High-resolution post-process corrected satellite AOD

Henri Taskinen[1], Arttu Väisänen[1], Lauri Hatakka[2], Timo H. Virtanen[2], Timo Lähivaara[1], Antti Arola[2], Ville Kolehmainen[1], and Antti Lipponen[2]

[1]Department of Applied Physics, University of Eastern Finland, Kuopio, Finland

[2]Atmospheric Research Centre of Eastern Finland, Finnish Meteorological Institute, Kuopio, Finland

Key Points:

- Satellites produce global aerosol data, however, these data often suffer from low accuracy and spatial resolution due to data aggregation.
- Machine-learning-based post-process correction leads to a significant improvement in the AOD accuracy over the conventional retrievals.
- Post-process correction approach is also efficient for spatial downscaling of satellite aerosol data leading to high resolution AOD.

Corresponding author: Antti Lipponen, `antti.lipponen@fmi.fi`

Abstract

Poor air quality poses a great threat to human health. Accurate high-resolution satellite remote sensing of atmospheric aerosols would highly benefit satellite-based air quality estimates. We have developed and validated a post-process correction and downscaling approach for satellite remote sensing of aerosols. We use NASA's Moderate Resolution Imaging Spectroradiometer (MODIS) aerosol optical depth (AOD) over Washington D.C. - Baltimore area during the Distributed Regional Aerosol Gridded Observation Networks (DRAGON) campaign in 2011 to evaluate our approach. We derive and evaluate the AOD fields with high 250 meter resolution. The results show that the post-process correction approach is suitable for deriving downscaled, high-resolution AOD estimates and significantly improves the accuracy of the AOD retrievals.

Plain Language Summary

Satellites collect information about our atmosphere. We often use satellite data to monitor, for example, the atmospheric aerosols, which are small solid and liquid particles in the air. However, the massive amount of satellite data and partially unknown atmospheric processes and land surface properties force us to use simplified computations in aerosol monitoring. Unfortunately, the simplified computations lead to sub-optimal accuracy and low resolution in aerosol data. In this work, we have developed a new machine-learning-based algorithm that is used together with conventional satellite data processing to improve the data. Our algorithm takes advantage of accurate ground-based measurements. As a result, it significantly improves the accuracy and resolution of the aerosol data. We improved the resolution of satellite-based aerosol data from 3 km to 250 meters. New high-resolution data may allow some new applications for the satellite data, such as street-level air quality monitoring.

1 Introduction

Poor air quality poses a great threat to human health. New World Health Organization (WHO) Global Air Quality Guidelines published in September 2021 provide clear evidence of air pollution's damage to human health at lower concentrations than previously conceived (World Health Organization, 2021). WHO estimates that people's exposure to air pollution causes 7 million premature deaths every year. A key indicator in monitoring air quality and epidemiological studies is the PM2.5 parameter, the dry

mass concentration of fine particulate matter with an aerodynamic diameter less than 2.5 micrometers (μg/m^3 of air). Fine particulate matter originates from vehicle emissions, coal-burning, industrial emissions, and other human and natural sources. Air quality monitoring networks often utilize in-situ measurements that measure air quality at pointwise locations. However, remote sensing and satellite observations are needed to get better spatial coverage of air quality estimates over large regions.

The most widely used high-resolution satellite-based air quality datasets are based on downscaled aerosol optical depth (AOD) estimates combined with auxiliary data (Shaddick et al., 2018; van Donkelaar et al., 2019; Hammer et al., 2020; van Donkelaar et al., 2021). Auxiliary data brings, for example, aerosol vertical distribution and composition information to the computations. Then, statistical methods such as land use regression or graphically weighted regression are used to combine the available information and obtain an estimate for the surface PM2.5. An improvement in AOD estimates' accuracy and resolution would directly translate into an improvement in PM2.5 estimates. Therefore, high accuracy, high-resolution satellite aerosol retrieval would greatly benefit satellite-based estimation of surface PM2.5.

Moderate spatial resolution imaging spectroradiometers, such as Moderate Resolution Imaging Spectroradiometer (MODIS), Ocean and Land Colour Instrument (OLCI), and Sea and Land Surface Temperature Radiometer (SLSTR), with a native resolution of some hundreds of meters at best, would be an excellent and openly available source of satellite data to be used in air quality retrievals. Compared to optical instruments with high-resolution of tens of meters or even better, these moderate resolution instruments have advantageous characteristics for air quality retrievals, including wide swath, frequent return times, good signal-to-noise ratio, and broad spectral coverage. In the operational aerosol products of these instruments, however, the aerosol properties are based on aggregated data due to high computational costs and biases related to the aerosol retrieval algorithms. As a result, the most widely used aerosol data products have a spatial resolution of 3-10 km (Levy et al., 2013).

