., Muna, D., Murphy, N. A., Nelson, S., Nguyen, G. H., Ninan, J. P., Nöthe, M., Ogaz, S., Oh, S., Parejko, J. K., Parley, N., Pascual, S., Patil, R., Patil, A. A., Plunkett, A. L., Prochaska, J. X., Rastogi, T., Reddy Janga, V., Sabater, J., Sakurikar, P., Seifert, M., Sherbert, L. E., Sherwood-Taylor, H., Shih, A. Y., Sick, J., Silbiger, M. T., Singanamalla, S., Singer, L. P., Sladen, P. H., Sooley, K. A., Sornarajah, S., Streicher, O., Teuben, P., Thomas, S. W., Tremblay, G. R., Turner, J. E. H., Terrón, V., van Kerkwijk, M. H., de la Vega, A., Watkins, L. L., Weaver, B. A., Whitmore, J. B., Woillez, J., Zabalza, V., and Astropy Contributors: The Astropy Project: Building an Open-science Project and Status of the v2.0 Core Package, , 156, 123, https://doi.org/10.3847/1538-3881/aabc4f, 2018.

Astropy Collaboration, Price-Whelan, A. M., Lim, P. L., Earl, N., Starkman, N., Bradley, L., Shupe, D. L., Patil, A. A., Corrales, L., Brasseur, C. E., N"othe, M., Donath, A., Tollerud, E., Morris, B. M., Ginsburg, A., Vaher, E., Weaver, B. A., Tocknell, J., Jamieson, W., van Kerkwijk, M. H., Robitaille, T. P., Merry, B., Bachetti, M., G"unther, H. M., Aldcroft, T. L., Alvarado-Montes, J. A., Archibald, A. M., B'odi, A., Bapat, S., Barentsen, G., Baz'an, J., Biswas, M., Boquien, M., Burke, D. J., Cara, D., Cara, M., Conroy, K. E., Conseil, S., Craig, M. W., Cross, R. M., Cruz, K. L., D'Eugenio, F., Dencheva, N., Devillepoix, H. A. R., Dietrich, J. P., Eigenbrot, A. D., Erben, T., Ferreira, L., Foreman-Mackey, D., Fox, R., Freij, N., Garg, S., Geda, R., Glattly, L., Gondhalekar, Y., Gordon, K. D., Grant, D., Greenfield, P., Groener, A. M., Guest, S., Gurovich, S., Handberg, R., Hart, A., Hatfield-Dodds, Z., Homeier, D., Hosseinzadeh, G., Jenness, T., Jones, C. K., Joseph, P., Kalmbach, J. B., Karamehmetoglu, E., Kaluszy'nski, M., Kelley, M. S. P., Kern, N., Kerzendorf, W. E., Koch, E. W., Kulumani, S., Lee, A., Ly, C., Ma, Z., MacBride, C., Maljaars, J. M., Muna, D., Murphy, N. A., Norman, H., O'Steen, R., Oman, K. A., Pacifici, C., Pascual, S., Pascual-Granado, J., Patil, R. R., Perren, G. I., Pickering, T. E., Rastogi, T., Roulston, B. R., Ryan, D. F., Rykoff, E. S., Sabater, J., Sakurikar, P., Salgado, J., Sanghi, A., Saunders, N., Savchenko, V., Schwardt, L., Seifert-Eckert, M., Shih, A. Y., Jain, A. S., Shukla, G., Sick, J., Simpson, C., Singanamalla, S., Singer, L. P., Singhal, J., Sinha, M., SipHocz, B. M., Spitler, L. R., Stansby, D., Streicher, O., Sumak, J., Swinbank, J. D., Taranu, D. S., Tewary, N., Tremblay, G. R., Val-Borro, M. d., Van Kooten, S. J., Vasovi'c, Z., Verma, S., de Miranda Cardoso, J. V., Williams, P. K. G., Wilson, T. J., Winkel, B., Wood-Vasey, W. M., Xue, R., Yoachim, P., Zhang, C.,

https://doi.org/10.5194/egusphere-2026-1987
Preprint. Discussion started: 4 June 2026
© Author(s) 2026. CC BY 4.0 License.

Zonca, A., and Astropy Project Contributors: The Astropy Project: Sustaining and Growing a Community-oriented Open-source Project and the Latest Major Release (v5.0) of the Core Package, , 935, 167, https://doi.org/10.3847/1538-4357/ac7c74, 2022.

Ba, J. L., Kiros, J. R., and Hinton, G. E.: Layer Normalization, https://arxiv.org/abs/1607.06450, 2016.

