Near-cloud aerosol retrieval using machine learning techniques, and implied direct radiative effects

Yang, C. K. ORCID: https://orcid.org/0000-0001-7685-4481, Chiu, J. C. ORCID: https://orcid.org/0000-0002-8951-6913, Marshak, A. ORCID: https://orcid.org/0000-0002-3973-1359, Feingold, G. ORCID: https://orcid.org/0000-0002-0774-2926, Várnai, T. ORCID: https://orcid.org/0000-0002-7419-2522, Wen, G. ORCID: https://orcid.org/0000-0003-2977-4993, Yamaguchi, T. ORCID: https://orcid.org/0000-0001-8059-0757 and Van Leeuwen, P. J. ORCID: https://orcid.org/0000-0003-2325-5340 (2022) Near-cloud aerosol retrieval using machine learning techniques, and implied direct radiative effects. Geophysical Research Letters, 49 (20). e2022GL098274. ISSN 0094-8276 doi: https://doi.org/10.1029/2022GL098274 Available at https://centaur.reading.ac.uk/109109/

Geophysical Research Letters

RESEARCH LETTER
10.1029/2022GL098274

Near-Cloud Aerosol Retrieval Using Machine Learning Techniques, and Implied Direct Radiative Effects

C. Kevin Yang[1], J. Christine Chiu[1], Alexander Marshak[2], Graham Feingold[3], Tamás Várnai[2,4], Guoyong Wen[2,5], Takanobu Yamaguchi[3,6], and Peter Jan van Leeuwen[1,7]

[1]Department of Atmospheric Science, Colorado State University, Fort Collins, CO, USA, [2]NASA Goddard Space Flight Center, Greenbelt, MD, USA, [3]NOAA Chemical Sciences Laboratory, Boulder, CO, USA, [4]Joint Center for Earth System Technology, University of Maryland Baltimore County, Baltimore, MD, USA, [5]GESTAR/Morgan State University, Baltimore, MD, USA, [6]Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, CO, USA, [7]Department of Meteorology, University of Reading, Reading, UK

Abstract There is a lack of satellite-based aerosol retrievals in the vicinity of low-topped clouds, mainly because reflectance from aerosols is overwhelmed by three-dimensional cloud radiative effects. To account for cloud radiative effects on reflectance observations, we develop a Convolutional Neural Network and retrieve aerosol optical depth (AOD) with 100–500 m horizontal resolution for all cloud-free regions regardless of their distances to clouds. The retrieval uncertainty is 0.01 + 5%AOD, and the mean bias is approximately −2%. In an application to satellite observations, aerosol hygroscopic growth due to humidification near clouds enhances AOD by 100% in regions within 1 km of cloud edges. The humidification effect leads to an overall 55% increase in the clear-sky aerosol direct radiative effect. Although this increase is based on a case study, it highlights the importance of aerosol retrievals in near-cloud regions, and the need to incorporate the humidification effect in radiative forcing estimates.

Plain Language Summary The presence of aerosols can heat or cool the atmosphere, depending on their interactions with clouds and radiation. These interactions remain one of the primary sources of uncertainty in climate change predictions. To understand the role of aerosols in climate and their interactions with clouds, reflectance measurements from satellites have been used to retrieve aerosol properties. However, these retrievals are typically available in regions far from clouds, but not in the vicinity of clouds, because the observed reflectance is dominated by nearby cloud scattering rather than by aerosols. Since more than half of cloud-free regions are within 4 km of low clouds, it is crucial to characterize the properties of aerosols near clouds. To tackle the issue, we developed a machine-learning based retrieval method. The new method characterizes cloud radiative effects, removes them from the observed reflectance, and then retrieves aerosol properties even in the vicinity of clouds. The retrieval uncertainty is comparable to benchmark products. This newly added capability allows us to fill the critical gap in current aerosol observations and better quantify how aerosols influence the Earth's radiation budget.

Key Points:

- A convolutional neural network is used to retrieve aerosol optical depth (AOD) with an uncertainty of 0.01 + 5%AOD in all cloud-free regions
- Due to aerosol hygroscopic growth, the optical depth of aerosols near clouds can be enhanced by 100% compared to those far from clouds
- The enhancement in AOD near clouds leads to an overall 55% increase in clear-sky aerosol direct radiative effects

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

C. K. Yang,
yang0920@rams.colostate.edu

Citation:

Yang, C. K., Chiu, J. C., Marshak, A., Feingold, G., Várnai, T., Wen, G., et al. (2022). Near-cloud aerosol retrieval using machine learning techniques, and implied direct radiative effects. *Geophysical Research Letters*, *49*, e2022GL098274. https://doi.org/10.1029/2022GL098274

Received 17 FEB 2022
Accepted 24 JUL 2022

1. Introduction

Aerosols continue to be one of the largest sources of uncertainty in quantifying the responses of our climate system to human impacts (Bellouin et al., 2020; Forster et al., 2021). Global aerosol observations largely rely on daytime satellite reflectance measurements. While aerosol retrievals in the region far from clouds have been developed and evaluated extensively, retrieving aerosol properties near clouds remains difficult. The difficulty arises, mainly because three-dimensional (3D) cloud radiative effects significantly enhance the observed reflectances and obscure signals from aerosols (e.g., Spencer et al., 2019; Stap et al., 2016; Várnai & Marshak, 2009; Wen et al., 2006). As a result, ∼20% of observed pixels in the Moderate Resolution Imaging Radiometer (MODIS) aerosol products are discarded (Schwarz et al., 2017). Since aerosols near clouds can be distinctly different from those far from clouds due to hygroscopic growth and other processes (Hoppel et al., 1986; Twohy et al., 2002, 2009), characterizing aerosols in the vicinity of clouds is crucial for estimating aerosol radiative forcing (Charlson et al., 2007; Koren et al., 2007).

To date, very few studies have attempted to retrieve aerosol properties in the vicinity of clouds and estimate the corresponding shortwave (SW) aerosol direct radiative effects (DRE). Using simplified aerosol property

estimates, Várnai and Marshak (2014) reported that aerosols near clouds increased instantaneous broadband reflected fluxes at the top of the atmosphere (TOA) by 6.4 W m^{-2} compared to those far from clouds. Additionally, Marshak et al. (2008, 2014) developed a two-layer model to remove reflectance enhancement induced by the interactions between clouds and the molecular layer above them. This model was applied to MODIS observations by Wen et al. (2013, 2016) for retrieving aerosol properties near clouds. Their method removes the 3D cloud radiative enhancements from the mean reflectance over a 10 × 10 km domain, which will require modifications to enable extracting near-cloud aerosol properties at higher horizontal resolutions.