Recent advances in new methods to combine conventional retrieval algorithms and machine learning have significantly improved satellite AOD estimate accuracy (Lipponen et al., 2021, 2022). For example, the post-process correction approach for satellite AOD retrieval uses a machine learning-based model to predict the retrieval error in the satel-

lite AOD and uses that prediction to correct the retrieval. Previous studies have shown that the combination of physics-based retrievals and machine learning leads to better accuracy than a machine learning-based approach alone.

This study develops and validates the post-process correction and downscaling approach for MODIS AOD over Washington D.C. - Baltimore area during the Distributed Regional Aerosol Gridded Observation Networks (DRAGON) campaign in 2011 (e.g. Garay et al. (2017); Virtanen et al. (2018)). The DRAGON campaign provides very dense coverage of accurate ground-based AERONET AOD measurements for validation of high-resolution AOD products. Therefore, the DRAGON campaign is a unique setting to validate the post-process correction of high-resolution AOD retrievals. Furthermore, the recently published post-process correction approaches have only been applied to the correction of the satellite retrievals and not for the additional downscaling of the data to high spatial resolution. In this study, we take advantage of the high-resolution MODIS level-1 observations and, in addition to correction, downscale the AOD to 250 meter spatial resolution.

2 Materials and Methods

In this study, we develop and validate the post-process correction and downscaling approach to satellite aerosol retrievals. We apply the correction to MODIS 3 km AOD and use the DRAGON 2011 campaign aerosol data to validate the satellite retrievals. We downscale the AOD spatial resolution by mapping the aggregated data to a high-resolution grid corresponding to the best native resolution of the MODIS instrument, 250 meters, and evaluate the accuracy of the corrected high-resolution AOD data.

2.1 Post-process correction satellite retrievals

Let

$$y = f(x) \tag{1}$$

be an accurate satellite retrieval algorithm. Here y denotes the retrieval algorithm outputs, such as AOD, f is an accurate retrieval algorithm, and x are the inputs for the retrieval algorithm such as measurement geometry information and satellite measured TOA reflectances. In reality, however, due to, for example, complex and partially unknown atmospheric parameters and surface reflectance, accurate retrieval algorithms do not ex-

ist. In practice, the retrievals are computed with an approximative retrieval algorithm $\tilde{f}$. The accurate retrieval algorithm (1) can be written as

$$
\begin{aligned}
y & = f(x) + \tilde{f}(x) - \tilde{f}(x) && (2) \\
& = \tilde{f}(x) + \left[f(x) - \tilde{f}(x) \right] && (3) \\
& = \tilde{f}(x) + e(x), && (4)
\end{aligned}
$$

where e denotes the retrieval error.

In the conventional supervised machine learning-based approach, the aim is to train a model to directly predict the satellite retrieval outputs y given the inputs x. These trained models approximate the accurate retrieval algorithm f. As these models rely on the use of machine learning models only, we refer to these models as fully learned models.

In the post-process correction approach, the aim is to train a machine-learning-based model to predict the retrieval error e and employ (4) to compute the retrieval output. This post-process correction approach combines the physics-based retrieval $\tilde{f}$ and machine learning. We expect the retrieval error e to be a less complex function than the full retrieval algorithm f to be learned from finite number of learning data, and thus expect the post-process correction to result in more accurate retrieval than the fully learned model.

In the development of the post-process correction approach, we noted an interesting similarity between the post-process correction model architecture and the widely used, recently developed neural network architecture of residual neural networks (ResNet). In ResNets, the network architecture is constructed so that skip connections are added to allow information skip over some neural network layers and act as inputs for the subsequent layers. AOD post-process correction can be thought of to have a similar skip connection for a subset of input data corresponding to the AOD to be corrected. The skip connection for AOD is created from the model inputs directly to the final output layer of the neural network. Having these similarities, however, the starting points of the ResNet and post-process correction are fundamentally different - post-process correction aims at correcting the output of a physics-based retrieval algorithm, and the ResNets have been developed to tackle the problem of vanishing gradients in the training of deep neural networks. As the practical implementations are quite similar between these two models, we also expect that the post-process correction model may be relatively tolerant against

the problem of vanishing gradients in the training of the neural networks. Therefore, we expect this feature to even further improve the accuracy of the post-process correction models. We also tested the ResNet-type of algorithm in this study and found it perform similarly as the post-process correction model and therefore do not show the results here.

2.2 Training and validation of the neural network models

The dense network of AERONET stations available in the DRAGON campaign allowed validation of the downscaled 250 m resolution aerosol product. In the validation, we used a cross-validation approach in which some of the AERONET stations were used for training the models, while others were used in the validation of the results. We train and validate both fully learned and post-process correction models to compare the performance between these two approaches. For training and validation, we randomly divided the MODIS-AERONET collocated pixels into three separate groups by AERONET station.

The division is carried out by AERONET station to avoid too similar data samples between the training and validation datasets. As is well known, too similar data samples could potentially lead to over-optimistic results. The evaluation of the accuracy of the models was carried out using cross-validation so that one group was used as training data, one group was used as validation to monitor the convergence of the training, and the resulting model was applied to the third independent group of test data. This was repeated three times so that each AERONET station was present once in the validation data. Training of every model was carried out 20 times with different random initial weights of the neural networks and the best performing model was always selected to be used in the evaluation of the test data.