Bahdanau, D., Cho, K., and Bengio, Y.: Neural Machine Translation by Jointly Learning to Align and Translate, https://doi.org/10.48550/arXiv.1409.0473, arXiv:1409.0473 [cs], 2016.

Belward, A. S., Estes, J. E., and Kline, K. D.: The IGBP-DIS global 1-km land-cover data set DISCover: A project overview, Photogrammetric Engineering and Remote Sensing, 65, 1013–1020, 1999.

Bishop, C. M. and Bishop, H.: Deep Learning: Foundations and Concepts, Springer International Publishing, Cham, https://doi.org/10.1007/978-3-031-45468-4, 2024.

Blaylock, B. K.: GOES-2-go: Download and display GOES-East and GOES-West data, https://github.com/blaylockbk/goes2go, 2023.

Clevert, D.-A., Unterthiner, T., and Hochreiter, S.: Fast and Accurate Deep Network Learning by Exponential Linear Units (ELUs), https://arxiv.org/abs/1511.07289, 2016.

Eck, T. F., Holben, B. N., Reid, J. S., Dubovik, O., Smirnov, A., O'Neill, N. T., Slutsker, I., and Kinne, S.: Wavelength dependence of the optical depth of biomass burning, urban, and desert dust aerosols, Journal of Geophysical Research: Atmospheres, 104, 31 333–31 349, https://doi.org/https://doi.org/10.1029/1999JD900923, 1999.

Ester, M., Kriegel, H.-P., Sander, J., and Xu, X.: A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise, in: Proceedings of the 2nd International Conference on Knowledge Discovery and Data Mining, pp. 226–231, AAAI Press, Portland, OR, 1996.

Fujisada, H., Urai, M., and Iwasaki, A.: Advanced methodology for ASTER DEM generation, IEEE transactions on geoscience and remote sensing, 49, 5080–5091, 2011.

Fujisada, H., Urai, M., and Iwasaki, A.: Technical methodology for ASTER global DEM, IEEE Transactions on Geoscience and Remote Sensing, 50, 3725–3736, 2012.

Giles, D. M., Sinyuk, A., Sorokin, M. G., Schafer, J. S., Smirnov, A., Slutsker, I., Eck, T. F., Holben, B. N., Lewis, J. R., Campbell, J. R., Welton, E. J., Korkin, S. V., and Lyapustin, A. I.: Advancements in the Aerosol Robotic Network (AERONET) Version 3 database – automated near-real-time quality control algorithm with improved cloud screening for Sun photometer aerosol optical depth (AOD) measurements, Atmospheric Measurement Techniques, 12, 169–209, https://doi.org/10.5194/amt-12-169-2019, 2019.

GOES-R Algorithm Working Group and GOES-R Series Program: NOAA GOES-R Series Advanced Baseline Imager (ABI) Level 2 Aerosol Optical Depth (AOD), https://doi.org/10.7289/V5BV7DSR, 2017.

GOES-R Calibration Working Group and GOES-R Series Program: NOAA GOES-R Series Advanced Baseline Imager (ABI) Level 1b Radiances, https://doi.org/10.7289/V5BV7DSR, 2017.

Goodfellow, I., Bengio, Y., and Courville, A.: Deep Learning, MIT Press, http://www.deeplearningbook.org, 2016.

Haywood, J. and Boucher, O.: Estimates of the direct and indirect radiative forcing due to tropospheric aerosols: A review, Reviews of Geophysics, 38, 513–543, https://doi.org/https://doi.org/10.1029/1999RG000078, 2000.

He, K., Zhang, X., Ren, S., and Sun, J.: Deep Residual Learning for Image Recognition, https://arxiv.org/abs/1512.03385, 2015.

Holben, B., Eck, T., Slutsker, I., Tanré, D., Buis, J., Setzer, A., Vermote, E., Reagan, J., Kaufman, Y., Nakajima, T., Lavenu, F., Jankowiak, I., and Smirnov, A.: AERONET—A Federated Instrument Network and Data Archive for Aerosol Characterization, Remote Sensing of Environment, 66, 1–16, https://doi.org/https://doi.org/10.1016/S0034-4257(98)00031-5, 1998.

https://doi.org/10.5194/egusphere-2026-1987
Preprint. Discussion started: 4 June 2026
© Author(s) 2026. CC BY 4.0 License.

Jarrett, K., Kavukcuoglu, K., Ranzato, M., and LeCun, Y.: What is the best multi-stage architecture for object recognition?, 2009 IEEE 12th International Conference on Computer Vision, pp. 2146–2153, https://api.semanticscholar.org/CorpusID:206769720, 2009.