Unlike the studies above mainly based on satellite measurements, Twohy et al. (2009) first examined how the properties of humidified aerosols vary with distance to clouds using aircraft in-situ measurements. Then, they derived the statistics of the fraction of cloud-free pixels that occurred at a given distance to clouds, using observations of the Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observation. By combining aerosol properties with cloud-free fraction statistics, they estimated the overall aerosol SW DRE can be 35%–65% larger if the humidification effects on aerosol hygroscopic growth were included. This approach, however, requires in-situ measurements, making it difficult to provide a global estimate of DRE of near-cloud aerosols.

To fill this critical gap in current aerosol observations in near-cloud regions, the objective of this study is to develop a new method that allows aerosol optical depth (AOD) retrievals in a fully 3D environment. This method uses a Convolutional Neural Network (CNN), trained by Large Eddy Simulation (LES) output from Jiang et al. (2009) and Yamaguchi et al. (2019), and reflectance calculated from a 3D radiative transfer model (Evans, 1998). We focus on cumulus regimes where 3D cloud radiative effects are expected to be most significant (Hogan & Shonk, 2013; Pincus et al., 2005). In Sections 2 and 3, a proof-of-concept case is presented, and an explainable Artificial Intelligence (AI) technique is used to understand how the CNN uses reflectance observations for retrievals. In Section 4, we retrieve AOD from MODIS observations and assess the SW DRE of near-cloud aerosols, using the same CNN but trained with different aerosol types from that in the proof-of-concept case.

2. The Convolutional Neural Network

2.1. The Training and the Configuration for the Proof-Of-Concept Case

The goal of our CNN is to predict AOD over a 2.5 × 2.5 km domain at a 100-m horizontal resolution from an input scene of reflectances at TOA. Due to 3D radiative effects, reflectances directly above this domain are influenced not only by aerosol and cloud properties underneath, but also by neighboring pixels. To incorporate information on neighboring pixels for appropriate AOD predictions, the input reflectance scene size must be larger than the output domain. Following the study of Okamura et al. (2017) for cloud retrievals, we chose an input scene of 4.5 × 4.5 km, that is, extending 1 km on all sides compared to the output domain.

Our training and testing datasets contain pairs of AOD fields and TOA reflectance fields; a flowchart is shown in Figure S1 in Supporting Information S1. To diversify aerosol conditions and cloud morphology in our training data set, four sets of LES output were used (see Figure 1 and Table S1 in Supporting Information S1). All datasets have a horizontal domain size of 48 × 48 km with a resolution of 100 m. The height of the domain is 5 km with a resolution of 40 m. Datasets 1 and 2 were generated using the forcing data collected from the Rain In Cumulus over Ocean (RICO) campaign (Rauber et al., 2007) with an initial ambient aerosol concentration of 200 cm^{-3} (Jiang et al., 2009). In contrast, Datasets 3 and 4 were generated from a cleaner environment with an initial aerosol concentration of 35 cm^{-3}, using forcing data from the Seven SouthEast Asian Studies (7SEAS) (Yamaguchi et al., 2019).

The aerosol type used in the LES is ammonium sulfate, and we use a cloud water content exceeding or equal to 0.01 g m^{-3} to define cloud. Cloud and aerosol optical properties in these snapshots were calculated using Mie theory, appropriate for ammonium sulfate (Curtis et al., 2007). With increasing humidity near clouds, the index of refraction of ammonium sulfate particles changes, and their effective radius increases due to water uptake. As a result, AOD is enhanced near cloud edges (see Figures 1e–1h).

The mean AODs from the 7SEAS and RICO snapshots are distinctly different (0.1 and 0.56, respectively), and the variability of AOD within each snapshot is small (Table S1 in Supporting Information S1). This introduces a large gap in AOD values in the training data set. Such a gap is not ideal because the CNN would not learn any

