

Earth and Space Science



RESEARCH ARTICLE

10.1029/2025EA004338

Key Points:

- We used the UNet 3+ to estimate aerosol optical depth in areas with missing satellite data, enhancing their usability in air quality applications
- The UNet 3+ can track the emissions and transport of wildfire smoke, even when clouds partially obscure it
- The UNet 3+ generalizes reasonably well over CONUS, but biases exist over regions with unique geographical properties

Supporting Information:

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

Correspondence to:

J. S. M. Lee,
jeffrey.lee-1@ou.edu

Citation:

Lee, J. S. M., Loria-Salazar, S. M., Holmes, H. A., & Sayer, A. M. (2025). Spatiotemporal gap-filling of NASA deep blue satellite aerosol optical depth over the contiguous United States (CONUS) using the UNet 3+ architecture. *Earth and Space Science*, 12, e2025EA004338. <https://doi.org/10.1029/2025EA004338>

Received 13 MAR 2025

Accepted 28 JUN 2025

Author Contributions:

Conceptualization: Jeffrey S. M. Lee, S. Marcela Loria-Salazar

Data curation: Jeffrey S. M. Lee

Formal analysis: Jeffrey S. M. Lee

Funding acquisition: S. Marcela Loria-Salazar

Investigation: Jeffrey S. M. Lee

Methodology: Jeffrey S. M. Lee, S. Marcela Loria-Salazar, Heather A. Holmes, Andrew M. Sayer

Project administration:

S. Marcela Loria-Salazar

Resources: S. Marcela Loria-Salazar, Heather A. Holmes, Andrew M. Sayer

Software: Jeffrey S. M. Lee

Supervision: S. Marcela Loria-Salazar

© 2025. The Author(s).

This is an open access article under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

Spatiotemporal Gap-Filling of NASA Deep Blue Satellite Aerosol Optical Depth Over the Contiguous United States (CONUS) Using the UNet 3+ Architecture

Jeffrey S. M. Lee¹ , S. Marcela Loria-Salazar¹ , Heather A. Holmes², and Andrew M. Sayer^{3,4} 

¹School of Meteorology, University of Oklahoma, Norman, OK, USA, ²Department of Chemical Engineering, University of Utah, Salt Lake City, UT, USA, ³Goddard Earth Sciences Technology and Research II, University of Maryland Baltimore County, Baltimore, MD, USA, ⁴NASA Goddard Space Flight Center, Greenbelt, MD, USA

Abstract Due to sensor and algorithmic constraints, satellite aerosol optical depth (AOD) retrievals are spatially incomplete and have gaps caused by clouds and bright surfaces. These gaps represent a barrier in characterizing daily aerosol loadings, which is important for air quality applications. In particular, recent studies in aerosol studies have shown satellite AOD to be a useful predictor of particulate matter, but are often limited to monthly or longer temporal resolution because of missing AOD retrievals. In this study, we propose using a UNet 3+ to fill gaps in satellite AOD retrievals. We tested the hypothesis that UNet 3+ trained on deep blue (DB) AOD and supplemental data sets (e.g., Modern-Era Retrospective analysis for Research and Applications, Version 2 reanalysis AOD, meteorological and land-use variables from North American Mesoscale Forecast System, and Hazard Mapping System smoke polygons) will improve the availability of AOD data accurately. We created spatiotemporal data sets of daily, gap-filled DB AOD from 2012 to 2023 over the Continental United States (CONUS) at a $12 \times 12 \text{ km}^2$ resolution. We were able to train the model and perform the gap-filling in ~ 10 hr, resulting in an increase of AOD data availability by 281%. We demonstrated that our approach is feasible over CONUS through quantitative and qualitative evaluations against AERONET and DB AOD. In statistical evaluations, our gap-filled AOD data set attained an RMSE ~ 0.09 and a $r \sim 0.87$ against collocated AERONET retrievals, compared to an RMSE ~ 0.11 and a $r \sim 0.86$ that the original DB AOD retrievals scored against AERONET. We plan to use this data set for future air quality and health investigations.