Based on results shown in Lipponen et al. (2022) and our preliminary tests, we utilized fully connected feed-forward networks and fixed both the fully learned and post-process correction neural network architectures to 3 hidden layers. In addition, we set the batch size for training to 8, used an initial learning rate of 5×10^{-5} with the Adam optimization algorithm, used mean square error loss, and selected rectified linear unit (ReLU) as the activation function for all hidden layers and linear activation for the output.

To determine the optimal number of neurons for each layer of the neural networks, the Asynchronous Successive Halving (ASHA) method was used (Li et al., 2020). The ASHA optimization tested all combinations of 8, 32, 128, 512, and 1024 neurons for each layer as a grid search and computed the validation losses for each trained network. The optimization was repeated for 10 different random initializations of the neural networks and the best average validation loss neural network structures were selected for the final models to be trained. The optimal number of neurons for the three layers in the fully learned model was found as 128, 1024, 128, and for the post-process correction model 512, 32, 512.

We also carried out full processing of the MODIS AOD data in the region of interest (ROI) during the whole DRAGON campaign period to produce AOD maps. In contrast to cross-validation evaluation of the model accuracies, for this use we trained separate fully learned and post-process correction models using full datasets. In the training of the full dataset models, eight AERONET stations were selected to be validation stations that were used to monitor the training convergence and the rest of the data were used as training data. The training of these models was also carried out 20 times with different random initial weights and the models with the smallest validation loss were selected as the final models for the data processing.

The dataset used for the training and validation of the models consisted of 2728 samples with 26 and 33 input parameters for the fully learned and post-process correction models, respectively. The model training and data processing used in this study were not computationally very expensive and were carried out using a regular laptop computer without GPU capabilities.

2.3 Region of interest and data gridding

We use Washington, D.C. - Baltimore, Maryland, USA, region as our region of interest (ROI). The size of the ROI is 120 by 120 km and it is divided into 480 by 480 pixel grid with 250 meter pixel size. The Universal Transverse Mercator (UTM) map projection zone 18 is used for constructing the grid. All spatially distributed parameters are projected to this grid using nearest-neighbor interpolation before the model training and evaluation computations. The ROI is shown in Figure 1.

2.4 Satellite data

We use MODIS data of both Terra and Aqua satellites in this study. As the top-of-atmosphere (TOA) reflectance data, MODIS collection 6.1 level-1b data of bands 1-13, 15, and 19-22 were used. Bands 14, 16-18 were not used as a significant portion of the data was missing. MODIS bands 1 and 2 have native spatial resolution of 250 meters, bands 3-7 500 meters, and other bands 1 km at nadir.

The physics-based AOD retrieval $\tilde{f}(x)$ we used was the MODIS collection 6.1 aerosol data product MOD04_3KM Dark Target AOD over land with native 3 km spatial resolution. The measurement geometry information, solar and view zenith and azimuth angles, the scattering and glint angles, and also the topographic land altitude were also taken from the MOD04_3KM data product to the aerosol retrieval data.

We only accepted pixels with the MODIS view zenith angle less than 50 degrees. This selection was made to restrict the pixels to the central part of swath pixels. As the MODIS pixel size grows towards the edges of the swath, this selection kept the pixel size reasonable and filtered out too large pixels to be used for high-resolution retrievals.

2.5 Ground-based AERONET AOD data

As an accurate ground-based reference AOD data, we used the sunphotometer-based AOD of AERONET stations in the ROI during the DRAGON 2011 campaign. The duration of the DRAGON campaign was from June 1 to August 15, 2011. The DRAGON campaign consisted of more than 40 AERONET stations deployed to Washington D.C. - Baltimore region. In this study, we use level 2.0 AERONET AOD at 550 nm which is computed from AOD measurement at 500 nm and Ångström exponent for 440-870 nm. AERONET AOD at visible wavelengths have been reported to have a low uncertainty of 0.01 and therefore we consider the AERONET AOD estimates accurate (Eck et al., 1999). We use AERONET AOD both for training and validation of our models.

2.6 Auxiliary high-resolution data

As we aim at high spatial resolution aerosol data, in addition to satellite-based data, we also used a high-resolution digital elevation model (DEM) as auxiliary data for machine-

learning-based models. These auxiliary data were added as additional input parameters in the models.

We used GMTED2010 DEM which has 7.5 arc-seconds (about 225 meters) resolution (Danielson & Gesch, 2011). For our use, GMTED2010 DEM data was interpolated to the 250 meter grid in our ROI.

In the post-process correction model, surface reflectance at three different wavelengths and AOD fields at four different wavelengths retrieved with Dark Target were added as auxiliary inputs. As the 3 km Dark Target AOD fields are noisy and therefore contain sharp changes between the neighboring 3 km pixels, we smoothed the AOD fields before the post-process correction. A 2D convolution using a Gaussian kernel with a 3 km standard deviation was used for the AOD field smoothing.