Just, A. C., De Carli, M. M., Shtein, A., Dorman, M., Lyapustin, A., and Kloog, I.: Correcting Measurement Error in Satellite Aerosol Optical Depth with Machine Learning for Modeling PM2.5 in the Northeastern USA, Remote Sensing, 10, https://doi.org/10.3390/rs10050803, 2018.

Kaufman, Y. J., Tanré, D., Remer, L. A., Vermote, E. F., Chu, A., and Holben, B. N.: Operational remote sensing of tropospheric aerosol over land from EOS moderate resolution imaging spectroradiometer, Journal of Geophysical Research: Atmospheres, 102, 17 051–17 067, https://doi.org/10.1029/96JD03988, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/96JD03988, 1997.

Kingma, D. P. and Ba, J.: Adam: A Method for Stochastic Optimization, https://arxiv.org/abs/1412.6980, 2017.

Klambauer, G., Unterthiner, T., Mayr, A., and Hochreiter, S.: Self-Normalizing Neural Networks, https://arxiv.org/abs/1706.02515, 2017.

Kondragunta, S., Laszlo, I., Zhang, H., Ciren, P., and Huff, A.: Chapter 17 - Air Quality Applications of ABI Aerosol Products from the GOES-R Series, in: The GOES-R Series, edited by Goodman, S. J., Schmit, T. J., Daniels, J., and Redmon, R. J., pp. 203–217, Elsevier, https://doi.org/https://doi.org/10.1016/B978-0-12-814327-8.00017-2, 2020.

Lanzaco, B. L., Olcese, L. E., Palancar, G. G., and Toselli, B. M.: A Method to Improve MODIS AOD Values: Application to South America, Aerosol and Air Quality Research, 16, 1509–1522, https://doi.org/10.4209/aaqr.2015.05.0375, 2016.

Lary, D. J., Remer, L. A., MacNeill, D., Roscoe, B., and Paradise, S.: Machine Learning and Bias Correction of MODIS Aerosol Optical Depth, IEEE Geoscience and Remote Sensing Letters, 6, 694–698, https://doi.org/10.1109/LGRS.2009.2023605, 2009.

Laszlo, I. and Liu, H.: EPS Aerosol Optical Depth (AOD) Algorithm Theoretical Basis Document, Tech. rep., NOAA NESDIS CENTER for SATELLITE APPLICATIONS and RESEARCH, 2020.

Laszlo, I., Ciren, P., Liu, H., Kondragunta, S., Tarpley, J. D., and Goldberg, M. D.: Remote sensing of aerosol and radiation from geostationary satellites, Advances in Space Research, 41, 1882–1893, https://doi.org/10.1016/j.asr.2007.06.047, 2008.

Lenoble, J., Remer, L., and Tanré, D., eds.: Aerosol Remote Sensing, Springer Berlin Heidelberg, Berlin, Heidelberg, https://doi.org/10.1007/978-3-642-17725-5, 2013.

Levy, R. C., Remer, L. A., and Dubovik, O.: Global aerosol optical properties and application to Moderate Resolution Imaging Spectroradiometer aerosol retrieval over land, Journal of Geophysical Research: Atmospheres, 112, https://doi.org/10.1029/2006JD007815, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2006JD007815, 2007.

Lipponen, A., Kolehmainen, V., Kolmonen, P., Kukkurainen, A., Mielonen, T., Sabater, N., Sogacheva, L., Virtanen, T. H., and Arola, A.: Model-enforced post-process correction of satellite aerosol retrievals, Atmospheric Measurement Techniques, 14, 2981–2992, https://doi.org/10.5194/amt-14-2981-2021, 2021.

Lipponen, A., Reinvall, J., Väisänen, A., Taskinen, H., Lähivaara, T., Sogacheva, L., Kolmonen, P., Lehtinen, K., Arola, A., and Kolehmainen, V.: Deep-learning-based post-process correction of the aerosol parameters in the high-resolution Sentinel-3 Level-2 Synergy product, Atmospheric Measurement Techniques, 15, 895–914, https://doi.org/10.5194/amt-15-895-2022, 2022.

Loshchilov, I. and Hutter, F.: SGDR: Stochastic Gradient Descent with Warm Restarts, https://arxiv.org/abs/1608.03983, 2017.

Loveland, T. R. and Belward, A.: The international geosphere biosphere programme data and information system global land cover data set (DISCover), Acta Astronautica, 41, 681–689, 1997.

Maas, A. L., Hannun, A. Y., Ng, A. Y., et al.: Rectifier nonlinearities improve neural network acoustic models, in: Proc. icml, vol. 30, p. 3, Atlanta, GA, 2013.

https://doi.org/10.5194/egusphere-2026-1987
Preprint. Discussion started: 4 June 2026
© Author(s) 2026. CC BY 4.0 License.