Plain Language Summary A major contributor to poor air quality is the presence of microscopic particles in the air, known as aerosols. In the U.S., the primary method of monitoring these aerosols is by using instruments to sample the air around ground-level, but high costs lead to insufficient coverage. An alternative solution is using aerosol data retrieved from satellites to derive surface concentrations of particulate matter (PM). These data sets cover most of the Earth at a daily resolution but have gaps in areas covered by clouds and bright surfaces. These gaps prevent the full estimation of surface concentrations of PM, making it difficult for health exposure assessments to be performed. In this study, we explore the use of deep learning to fill these gaps. Our novel workflow utilizes satellite, reanalysis, and human-labeled data sets and a custom data preprocessing and training methodology to provide a best-estimate in inaccessible regions. Evaluation against ground-based instruments suggests that our data product performs better than the original satellite data set as well as the reanalysis data set.

1. Introduction

Atmospheric aerosols are small airborne solid or liquid particles that range in size from nanometers to micrometers (Pöschl, 2005). Aerosols significantly impact the Earth by altering weather and climate through light absorption and scattering, as well as modifying cloud microphysics. These aerosols come in diverse forms (e.g., dust, sea salt, ash, smoke, pollen) and are emitted from both natural and anthropogenic sources (Andreae & Rosenfeld, 2008; Jacobson et al., 2000). Man-made contributions to global aerosol emissions have steadily increased since the Industrial Revolution (Bond et al., 2007), a factor that plays a prominent role in today's changing climate and human health effects (Charlson et al., 1992; K.-H. Kim et al., 2015; H. Lee et al., 2023; Leibensperger et al., 2012; Wilcox et al., 2013).

For aerosol-related health studies, aerosols are known synonymously as particulate matter (PM) and are categorized by aerodynamic diameter ($\text{PM}_{2.5}$ is defined as PM with an aerodynamic diameter of less than 2.5 microns,

Validation: Jeffrey S. M. Lee,
S. Marcela Loria-Salazar, Andrew
M. Sayer

Visualization: Jeffrey S. M. Lee

Writing – original draft: Jeffrey
S. M. Lee

Writing – review & editing: Jeffrey
S. M. Lee, S. Marcela Loria-Salazar,
Heather A. Holmes, Andrew M. Sayer

etc.). A growing body of research finds associations between $PM_{2.5}$ inhalation and negative health effects, for example, with wildfire smoke (Brook et al., 2010; Grant & Runkle, 2022; Y. Kim et al., 2017; Martinelli et al., 2013; Matz et al., 2020; C. E. Reid et al., 2016; Uda et al., 2019), dust aerosols (Bach et al., 1996; Cheng et al., 2008; Chien et al., 2012; Gillette & Hanson, 1989; Herrera-Molina et al., 2024; Kang et al., 2012; H. Lee et al., 2013; Panikkath et al., 2013; Péwé & American Association for the Advancement of Science, 1981), and urban emissions (Amato et al., 2014; Buckeridge et al., 2002; Guttikunda & Kopakka, 2014).

A significant advancement in understanding the role of aerosols in human health and the climate has been the use of satellite aerosol products such as aerosol optical depth (AOD). Satellites retrieve AOD and not $PM_{2.5}$ directly because AOD is a column optical metric most directly related to what the satellite measures (i.e., reflected or emitted solar/thermal radiation), and $PM_{2.5}$ is a surface-level mass concentration. Despite these differences, studies have demonstrated the feasibility of modeling surface $PM_{2.5}$ concentrations using AOD from multiple satellite algorithms (such as deep blue (DB), Dark Target (DT), and Multi-Angle Implementation of Atmospheric Correction (MAIAC)) as a spatial predictor (Chang et al., 2014; G. Chen et al., 2021; Hu et al., 2014). However, they noted a limitation of their approach in the lack of a method to fill the gaps in areas where AOD is not retrieved. Many $PM_{2.5}$ studies cannot generate predictions for areas with missing AOD data without utilizing a secondary chemical transport model, increasing the computational resources needed (van Donkelaar et al., 2016).