2.7 MODIS-AERONET collocation

For the MODIS-AERONET collocation, we followed a similar protocol as in Bilal et al. (2013). That is, we required the distance between the high-resolution grid pixel center and AERONET station to be less than 750 meters and restricted our data to a maximum number of nine pixels per overpass around an AERONET station. For temporal collocation, we used a maximum time difference of $\pm$ 250 seconds. With our collocation protocol and data criteria, we ended up having data from 37 different AERONET stations. The excluded stations did not contain any valid pixels.

3 Results

Figure 2 shows the MODIS-AERONET AOD comparison for the Dark Target, fully learned model, and the post-process corrected model. The post-process corrected model is clearly the best performing model. The post-process corrected model has all the best metrics with the only exception of bias. The bias of the fully learned model and the post-process corrected models are the same (-0.004). The root mean squared error (RMSE) of the post-process corrected data is only 0.038, which is 28% smaller than the fully learned model RMSE and about 64% smaller than the Dark Target RMSE. The fully learned model has some problems in predicting large AOD values and the highest AOD values are significantly underestimated whereas the highest AODs predicted by the post-process corrected model are all within the Dark Target expected error (EE) envelope. The Dark

Target, on the contrary, is overestimating AOD in almost all AERONET overpasses. This type of systematic over- or underestimation is not observed in machine-learning-based AODs. The maximum absolute value of the retrieval error in the post-process corrected AOD is less than half of the one obtained with the fully learned model. The fraction of retrievals within the Dark Target EE envelope is also clearly better in the post-process corrected AOD (96.4%) than in the fully learned AOD retrievals (91.6%). The Dark Target fraction within EE envelope was 67.2%.

All MODIS overpasses of the DRAGON campaign were processed using the final trained fully learned and post-process correction models. Figure 1 shows the average AOD and the average of daily AOD anomalies with respect the daily mean of the AOD in the ROI for the DRAGON campaign duration for the Dark Target, fully learned model, and post-process correction model. Average AERONET AOD for the DRAGON campaign duration collocated with MODIS overpasses are also shown. The AOD anomalies were constructed by first computing daily anomalies as the differences of the full AOD fields and the daily ROI average AODs and then temporally averaging the daily anomalies over the whole duration of the DRAGON campaign. On average, there were 2.8 daily MODIS overpasses during the campaign. The map figure shows that the Dark Target AODs are significantly higher than the AOD obtained with fully learned or post-process correction approaches. The machine-learning-based AOD datasets match better the AOD from the AERONET stations. Near the coastline both of the machine-learning-based approaches seem to work well and do not show any clear anomalies near the coast whereas Dark Target shows elevated AOD values near the coast with no clear physical explanation to it. AERONET observations do not show elevated AOD values near the coastline. The AOD anomaly maps show higher AOD values over the urban areas in Washington D.C. and Baltimore and between them. The average positive AOD anomaly is stronger in the DT dataset, about 0.2 over the densest urban areas, than those based on machine learning that have anomaly value of about 0.1 over the densest urban areas. Both of the machine-learning-based datasets have a good agreement with the AERONET over urban areas.