NASA: ABI/GOES-16 Dark Target Aerosol 10-Min L2 Full Disk 10 km, https://doi.org/10.5067/MEaSUREs/GLDTA/XAERDT_L2_ABI_G16.001, https://doi.org/10.5067/MEaSUREs/GLDTA/XAERDT_L2_ABI_G16.001, 2019a.

NASA: ABI/GOES-17 Dark Target Aerosol 10-Min L2 Full Disk 10 km, https://doi.org/10.5067/MEaSUREs/GLDTA/XAERDT_L2_ABI_G17.001, https://doi.org/10.5067/MEaSUREs/GLDTA/XAERDT_L2_ABI_G17.001, 2019b.

NASA/METI/AIST/Japan Spacesystems, and US/Japan ASTER Science Team: ASTER Global Digital Elevation Model V003, distributed by NASA EOSDIS Land Processes DAAC, 2019.

Nguyen, T. N. T., Mantovani, S., Campalani, P., and Limone, G. P.: DOWNSCALING AEROSOL OPTICAL THICKNESS TO 1 KM2 SPATIAL RESOLUTION USING SUPPORT VECTOR REGRESSION REPLIED ON DOMAIN KNOWLEDGE, in: Proceedings of the 1st International Conference on Pattern Recognition Applications and Methods - Volume 2: ICPRAM,, pp. 230–239, INSTICC, SciTePress, https://doi.org/10.5220/0003791302300239, 2012.

Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., Desmaison, A., Köpf, A., Yang, E., DeVito, Z., Raison, M., Tejani, A., Chilamkurthy, S., Steiner, B., Fang, L., Bai, J., and Chintala, S.: PyTorch: An Imperative Style, High-Performance Deep Learning Library, https://arxiv.org/abs/1912.01703, 2019.

Remer, L. A., Kaufman, Y., Tanré, D., Mattoo, S., Chu, D., Martins, J. V., Li, R.-R., Ichoku, C., Levy, R., Kleidman, R., et al.: The MODIS aerosol algorithm, products, and validation, Journal of the atmospheric sciences, 62, 947–973, 2005.

Rosenblatt, F.: The perceptron: a probabilistic model for information storage and organization in the brain., Psychological review, 65 6, 386–408, https://api.semanticscholar.org/CorpusID:12781225, 1958.

Sayeed, A., Gupta, P., Henderson, B., Kondragunta, S., Zhang, H., and Liu, Y.: GOES-R PM2.5 Evaluation and Bias Correction: A Deep Learning Approach, Earth and Space Science, 12, https://doi.org/10.1029/2024EA004012, 2025.

Sulla-Menashe, D. and Friedl, M. A.: User guide to collection 6 MODIS land cover (MCD12Q1 and MCD12C1) product, USGS: Reston, VA, USA, 1, 18, 2018.

Taskinen, H., Väisänen, A., Hatakka, L., Virtanen, T. H., Lähivaara, T., Arola, A., Kolehmainen, V., and Lipponen, A.: High-Resolution Post-Process Corrected Satellite AOD, Geophysical Research Letters, 49, e2022GL099 733, https://doi.org/https://doi.org/10.1029/2022GL099733, e2022GL099733 2022GL099733, 2022.

Thangavel, P., Park, D., and Lee, Y.-C.: Recent Insights into Particulate Matter (PM2.5)-Mediated Toxicity in Humans: An Overview, International Journal of Environmental Research and Public Health, 19, 7511, https://doi.org/10.3390/ijerph19127511, 2022.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., and Polosukhin, I.: Attention Is All You Need, https://doi.org/10.48550/arXiv.1706.03762, arXiv:1706.03762 [cs], 2023.

Zhang, C., Niu, X., Wu, H., Ding, Z., Chan, K. L., Kim, J., Wagner, T., and Liu, C.: Unleashing the potential of geostationary satellite observations in air quality forecasting through artificial intelligence techniques, Atmospheric Chemistry and Physics, 25, 759–770, https://doi.org/10.5194/acp-25-759-2025, publisher: Copernicus GmbH, 2025.

Zhang, H. and Kondragunta, S.: Daily and Hourly Surface PM2.5 Estimation From Satellite AOD, Earth and Space Science, 8, e2020EA001 599, https://doi.org/10.1029/2020EA001599, _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2020EA001599, 2021.

Zhang, H., Wei, Z., Henderson, B. H., Anenberg, S. C., O'Dell, K., and Kondragunta, S.: Nowcasting applications of geostationary satellite hourly surface PM 2.5 Data, Weather and forecasting, 37, 2313–2329, 2022.