This sparsity exists because AOD (dependent on sensor and algorithm) cannot be retrieved accurately from satellite sensors under certain conditions. Radiometers cannot generally retrieve AOD for a column of the atmosphere if a section of it is obscured by cloud cover, so all satellite AOD retrieval algorithms include some form of cloud masking (Hsu et al., 2004; Kaufman et al., 1997). Another condition is if the underlying surface is too bright/reflective, such as with snow cover, salt pans, and deserts (Loria-Salazar et al., 2016, 2021; Lyapustin et al., 2018; Peterson et al., 2014, 2017). In some cases, this is due to inadequate surface reflectance modeling; in others, it is due to limitations of the cloud masking process, which can inaccurately flag some of these pixels as cloudy.

AOD reanalyses, such as Modern-Era Retrospective analysis for Research and Applications, Version 2 (MERRA-2), have been a complementary tool in investigations that require spatially complete AOD data sets. However, air quality exposure studies that employ AOD from satellite retrievals to infer $PM_{2.5}$ ideally need a spatial resolution (finer than $12 \times 12 \text{ km}^2$) higher than what MERRA-2 offers for its AOD product ($0.5^\circ \times 0.625^\circ$) to meet research standards regarding population exposure. This is because localized high AOD values can get lost after averaging over such a large spatial area given the right skewed distribution of AOD.

Hence, when it comes to studies requiring daily AOD data sets, researchers are forced to choose between low-resolution AOD from reanalysis and incomplete AOD from satellites. In this study, we aim to address these shortcomings by creating a data set of AOD at the 550 nm wavelength that is both high-resolution ($12 \times 12 \text{ km}^2$) and spatially complete. We were motivated to explore a deep-learning approach to address this problem for a few reasons. One reason is that deep learning (DL) seems to be a particularly well-suited solution given the complex nature of the problem due to diverse aerosol sources, transport and evolution, missing data, and deposition. This is because DL models are characterized by extremely large numbers of parameters (often millions to billions), which, for the purposes of AOD gap-filling, are beneficial for capturing the highly nonlinear spatiotemporal relationships between the many possible input features with AOD. Another reason is that DL has been demonstrated to be adept at handling tasks with an innate missing value component. Relevant examples include image inpainting and uncropping, which involve generating predictions for missing pixels in the interior and exterior of an image, respectively (Saharia et al., 2022). In particular, a class of DL architecture known as UNets has shown proficiency for these tasks. These UNets are convolutional neural networks and are highly suited to processing gridded data with features at multiple scales due to their large number of interconnected encoders and decoders. As such, we have chosen to use the UNet 3+, an updated version of the original UNet, for our study (Huang et al., 2020).

Numerous methods have been used to perform AOD gap-filling, including Bayesian methods (M. K. Singh et al., 2017; Tang et al., 2016), regression (sometimes combined with additional techniques such as Bayesian methods or Kriging) (Lv et al., 2016; Wei et al., 2019; Xiao et al., 2018), random forest (J. Chen et al., 2021), tensor completion (Bai et al., 2022), and many more. However, our study brings value in two major aspects. One is that increasing wildfire activity will place an extra burden on air quality monitoring efforts in the U.S., however, the majority of recent AOD gap-filling studies are focused on Asia where air pollution sources, terrain, and land-

cover types may impact the gap-filling model results (Burke et al., 2021). Second, DL is still an underexplored approach for AOD gap-filling, despite its potential to improve AOD gap-filling approaches, both in terms of viability and model performance.