To evaluate the possible contribution of surface reflectance to the retrieved AOD, we studied the correlation between the MODIS nadir bidirectional reflectance distribution function (BRDF)-adjusted surface reflectance from the MCD43A4 data product and AOD over the ROI. This surface reflectance dataset is based on atmospheric correction that treats the aerosols independently of the DT algorithm. The same daily surface re-
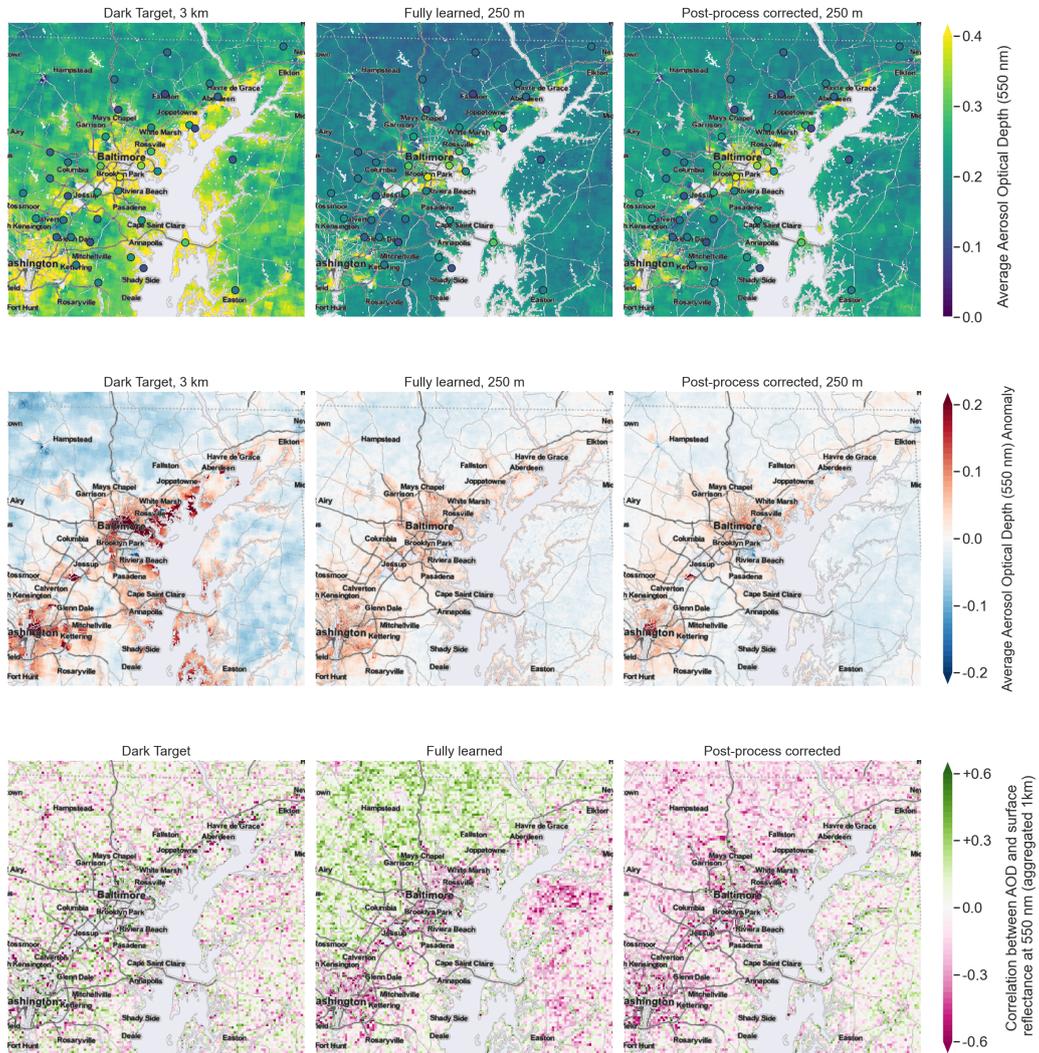
Figure 1. Top row: average MODIS AOD at 550 nm during the DRAGON campaign. Middle row: Average daily MODIS AOD at 550 nm anomaly. Bottom row: Correlation between AOD and MODIS surface reflectance at 550 nm. Correlation is shown with data aggregated to 1 km spatial resolution. Left column: Dark Target. Middle column: Fully learned model. Right column: Post-process corrected Dark Target.
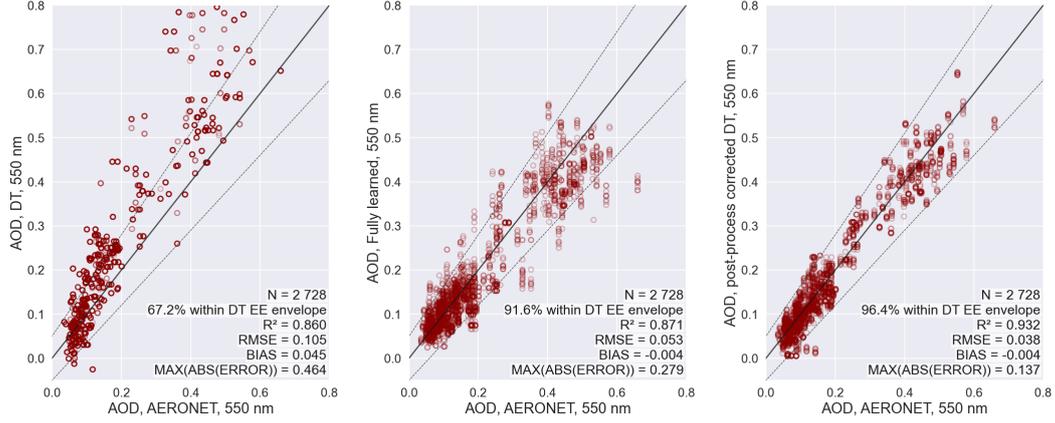
Figure 2. MODIS-AERONET AOD at 550 nm comparison. Left: Dark Target (DT) 3 km. Middle: Fully learned model. Right: Post-process corrected Dark Target. Each marker in the image corresponds to a single collocated MODIS-AERONET overpass. The following metrics are shown for each model: the number of samples N, the fraction of retrievals within the Dark Target expected error (DT EE) envelope $0.05 \pm 15\%$, root mean squared error (RMSE), mean bias (BIAS), and the maximum absolute error (MAX(ABS(ERROR))).

flectance data were used for computing the correlations for all AOD datasets. The correlations between AOD and surface reflectance at 550 nm is shown in Figure 1. To reduce the noise in the correlation maps, the correlations were computed with data aggregated to 1 km spatial resolution. The maps showing the correlation between the surface reflectance and AOD for DT and post-process corrected AOD do not have a clear contrast between the dense urban and surrounding regions. Both of the machine-learning-based datasets generally show opposite signed correlation between the southeast and northwest corners of the ROI. The fully learned model results in mainly positive correlation on the northwest corner and negative on the southeast corner, and the post-process corrected AOD behaves the opposite way. We did not find a clear explanation for this behavior. For the fully learned model AOD and surface reflectance correlation, the urban areas of Washington D.C. and Baltimore are visible as they have a negative correlation, whereas the nearby surrounding regions have a positive correlation. Furthermore, the fully learned model AOD was observed to be underestimated in most retrievals over AERONET stations associated with the highest surface reflectances. This type of behaviour was not observed with DT or the post-process corrected AOD.