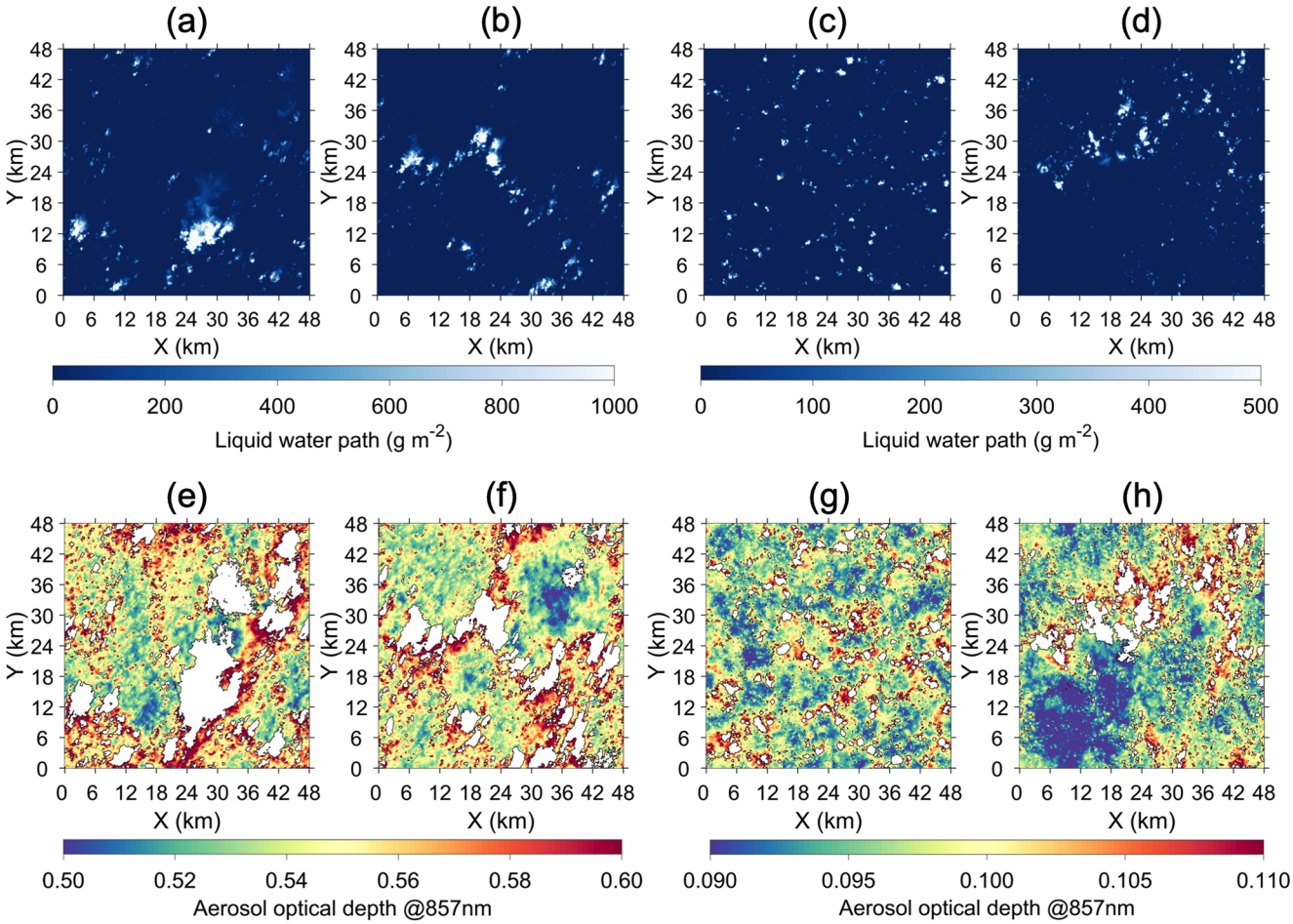

Figure 1. Snapshots of liquid water paths (a–d) and aerosol optical depths (e–h) for Datasets 1–4 (left to right, respectively). The boundaries of clouds (defined by liquid water path greater than zero) are outlined by black lines in (e–h). Note that the color scales are different between datasets.

AOD values in between, which may lead to large errors in predictions. To alleviate this issue, we perturbed the aerosol mass concentration fields homogeneously everywhere by a factor ranging between 0.05 and 1.0 in the RICO snapshots, and by a factor ranging between 0.5 and 5.0 in the 7SEAS snapshots. These factors were chosen to ensure some AOD overlap between RICO and 7SEAS snapshots. These perturbations led to 50 snapshots, 30 from RICO and 20 from 7SEAS.

From these 50 snapshots at their resolutions of 100 m, we then calculated the corresponding reflectance fields at 857 nm wavelength, one of the MODIS bands. The radiative transfer calculations were performed using the Spherical Harmonic Discrete Ordinate Method (Evans, 1998) under a 3D environment, assuming a solar zenith angle of 50°, a solar azimuth angle of 210° clockwise from the north, and a viewing zenith angle of 0°. Water vapor absorption and Rayleigh scattering were included, and ocean surface reflectances were computed with an assumed wind speed of 6 m s^{-1}. Once the pairs of AOD and reflectance fields were produced, we randomly sampled 4.5 × 4.5 km scenes from each 48 × 48 km snapshot to generate the training data set. For the testing data set, we sampled the scenes using a different data set—we rotated the 50 snapshots by 90° clockwise and recomputed the corresponding reflectance. The rotation effectively changes the sun-viewing geometry, introduces different 3D cloud radiative effects from the training data set, and ensures no overlapping scenes between the training and testing datasets. The final sample size is 185,000 for the training and 46,000 for the testing.

To capture the complexity of the input reflectance scene, we used a CNN that is commonly used to identify spatial patterns in images (Okamura et al., 2017; Zeiler & Fergus, 2014). Our CNN has two layers with convolution, followed by two fully connected layers. A schematic of the CNN architecture can be found in Figure S2 in Supporting Information S1. A Rectified Linear Unit (ReLU; Agarap, 2018) activation function was used at

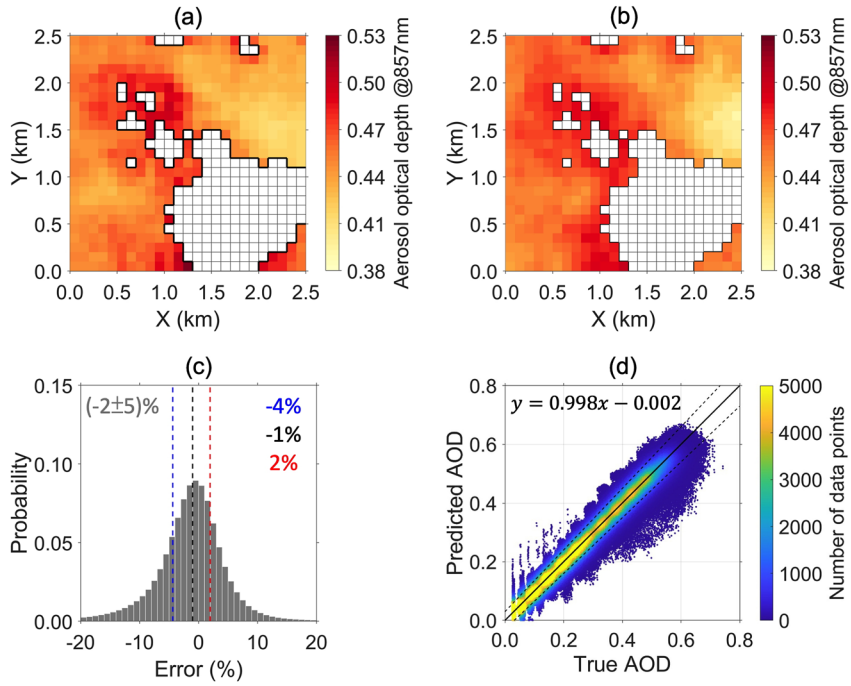

Figure 2. Panel (a) is the true aerosol optical depth (AOD) at 857 nm and (b) is the prediction from the Convolutional Neural Network (CNN) for a testing scene. Using the entire testing data set, (c) is the histogram of the corresponding relative errors (%), and (d) is the density scatter plot of predicted AOD versus the truth for all cloud-free pixels. In (c), the mean error ± the mean absolute deviation (%, using the mean as the center point) are denoted in gray; the 25th, 50th, and 75th percentiles of the errors are represented by blue, black, and red dashed lines and numbers, respectively. In (d), the solid line represents the 1:1 line, while the dotted lines represent ±(0.03 + 5%AOD). The corresponding linear regression equation is listed in the upper-left corner.

each node across all layers. The input (i.e., reflectance) and output (i.e., AOD) values were scaled using the mean normalization method, that is, subtracting the mean and dividing by the full range of the values. Additionally, the training of the CNN was performed by the Adam optimizer (Kingma & Ba, 2015), with the loss function defined as the mean squared error between the true and the predicted AOD. We used early stopping (Prechelt, 1998) to avoid overfitting the training data set.