This study presents our approach to gap-filling daily satellite AOD over CONUS from 2012 to 2023 by using a DL architecture called UNet 3+. Section 2 describes the data sets used to train the UNet 3+. Section 3 presents the rationale behind the choice of architecture, data pre-processing, model training, and model evaluation. Section 4 includes the gap-filling results and discusses the gap-filled DB AOD performance. Finally, Section 5 summarizes the findings of our research.

2. Data Sets

This section explains the instrumentation, models, and data sets used to train the UNet 3+ AOD gap-filling model. We used several different satellite instruments and model reanalyses to train the UNet 3+ and ground-based instruments to validate it. Table S1 in Supporting Information S1 summarizes all the data sets used. Further details regarding each instrument/reanalysis are provided in the following subsections.

2.1. Aerosol Optical Depth (AOD) and Ångström Exponent (AE)

2.1.1. Aerosol RObotic NETwork (AERONET)

NASA operates a global network of Cimel-CE 318 sun photometers called the AErosol RObotic NETwork (AERONET). The Cimel CE-318 is a photometer that can measure direct (direct Sun scan) and sky (almucantar, principal plane, and hybrid scans) irradiance using the sun as the light source under non-cloudy conditions (Dubovik et al., 2000, 2006; Holben et al., 1998). Each instrument takes direct Sun measurements every ~3–15 min, providing spectral AOD, as well as sky scans every hour to retrieve additional aerosol properties (Sinyuk et al., 2020). The AOD retrievals from AERONET are a more direct measurement of AOD compared to the area-averaged retrievals from satellites (Holben et al., 1998), but do not reveal much about the fine-scale spatial distribution of AOD since they are point measurements (Gupta et al., 2006). AERONET has grown to be a very expansive network, with 462 stations included in our study domain, and is a valuable evaluation and calibration data set. As such, we used AERONET AOD data to evaluate the UNet 3+ but not as input. We used AOD₆₇₅, AE_{440–675}, and AE_{440–870} from AERONET Version 3 Level 2 (quality assured) between 10:00 a.m. and 02:00 p.m. local time (Bhattacharjee et al., 2023; Giles et al., 2019).

The Ångström exponent (AE) describes AOD's dependence on wavelength, which is a qualitative indicator of particle size (Ångström, 1929). Since AERONET does not offer AOD retrievals at 550 nm, we used AERONET AE_{440–675} to interpolate AOD to 550 nm, following Eck et al. (2003). In addition, we categorized smoke aerosols using AE_{440–870} for evaluation purposes. AE_{440–870} ≥ 1.7 implies the optical dominance of smoke, which we use to evaluate model performance during wildfire events. Note that wildfire smoke from wet fuels (smoldering smoke) can have an AE lower than 1.7 as the smoldering phase generates larger particles compared to the flaming phase (J. S. Reid et al., 2005). Studies have noted smoke events with AE values as low as 1.5 (Wu et al., 2009). However, for our evaluation, we are not taking these cases into account because we are not using any composition measurement to prove that these aerosols come from smoldering versus other sources that lie in the same AE range, such as urban emissions (Russell et al., 2010; R. Singh et al., 2023).

2.1.2. Deep Blue (DB)

We input DB AOD retrievals from the twin Moderate Resolution Imaging Spectroradiometer (MODIS) instruments aboard Terra (approx. 10:30 a.m. local overpass time) and Aqua (approx. 01:30 p.m. local overpass time) satellites, as well as the Visible Infrared Imaging Radiometer Suite (VIIRS) instrument from the Suomi National Polar-orbiting Partnership (S-NPP) satellite (approx. 01:30 p.m. local overpass time). We use the current Collection 6.1 (C6.1) for Terra and Aqua, which offers many improvements over the previous Collection 6 (C6) including better radiometric calibration, heavy smoke detection, artifact reduction over heterogeneous terrain, improved surface modeling for elevated terrain, bug fixes, and updated regional/seasonal aerosol optical models (Hsu et al., 2019; Sayer et al., 2019). The DB AOD at 550 nm from C6.1, which is at 10 × 10 km² resolution at the sub-satellite point.