Regardless of the DRAGON campaign being a unique campaign for high-resolution validation of satellite data, the distance between the stations is still not good enough for very high resolution evaluation. In this study, regardless of the dense and unique AERONET setting in the ROI, the average distance between two AERONET stations was 1.2 km, and the average distance of a pixel in our ROI to the nearest AERONET station about 12 km. As the distances are significantly larger than our pixel size 250 meters, we need to visually inspect these retrieval maps and see some features to assess the results. We do not observe any very local and clearly distinctive AOD features in any of the average AOD datasets. This was expected as aerosols are easily transported and mixed locally in the atmosphere and we expected the average AOD fields to be relatively smooth. Over cities, the average satellite-based AOD fields are clearly higher than in the surrounding regions in all datasets.

We used SHapley Additive exPlanations (SHAP) to explain the variables that have the largest impact on the correction of AOD (Lundberg & Lee, 2017). We used the DeepExplainer model of the Shap Python library and computed the average SHAP values for 10000 randomly selected pixels based on set of background values of another 10000 randomly selected pixels. The results show that, on average, the most significant variables to explain the retrieval error correction terms are the AOD at 440 nm, 675 nm, and 550 nm, the GMTED2010 surface elevation, TOA reflectances at bands 11 and 9, and the AOD at 2100 nm. The mentioned variables explain about 60% of the AOD correction term. As the DT AOD is typically overestimated, it is expected that AOD input terms explain quite a large fraction of the correction. Topographic altitude terms in the list for this ROI probably act as a proxy for some other quantity and indicate some indirect effect, such as distance from the coastline, not the real dependency of the AOD correction to surface altitude. The most important TOA reflectance bands correspond to wavelengths 526-536 nm and 438-448 nm.

Figure 3 shows AOD time series for the DRAGON_ARNLS station and AOD fields corresponding to two MODIS overpasses for the Dark Target, the fully learned model, and the post-process correction. The DRAGON_ARNLS station was one of the 8 stations not used in the actual training of the models but the convergence monitoring only. The figure shows Dark Target mostly overestimates AOD at this location. Both of the machine-learning-based models follow well the changes in AOD over the whole campaign duration. The AOD maps corresponding to single overpasses clearly show the coarser

3 km resolution of the Dark Target. Both the fully learned and post-process corrected

models result in smooth AOD fields. Regardless of the high spatial resolution and the

independent treatment of all pixels in both fully learned and post-process correction ap-

proaches, the AOD fields are highly consistent and look feasible. There are no single pix-

els that would have significantly higher or lower AOD than the neighboring pixels. The

AOD fields do not seem to contain significant noise. The MODIS overpass correspond-

ing to 24 July 2011 15:45Z has a significant fraction of missing data. This indicates most

of the ROI was covered by clouds at the satellite overpass time. The maps show that all

retrieval approaches do work well in the vicinity of clouds and do not show any behav-

ior correlated for example with the distance from the cloud.

4 Summary and Conclusions

In this study, we developed a machine-learning-based post-process correction ap-

proach to correct and downscale the MODIS AOD. The post-process correction approach

combines both the physics-based conventional retrieval algorithms and machine learn-

ing. In the development, we concentrated on downscaling and improving the accuracy

of the satellite-based AOD. We used the AERONET year 2011 DRAGON campaign over

the Baltimore - Washington D.C. area to evaluate our approach.

We compared our post-process corrected AOD to the 3 km MODIS Dark Target

aerosol data product and the fully-learned machine learning approach in which the machine-

learning model was trained to carry out the AOD retrieval. The results show that our

post-process correction approach is suitable for the downscaling of existing aerosol data

products and results in clearly improved AOD accuracy over both the Dark Target and

fully learned AOD.

Accurate high-resolution AOD would be highly beneficial, for example, for the air

quality applications. Improvements in satellite-based AOD estimates will directly trans-

late into improved PM2.5 estimates. The post-process corrected AOD serves as a highly

promising starting point for satellite-based air quality PM2.5 estimates.