2.2. Performance Evaluations

Figures 2a and 2b show the ability of our CNN to retrieve AOD at 857 nm in a testing scene. All clear-sky pixels (defined as those with zero liquid water path) agree with the true AOD to be better than 0.03, except for seven pixels. Importantly, the enhanced AODs near clouds are captured by the CNN. Considering the entire testing data set, the mean error is −2% (Figure 2c), and the scatter plot in Figure 2d shows that most data points fall on the 1:1 line. By binning the AOD and analyzing the standard deviation of the prediction errors in each AOD bin, the uncertainty in the CNN retrievals is estimated as 0.01 + 5%AOD. This uncertainty meets the pre-launch expectation for MODIS retrievals, which is 0.03 + 5%AOD (Remer et al., 2020).

3. Explaining the CNN

To understand how the CNN combined input information to predict AOD near clouds, we applied Layer-wise Relevance Propagation analyses (Bach et al., 2015), one of the explainable AI methods. Specifically, we used the method called LRPz (Montavon et al., 2017), which has shown promising results in regression problems (Mamalakis et al., 2022). For a CNN with N layers, let us denote the layer with index l that ranges from 1 (the input layer) to N (the output layer). The relevance, a measure of the contribution of a single pixel to the prediction, is calculated starting from the output layer, and then propagated backward, layer-by-layer, to the input layer, that is,

$$R_i^{(l)} = \sum_{\substack{k\ :\ \text{neurons} \in (l+1)\ and \\ \text{have received input from neuron } i \in (l)}} R_{i\leftarrow k}^{(l,l+1)}, \tag{1}$$

where $R_i^{(l)}$ represents the relevance of neuron i at Layer l to the prediction of interest, and $R_{i\leftarrow k}^{(l,l+1)}$ represents the relevance propagated backward from Layer $l+1$ to Layer l, from neuron k to neuron i. At Layer $l+1$, any neuron that has received input from neuron i should be counted in calculations of $R_i^{(l)}$, which is indicated by the summation in Equation 1.

Now, let us discuss how to calculate $R_{i\leftarrow k}^{(l,l+1)}$ for Equation 1. The information that flows forward from neuron i in Layer l to neuron k in Layer $l+1$ is the input from i, denoted by x_i, times the weight applied to that input at neuron k, denoted by w_{ik}. The relative contribution of neuron i to neuron k compared to all information flowing from Layer l to neuron k is $x_i w_{ik} / \sum_{j\in(l)} x_j w_{jk}$. This relative contribution shows how neuron i and neuron k are connected. To determine the relevance propagated backward from k to i, we multiply this relative contribution by the relevance of neuron k, leading to

$$R_{i\leftarrow k}^{(l,l+1)} = R_k^{(l+1)} \frac{x_i w_{ik}}{\sum_{j\in(l)} x_j w_{jk}}. \tag{2}$$

In short, the relevance can be thought of as the local contribution of in input pixel to the prediction. Since ReLU has been used as the activation function, the attribution from LRPz is equivalent to the product of input pixel value and the gradient of the CNN at the input pixel (Kindermans et al., 2016). A higher absolute value indicates a higher relevance. The sign of the relevance is positive if the local contribution of the neuron has the same sign as the sum of the contribution from all input neurons. In contrast, the negative sign indicates that the signs of the local contribution and the aggregated contribution are opposite.

Figure 3 shows examples of the relevance heat maps for a few output pixels using the testing scene in Figure 2a. For convenience, the input reflectance field for the scene is shown in Figure 3a to relate the pixels of interest to cloud locations. The main findings are noted here:

- Relevance from clouds is evident no matter whether the pixel of interest is near clouds or relatively far from clouds. Considering that the optical influence of clouds can enhance the reflectance of cloud-free pixels as far as 10–15 km (Várnai & Marshak, 2012), the relevance from clouds found in a 4.5 × 4.5 km domain is reasonable.
- For pixels relatively far from clouds (Figures 3d and 3f), we see strong influence from the local surrounding pixels, with a negative sign. Recall that the negative sign indicates that the contribution has an opposite sign from that of the aggregated contribution. Since pixels far from clouds typically have lower AOD and reflectance compared to near-cloud pixels that are strongly influenced by humidity and clouds, the local contribution needs to be of opposite sign to lower the AOD so as to match what has been observed in the training data set. The strong negative local relevance fades away when moving the pixel of interest toward clouds (e.g., Figures 3b and 3c), since the influence from clouds becomes more significant.
- Unlike Figures 3b and 3c, the near-cloud pixel in Figure 3e receives strong negative influence from nearby cloud-free pixels. This is because those nearby pixels are in the shadow cast by clouds, and thus have low input reflectance values. This leads to a lower corresponding relevance compared to other pixels and thus is associated with a negative sign.

Although results from LRPz analyses generally match our understanding of radiative transfer, the relatively small contributions from cloud-free pixels to predictions at near-cloud pixels are surprising. As explained above, the attribution from LRPz can largely depend on the input pixel value. Hence, the generally small contributions from cloud-free pixels can be the consequence of the LRPz technique itself, but the role of near-cloud reflectance in the CNN remains unclear and needs further investigations as discussed next.

We propose two hypotheses to explain how the CNN predicts AOD. The first hypothesis is that the CNN identifies cloud locations, retrieves AOD for pixels far away from clouds, and then interpolates AOD based on the relationships between AOD and distances to nearest clouds. In this hypothesis, the CNN does not need to actively account for 3D cloud radiative effects in observed reflectance values and hence the relevance of pixels in near-cloud