Extension of the post-process correction approach is straightforward to other re-

trieved quantities and instruments as long as suitable training data is available. We ex-

pect the post-process correction approach to be highly beneficial, especially for appli-

cations with scarce training data. With a small number of training data samples, the fully
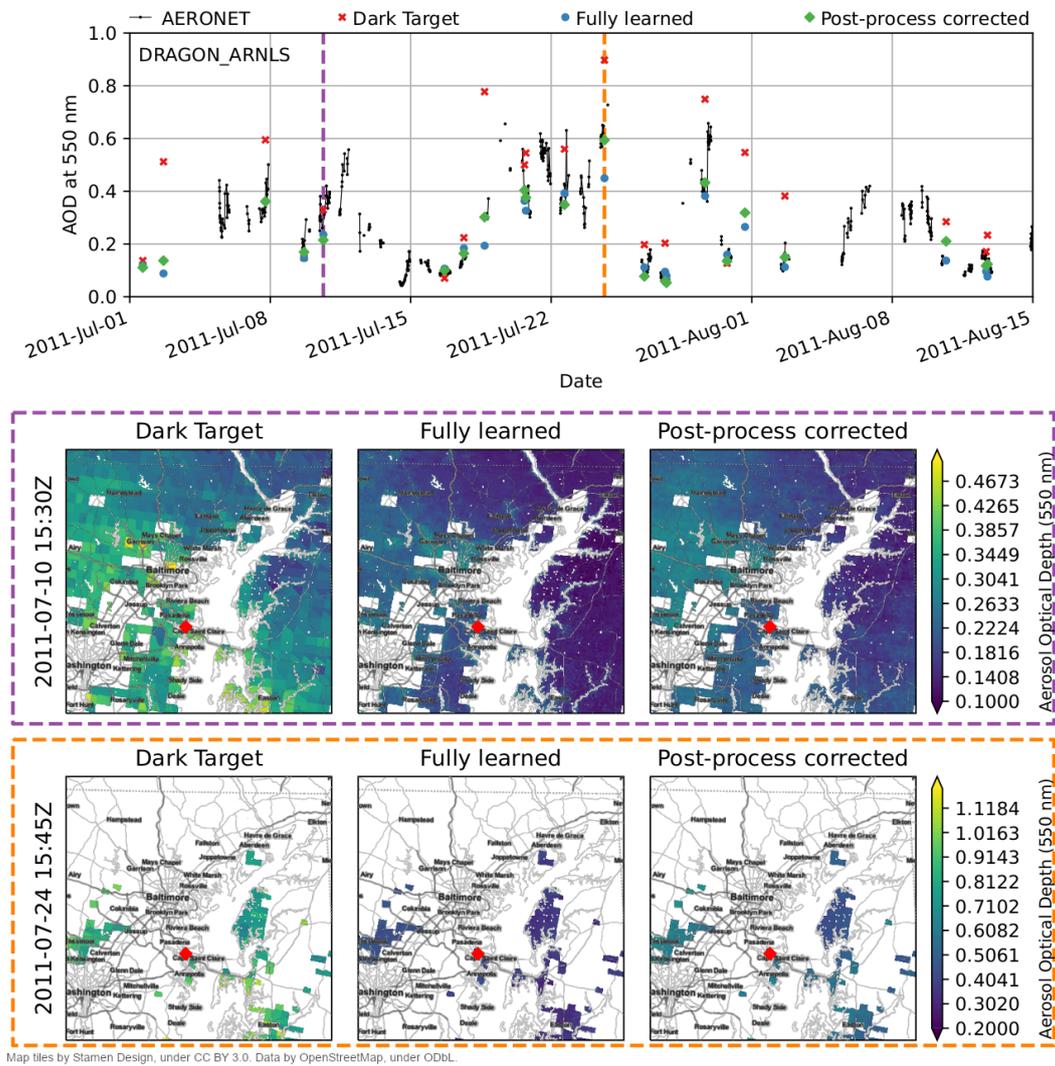
Figure 3. Top row: AOD 550 nm time series for the DRAGON_ARNLS station. Middle row: MODIS overpass on 10 July 2011 at 15:30Z. Left: Dark Target AOD. Middle: Fully learned AOD. Right: Post-process corrected AOD. Bottom row: MODIS overpass on 24 July 2011 at 15:45Z. Left: Dark Target AOD. Middle: Fully learned AOD. Right: Post-process corrected AOD. The purple and orange dashed lines on the top time series figure correspond to middle and bottom row overpasses, respectively. The red diamond symbol on the maps indicates the location of the AERONET station.

learned models may generalize poorly and fail with inputs not present in the training data. By the support of physics-based retrieval algorithms, we expect better generalization in the prediction of the retrieval error than the retrieval itself and therefore expect post-process correction to be a successful approach in many applications.

5 Open Research

The AERONET DRAGON 2011 campaign data used in this study are openly available and were obtained from the NASA AERONET website `https://aeronet.gsfc.nasa.gov/`. All MODIS data used in this study are open data and were obtained from the NASA Level-1 and Atmosphere Archive & Distribution System Distributed Active Archive Center (LAADS DAAC) `https://ladsweb.modaps.eosdis.nasa.gov/`. The GMTED2010 digital elevation model data is public domain and were obtained from the U.S. Geological Survey Earth Resources Observation and Science (EROS) Center website `https://www.usgs.gov/centers/eros/science/usgs-eros-archive-digital-elevation-global-multi-resolution-terrain-elevation`.

Software (version 1.0.0 / 24 May 2022) associated with this manuscript to carry out and reproduce the data download, pre-processing, downscaling and post-process correction and validation of AOD is licensed under MIT and published on GitHub `https://github.com/TUT-ISI/DRAGONcorr/` (Lipponen, 2022).

References

Bilal, M., Nichol, J. E., Bleiweiss, M. P., & Dubois, D. (2013). A simplified high resolution MODIS aerosol retrieval algorithm (SARA) for use over mixed surfaces. *Remote sensing of environment*, *136*, 135–145.

Danielson, J. J., & Gesch, D. B. (2011). *Global multi-resolution terrain elevation data 2010 (GMTED2010)*. US Department of the Interior, US Geological Survey Washington, DC, USA.

Eck, T. F., Holben, B., Reid, J., Dubovik, O., Smirnov, A., O'neill, N., ... Kinne, S. (1999). Wavelength dependence of the optical depth of biomass burning, urban, and desert dust aerosols. *Journal of Geophysical Research: Atmospheres*, *104*(D24), 31333–31349.

Garay, M. J., Kalashnikova, O. V., & Bull, M. A. (2017). Development and assessment of a higher-spatial-resolution (4.4 km) MISR aerosol optical depth product using AERONET-DRAGON data. *Atmospheric Chemistry and Physics*, *17*(8), 5095–5106.

Hammer, M. S., van Donkelaar, A., Li, C., Lyapustin, A., Sayer, A. M., Hsu, N. C., ... Martin, R. V. (2020). Global estimates and long-term trends of fine particulate matter concentrations (1998–2018). *Environmental Science & Technology*, *54*(13), 7879–7890.

Levy, R., Mattoo, S., Munchak, L., Remer, L., Sayer, A., Patadia, F., & Hsu, N. (2013). The collection 6 MODIS aerosol products over land and ocean. *Atmospheric Measurement Techniques*, *6*(11), 2989–3034.

Li, L., Jamieson, K., Rostamizadeh, A., Gonina, E., Ben-Tzur, J., Hardt, M., ... Talwalkar, A. (2020). A system for massively parallel hyperparameter tuning. *Proceedings of Machine Learning and Systems*, *2*, 230–246.

Lipponen, A. (2022). MODIS aerosol optical depth spatial downscaling and post-process correction for the DRAGON campaign 2011. 24 May 2022 release (version 1.0.0) [Software]. *Zenodo*. doi: 10.5281/zenodo.6575797

Lipponen, A., Kolehmainen, V., Kolmonen, P., Kukkurainen, A., Mielonen, T., Sabater, N., ... Arola, A. (2021). Model-enforced post-process correction of satellite aerosol retrievals. *Atmospheric Measurement Techniques*, *14*(4), 2981–2992.

Lipponen, A., Reinvall, J., Väisänen, A., Taskinen, H., Lähivaara, T., Sogacheva,

L., ... Kolehmainen, V. (2022). Deep-learning-based post-process correction of the aerosol parameters in the high-resolution Sentinel-3 Level-2 Synergy product. *Atmospheric Measurement Techniques*, *15*(4), 895–914.

Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. In I. Guyon et al. (Eds.), *Advances in neural information processing systems 30* (pp. 4765–4774). Curran Associates, Inc. Retrieved from http://papers.nips.cc/paper/7062-a-unified-approach-to-interpreting-model-predictions.pdf

Shaddick, G., Thomas, M. L., Amini, H., Broday, D., Cohen, A., Frostad, J., ... Brauer, M. (2018). Data integration for the assessment of population exposure to ambient air pollution for global burden of disease assessment. *Environmental Science & Technology*, *52*(16), 9069–9078.

van Donkelaar, A., Hammer, M. S., Bindle, L., Brauer, M., Brook, J. R., Garay, M. J., ... Martin, R. V. (2021). Monthly global estimates of fine particulate matter and their uncertainty. *Environmental Science & Technology*, *55*(22), 15287–15300.

van Donkelaar, A., Martin, R. V., Li, C., & Burnett, R. T. (2019). Regional estimates of chemical composition of fine particulate matter using a combined geoscience-statistical method with information from satellites, models, and monitors. *Environmental Science & Technology*, *53*(5), 2595–2611.

Virtanen, T. H., Kolmonen, P., Sogacheva, L., Rodríguez, E., Saponaro, G., & de Leeuw, G. (2018). Collocation mismatch uncertainties in satellite aerosol retrieval validation. *Atmospheric Measurement Techniques*, *11*(2), 925–938.

World Health Organization. (2021). *Who global air quality guidelines: particulate matter (PM2.5 and PM10), ozone, nitrogen dioxide, sulfur dioxide and carbon monoxide* [Publications]. World Health Organization.