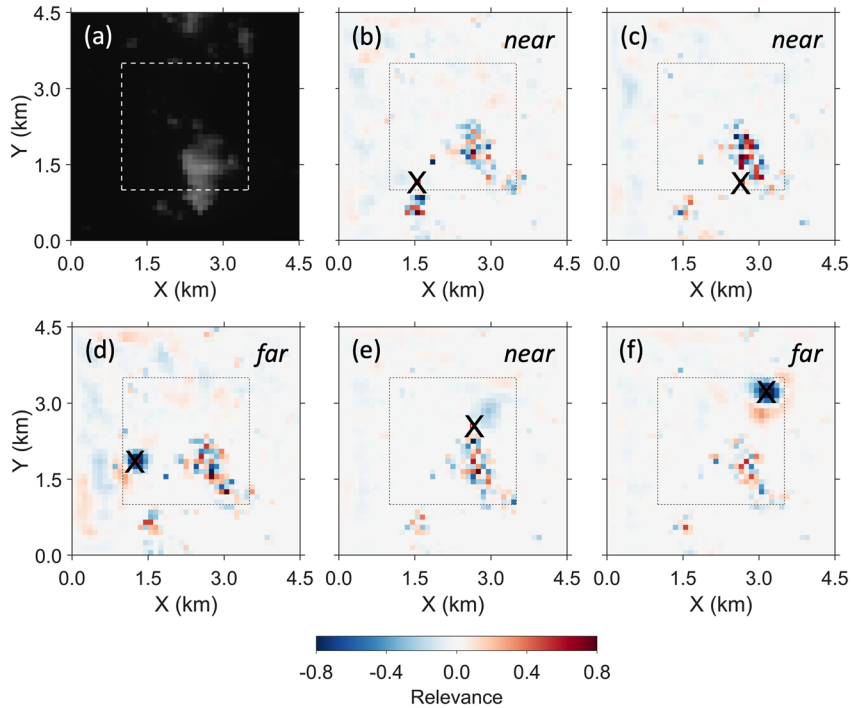

Figure 3. (a) Input reflectance field for the testing scene used in Figure 2a. (b–f) are the corresponding heat maps of relevance based on our Convolutional Neural Network (CNN) for selected pixels marked by black crosses. These pixels are near clouds in (b, c) and e), and farther from clouds in (d and f). Dotted boxes represent the output aerosol optical depth (AOD) domain. The relevance has been normalized by the corresponding maximum absolute value within each map, constraining the range to always fall within ±1. For illustration purposes, we use a color scale of ±0.8.

regions is minimal. The second hypothesis is that the CNN does actively correct the 3D cloud radiative effects in observed reflectances and uses the corrected reflectance field to retrieve AOD. To test which hypothesis is valid, we conducted the following experiment.

For the scene shown in Figure 2a, we used the same input reflectance field, but manipulated reflectance for cloud-free pixels (Figure S3 in Supporting Information S1). A constant reflectance value that corresponds to an AOD of 0.44 (the median of the AOD range in Figure 2a) was applied. If the first hypothesis is correct, that is, the CNN only minds the cloud locations and the background reflectance value, then the retrieved AOD field would still show the enhancement near clouds, even though there is no enhancement in the input reflectance field. However, if the second hypothesis is correct, that is, the CNN actively accounts for 3D cloud radiative effects, then the CNN will tend to make a larger reflectance correction in near-cloud regions than in regions far from clouds. As a result, the trend of AOD versus distance to nearest clouds should disappear or even reverse. Additionally, since the reflectance on the shadow side is supposed to be lower than the illuminated side for a given AOD, the CNN will need to account for the shadowing effect during the prediction process. Hence, if the CNN indeed recognizes the sun-viewing geometry, our manipulated constant reflectance field will lead to a higher AOD on the shadowing side than the illuminated side, given the same reflectance on both sides.

As shown in Figure S3b in Supporting Information S1, the retrievals from the manipulated reflectance input show no clear evidence of AOD enhancement near clouds. Instead, the retrievals on the shadowing side are ∼18% larger than those on the illuminated side. These findings suggest that the second hypothesis is more plausible. It appears that our CNN actively accounts for 3D cloud radiative effects and corrects them in the input reflectance for predicting AOD. In short, it is not surprising that the LRPz technique made a large attribution to cloudy pixels in AOD predictions, since the largest contribution in reflectance enhancement comes from clouds. However, caution should be exercised in interpreting the small attributions in clear-sky regions. Our experiment with a manipulated reflectance field suggests that the reflectances near clouds remain important for the CNN to retrieve AOD.

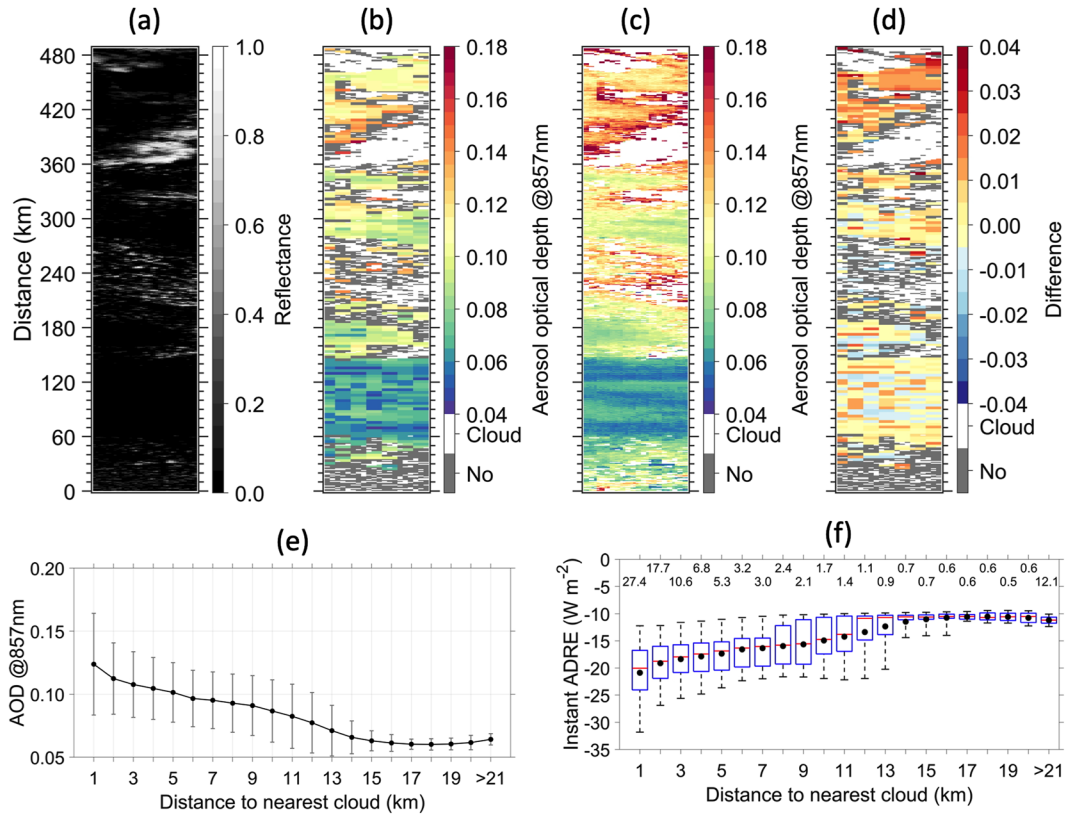

Figure 4. A case observed by MODIS onboard the Aqua Satellite at 17:25 UTC on 22 October 2009, showing (a) reflectance at 857 nm wavelength, (b) aerosol optical depth (AOD) from the operational product, (c) AOD from the Convolutional Neural Network (CNN), and (d) the difference computed by subtracting (b) from (c). White color represents pixels in which operational cloud retrieval is available, but operational AOD retrieval is not. Gray color represents pixels in which both operational cloud and aerosol retrievals are unavailable. Note that (a and c) are plotted at their native resolution of 500 m, while (b and d) are plotted at 3-km resolution. Panel (e) shows the mean retrieved AOD versus distance to clouds; the error bar represents one standard deviation. Panel (f) is same as (e), but for the box plot of the instantaneous mean aerosol direct radiative effect. The bottom and top of each box represent the 25% and 75% percentiles, and the red line inside the box represents the median. The whiskers mark the 5th and 95th percentiles. The black dots represent the means. The fraction (in %) of pixel number with respect to the total cloud-free pixels in each bin is listed in the figure.

4. A Case Study Using MODIS Observations

4.1. Comparison to the MODIS Operational Product

In this section, we apply our CNN approach to an Aqua satellite overpass near Bermuda and compare our predicted AOD to MODIS operational Level 2 Collection 6 aerosol products. The operational product (MYD04_3K) has a horizontal resolution of 3 km with a pre-launch expected uncertainty of 0.03 + 5%AOD over oceans (Levy et al., 2013; Remer et al., 2005, 2013).

We selected a 20-km wide scene over the ocean, away from sun-glint, comprising scattered shallow cumulus (see Figure 4a). The solar zenith angles range between 47° and 51°, and the solar azimuth angle is ~210° clockwise from the north. The viewing zenith angle is within 1° from the nadir. Our comparison will focus on regions far away from clouds (e.g., at distances of 60–120 km in Figure 4a) because that is the region where the operational product performs best.

For this application, we use the same CNN configuration as in Section 2 but trained with a different data set to accommodate the difference in ambient aerosol types, pixel resolution, and sun-geometry. We use MODIS pre-defined aerosol models to minimize the comparison discrepancy. For the selected scene, most MODIS retrievals suggest that the dominant fine mode is wet water-soluble particles, and the dominant coarse mode is sea salt. The fine mode fraction varies from 0.3 to 0.7. For simplicity, we used a fine mode fraction of 0.5. We

then recalculate reflectance for MODIS resolution of 500 m to regenerate our training data set and retrain the CNN. More details can be found in Text S1 in Supporting Information S1.

A total of 49 CNNs covered all observed SZAs in Figure 4a were built and applied to cloud-free pixels, based on cloud masks in the MODIS Level 2 Collection 6 cloud products. The CNN predictions are averaged from 500 m to 3 km to match the resolution of the aerosol operational product. In general, the AOD difference between two retrieval sets is within MODIS retrieval uncertainty except for five pixels (see Figure 4d). Considering that these two methods are based on different radiative transfer calculations (3D vs. 1D), and that some pixels are based on different fine mode fractions, the agreement within the expected MODIS retrieval uncertainty in the far-from-cloud regions suggests that our CNNs work well.

4.2. Radiative Effects of Near-Cloud Aerosols and Their Impacts

Using retrievals from the CNN in Figure 4c, we investigate how AOD and aerosol DRE at TOA vary with distance to nearest clouds. Details for DRE calculations and flux evaluations against observations from the Clouds and the Earth's Radiant Energy System can be found in Text S2 and Figure S4 in Supporting Information S1. As shown in Figure 4e, when approaching clouds from far away (e.g., further than 20 km) to near clouds (e.g., $\lesssim$1 km), AOD at 857 nm increases from 0.06 to 0.12, which is a 100% enhancement. Similarly, the magnitude of the instantaneous DRE at TOA increases by 91%, changing from -11 to -21 W m^{-2}.

The overall radiative effect of near-cloud aerosols in this scene is calculated as:

$$\mathrm{DRE} = \sum_j f_j \cdot \mathrm{DRE}_j \tag{3}$$

where f_j is the fraction of cloud-free pixels that occur at certain distances bin j. Using Equation 3 and results in Figure 4f, we found the final mean DRE is -17 W m^{-2}. If aerosols were not humidified, the DRE across all distance bins would be close to the value far away from clouds, for example, -11 W m^{-2}. Hence, the DRE enhancement due to humidification effects is about 55%, consistent with the (35%–65%) estimate made by Twohy et al. (2009). This consistency is encouraging, because their aerosol properties are based on in-situ measurements and thus not affected by 3D cloud radiative effects. It provides confidence that the CNN can be further extended for global satellite observations over oceans to improve the estimate of aerosol DRE.

5. Summary

We have developed a new CNN for retrieving AOD, with a focus on near-cloud regions in which AOD is enhanced due to humidification effects. Our CNN was trained using large-eddy simulations of marine cumuli under various atmospheric and aerosol environments, and the reflectance at 857 nm wavelength was computed by a 3D radiative transfer model. In the testing data set, the CNN predicts AOD at 100 m resolution with an uncertainty of 0.01 + 5%AOD, and the mean relative bias is -2%. For fields of scattered cumuli observed by Aqua satellite over Bermuda, the predicted AOD generally agrees with the MODIS operational product within the retrieval uncertainty for regions far away from clouds. Importantly, the CNN provides retrievals of near-cloud aerosols. For this case, the AOD enhancement due to humidification near clouds leads to a 55% increase in clear-sky aerosol radiative effect estimates.

We have also applied an explainable AI method and conducted an additional experiment to understand how the CNN uses the reflectance scene for AOD retrievals. Results indicate that the CNN recognizes the sun-viewing geometry, actively removes the 3D cloud radiative effects from the input reflectance field, and then retrieves AOD from the corrected reflectance. Although the current CNN is designed for marine low clouds, the method can be extended to other cloud types and incorporate various sun-viewing geometry, aerosol models, surface conditions, and wavelengths of interest. The extensions of the CNN will greatly help improve aerosol radiative forcing estimates that include humidification effects near clouds.

Data Availability Statement

NASA aerosol products are available via http://dx.doi.org/10.5067/MODIS/MYD04_3K.061 (Levy & Hsu, 2015) and cloud products are available via http://dx.doi.org/10.5067/MODIS/MYD06_L2.061 (Platnick et al., 2015). The LES simulations and the machine learning code can be accessed through the Mountain Scholar data repository (http://dx.doi.org/10.25675/10217/235755).

Acknowledgments
This research is supported by National Aeronautics and Space Administration Projects 80NSSC20K0596 and 80NSSC20K1719. Yang thanks the Cooperative Institute for Research in the Atmosphere for support. Feingold and Yamaguchi are supported by U.S. Department of Energy, Office of Science, Atmospheric System Research Program Interagency Award 89243020SSC000055. Van Leeuwen is supported by the European Research Council under the Causality relations Using Nonlinear Data Assimilation project 694509. The authors thank Christian Kummerow, Steven Miller and Imme Ebert-Uphoff for stimulating discussions on explaining the CNN. The authors also thank Sonia Kreidenweis and Ernie Lewis for discussions on aerosol growth factors. The authors acknowledge high-performance computing support from Cheyenne (https://doi.org/10.5065/D6RX99HX), provided by the Computational and Information Systems Laboratory at National Center for Atmospheric Research and sponsored by the National Science Foundation. This work also utilized resources from the University of Colorado Boulder Research Computing Group, which is supported by the National Science Foundation (awards ACI-1532235 and ACI-1532236), the University of Colorado Boulder, and Colorado State University.

References

Agarap, A. F. (2018). Deep learning using rectified linear units (RELU). *arXiv:1803.08375*. https://doi.org/10.48550/arXiv.1803.08375

Bach, S., Binder, A., Montavon, G., Klauschen, F., Müller, K. R., & Samek, W. (2015). On pixel-wise explanations for non-linear classifier decisions by layer-wise relevance propagation. *PLoS One*, *10*(7), e0130140. https://doi.org/10.1371/journal.pone.0130140

Bellouin, N., Quaas, J., Gryspeerdt, E., Kinne, S., Stier, P., Watson-Parris, D., et al. (2020). Bounding global aerosol radiative forcing of climate change. *Reviews of Geophysics*, *58*(1), e2019RG000660. https://doi.org/10.1029/2019rg000660

Charlson, R. J., Ackerman, A. S., Bender, F. A. M., Anderson, T. L., & Liu, Z. (2007). On the climate forcing consequences of the albedo continuum between cloudy and clear air. *Tellus B: Chemical and Physical Meteorology*, *59*(4), 715–727. https://doi.org/10.3402/tellusb.v59i4.17051

Curtis, D. B., Aycibin, M., Young, M. A., Grassian, V. H., & Kleiber, P. D. (2007). Simultaneous measurement of light-scattering properties and particle size distribution for aerosols: Application to ammonium sulfate and quartz aerosol particles. *Atmospheric Environment*, *41*(22), 4748–4758. https://doi.org/10.1016/j.atmosenv.2007.03.020

Evans, K. F. (1998). The spherical harmonics discrete ordinate method for three-dimensional atmospheric radiative transfer. *Journal of the Atmospheric Sciences*, *55*(3), 429–446. https://doi.org/10.1175/1520-0469(1998)055<0429:tshdom>2.0.co;2

Forster, P., Storelvmo, T., Armour, K., Collins, W., Dufresne, J. L., Frame, D., et al. (2021). The Earth's energy budget, climate feedbacks, and climate sensitivity. In P. Zhai, A. Pirani, S. L. Connors, C. Péan, S. Berger, N. Caud, et al. (Eds.), *Climate change 2021: The physical science basis. Contribution of working group I to the sixth assessment report of the intergovernmental panel on climate change [Masson-Delmotte, V.]*. Cambridge University Press.

Hogan, R. J., & Shonk, J. K. P. (2013). Incorporating the effects of 3D radiative transfer in the presence of clouds into two-stream multilayer radiation schemes. *Journal of the Atmospheric Sciences*, *70*(2), 708–724. https://doi.org/10.1175/jas-d-12-041.1

Hoppel, W. A., Frick, G. M., & Larson, R. E. (1986). Effect of nonprecipitating clouds on the aerosol size distribution in the marine boundary layer. *Geophysical Research Letters*, *13*(2), 125–128. https://doi.org/10.1029/gl013i002p00125

Jiang, H., Feingold, G., & Koren, I. (2009). Effect of aerosol on trade cumulus cloud morphology. *Journal of Geophysical Research*, *114*(D11), D11209. https://doi.org/10.1029/2009JD011750

Kindermans, P. J., Schütt, K., Müller, K. R., & Dähne, S. (2016). Investigating the influence of noise and distractors on the interpretation of neural networks. *arXiv:1611.07270*. https://doi.org/10.48550/arXiv.1611.07270

Kingma, D. P., & Ba, J. (2015). Adam: A method for stochastic optimization. *Proceedings of the 3rd international conference for learning representations*. https://doi.rog/10.48550/arXiv.1412.6980

Koren, I., Remer, L. A., Kaufman, Y. J., Rudich, Y., & Martins, J. V. (2007). On the twilight zone between clouds and aerosols. *Geophysical Research Letters*, *34*(8). https://doi.org/10.1029/2007gl029253

Levy, R., & Hsu, C. (2015). MODIS atmosphere L2 aerosol product. In *NASA MODIS adaptive processing system*. Goddard Space Flight Center. https://doi.org/10.5067/MODIS/MYD04_L2.061

Levy, R. C., Mattoo, S., Munchak, L. A., Remer, L. A., Sayer, A. M., Patadia, F., et al. (2013). The Collection 6 MODIS aerosol products over land and ocean. *Atmospheric Measurement Techniques*, *6*(11), 2989–3034. https://doi.org/10.5194/amt-6-2989-2013

Mamalakis, A., Ebert-Uphoff, I., & Barnes, E. A. (2022). Neural network attribution methods for problems in geoscience: A novel synthetic benchmark dataset. *Environmental Data Science*, *1*, E8. https://doi.org/10.1017/eds.2022.7

Marshak, A., Evans, K. F., Várnai, T., & Wen, G. (2014). Extending 3D near-cloud corrections from shorter to longer wavelengths. *Journal of Quantitative Spectroscopy and Radiative Transfer*, *147*, 79–85. https://doi.org/10.1016/j.jqsrt.2014.05.022

Marshak, A., Wen, G., Coakley, J. A., Jr., Remer, L. A., Loeb, N. G., & Cahalan, R. F. (2008). A simple model for the cloud adjacency effect and the apparent bluing of aerosols near clouds. *Journal of Geophysical Research*, *113*(D14), D14S17. https://doi.org/10.1029/2007JD009196

Montavon, G., Lapuschkin, S., Binder, A., Samek, W., & Müller, K. R. (2017). Explaining nonlinear classification decisions with deep Taylor decomposition. *Pattern recognition*, *65*, 211–222.

Okamura, R., Iwabuchi, H., & Schmidt, K. S. (2017). Feasibility study of multi-pixel retrieval of optical thickness and droplet effective radius of inhomogeneous clouds using deep learning. *Atmospheric Measurement Techniques*, *10*(12), 4747–4759. https://doi.org/10.5194/amt-10-4747-2017

Pincus, R., Hannay, C., & Evans, K. F. (2005). The accuracy of determining three-dimensional radiative transfer effects in cumulus clouds using ground-based profiling instruments. *Journal of the atmospheric sciences*, *62*(7), 2284–2293.

Platnick, S., Ackerman, S. A., King, M. D., Meyer, K., Menzel, W. P., Holz, R. E., et al. (2015). MODIS atmosphere L2 cloud product (06_L2). In *NASA MODIS adaptive processing system*. Goddard Space Flight Center. https://doi.org/10.5067/MODIS/MYD06_L2.061

Prechelt, L. (1998). Early stopping-but when?. In *Neural Networks: Tricks of the trade* (pp. 55–69). Springer.

Rauber, R. M., Stevens, B., Ochs, H. T., III., Knight, C., Albrecht, B. A., Blyth, A. M., et al. (2007). Rain in shallow cumulus over the ocean: The RICO campaign. *Bulletin of the American Meteorological Society*, *88*(12), 1912–1928. https://doi.org/10.1175/bams-88-12-1912

Remer, L. A., Kaufman, Y. J., Tanré, D., Mattoo, S., Chu, D. A., Martins, J. V., et al. (2005). The MODIS aerosol algorithm, products, and validation. *Journal of the Atmospheric Sciences*, *62*(4), 947–973. https://doi.org/10.1175/jas3385.1

Remer, L. A., Levy, R. C., Mattoo, S., Tanré, D., Gupta, P., Shi, Y., et al. (2020). The dark target algorithm for observing the global aerosol system: Past, present, and future. *Remote Sensing*, *12*(18), 2900. https://doi.org/10.3390/rs12182900

Remer, L. A., Mattoo, S., Levy, R. C., & Munchak, L. A. (2013). MODIS 3 km aerosol product: Algorithm and global perspective. *Atmospheric Measurement Techniques*, *6*(7), 1829–1844. https://doi.org/10.5194/amt-6-1829-2013

Schwarz, K., Cermak, J., Fuchs, J., & Andersen, H. (2017). Mapping the twilight zone—What we are missing between clouds and aerosols. *Remote Sensing*, *9*(6), 577. https://doi.org/10.3390/rs9060577

Spencer, R. S., Levy, R. C., Remer, L. A., Mattoo, S., Arnold, G. T., Hlavka, D. L., et al. (2019). Exploring aerosols near clouds with high-spatial-resolution aircraft remote sensing during SEAC4RS. *Journal of Geophysical Research: Atmospheres*, *124*(4), 2148–2173. https://doi.org/10.1029/2018jd028989

Stap, F. A., Hasekamp, O. P., Emde, C., & Röckmann, T. (2016). Multiangle photopolarimetric aerosol retrievals in the vicinity of clouds: Synthetic study based on a large eddy simulation. *Journal of Geophysical Research: Atmospheres*, *121*(21), 12914–12935. https://doi.org/10.1002/2016jd024787

Twohy, C. H., Clement, C. F., Gandrud, B. W., Weinheimer, A. J., Campos, T. L., Baumgardner, D., et al. (2002). Deep convection as a source of new particles in the midlatitude upper troposphere. *Journal of Geophysical Research*, *107*(D21), 4560. https://doi.org/10.1029/2001JD000323

Twohy, C. H., Coakley, J. A., Jr., & Tahnk, W. R. (2009). Effect of changes in relative humidity on aerosol scattering near clouds. *Journal of Geophysical Research*, *114*(D5), D05205. https://doi.org/10.1029/2008JD010991

Várnai, T., & Marshak, A. (2009). MODIS observations of enhanced clear sky reflectance near clouds. *Geophysical Research Letters*, *36*(6), L06807. https://doi.org/10.1029/2008GL037089

Várnai, T., & Marshak, A. (2012). Analysis of co-located MODIS and CALIPSO observations near clouds. *Atmospheric Measurement Techniques*, *5*(2), 389–396. https://doi.org/10.5194/amt-5-389-2012

Várnai, T., & Marshak, A. (2014). Near-cloud aerosol properties from the 1 km resolution MODIS ocean product. *Journal of Geophysical Research: Atmospheres*, *119*(3), 1546–1554. https://doi.org/10.1002/2013JD020633

Wen, G., Marshak, A., & Cahalan, R. F. (2006). Impact of 3-D clouds on clear-sky reflectance and aerosol retrieval in a biomass burning region of Brazil. *IEEE Geoscience and Remote Sensing Letters*, *3*(1), 169–172. https://doi.org/10.1109/lgrs.2005.861386

Wen, G., Marshak, A., Levy, R. C., Remer, L. A., Loeb, N. G., Várnai, T., et al. (2013). Improvement of MODIS aerosol retrievals near clouds. *Journal of Geophysical Research: Atmospheres*, *118*(16), 9168–9181. https://doi.org/10.1002/jgrd.50617

Wen, G., Marshak, A., Várnai, T., & Levy, R. (2016). Testing the two-layer model for correcting near-cloud reflectance enhancement using LES/SHDOM-simulated radiances. *Journal of Geophysical Research: Atmospheres*, *121*(16), 9661–9674. https://doi.org/10.1002/2016jd025021

Yamaguchi, T., Feingold, G., & Kazil, J. (2019). Aerosol-cloud interactions in trade wind cumulus clouds and the role of vertical wind shear. *Journal of Geophysical Research: Atmospheres*, *124*(22), 12244–12261. https://doi.org/10.1029/2019jd031073

Yang, C. K., Chiu, J. C., Marshak, A., Feingold, G., Várnai, T., Wen, G., et al. (2022). Dataset associated with "Near-Cloud Aerosol Retrieval Using Machine Learning Techniques, and Implied Direct Radiative Effects". http://dx.doi.org/10.25675/10217/235755

Zeiler, M. D., & Fergus, R. (2014). Visualizing and understanding convolutional networks. In *European conference on computer vision* (pp. 818–833). Springer.