For VIIRS we used the Version 1.1 data set over the more recently released Version 2 because Version 1.1 DB AOD is algorithmically identical to Terra and Aqua C6.1 (Hsu et al., 2019). This allows the three sensors to be more reasonably simply averaged into a consistent record, and we use this average as a training target. The DB AOD data from Version 1.1 is given at a $6 \times 6 \text{ km}^2$ resolution at the sub-satellite point. In addition to the AOD retrievals, the UNet 3+ was given three supplemental DB variables collected by VIIRS. The first is aerosol type, a categorical variable with nine levels: no retrieval, dust, smoke, high altitude smoke, pyro cumulonimbus clouds, non-smoke fine mode, mixed (land and ocean), background (land and ocean maritime), and fine dominated. The second is the algorithm flag over land, a categorical variable with four levels: no retrieval, arid DB, vegetated, and mixed. The last is the algorithm flag over the ocean, a categorical variable with four levels: no retrieval, full retrieval, turbid/shallow, and mixed (Hsu et al., 2019; Sayer et al., 2018). These additional variables were included to provide the UNet 3+ with further information to track aerosol transport and evolution. We chose to use DB for our study for two reasons. (a) The DT algorithm (Remer et al., 2005) is only intended for use over dark surfaces such as vegetated land areas and the oceans. This is a complex challenge for areas such as the western U. S., which has bright surface regions (e.g., deserts, snowpack, and salt pans) (Loría-Salazar et al., 2016). (b) MAIAC Collection 6, which was the latest collection at the time of our data selection process, was not chosen as it was found to underestimate high AOD loadings such as smoke in comparison to ground-truth AERONET AOD (Loría-Salazar et al., 2021).

2.1.3. Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2)

MERRA-2 is a global reanalysis developed by NASA's Global Modeling and Assimilation Office. It offers many improvements over the previous version, which, most notably for this project, includes aerosol data assimilation and analysis (Gelaro et al., 2017; Randles et al., 2017). Most of MERRA-2's aerosol assimilation data comes from a neural-network AOD retrieval based on MODIS Terra and Aqua reflectances, with additional AOD inputs from the Advanced Very High Resolution Radiometer, Multiangle Imaging SpectroRadiometer, and AERONET. These AOD retrievals are then assimilated into MERRA-2 using the Goddard Aerosol Assimilation System, which compares the assimilation AOD to MERRA-2 simulated aerosols every 3 hr (Buchard et al., 2015; Randles et al., 2017). From the MERRA-2 data set, an hourly time-averaged single-level assimilation aerosol diagnostics data set, we used the reanalysis AOD at 550 nm product that is spatiotemporally complete and offered at the native resolution of $0.5^\circ \times 0.625^\circ$.

2.2. Weather and Land-Use Variables

We incorporated meteorological reanalysis products from the North American Mesoscale Forecast System (NAM) 12 km (Janjic, 2003; Janjic et al., 2005) to provide the UNet 3+ the building blocks for deciphering the nature of aerosol transport and evolution. NAM is an operational reanalysis model run by the National Oceanic and Atmospheric Administration National Centers for Environmental Prediction. The NAM data sets were selected because (a) a weather model offers spatially complete data on a regularly spaced grid, unlike data collected from radiosondes and weather stations and (b) a mesoscale model is gridded at roughly the same horizontal resolution as satellite products. From NAM, we used surface pressure (P_0), mean sea level pressure, surface temperature (T_0), 2-m temperature (2-T), 2-m dewpoint temperature (2-DewT), 2-m relative humidity (2-RH), 10-m wind speeds (10-WS), planetary boundary layer height, orography, and vegetation.

2.3. Fire Radiative Power (FRP)

Fire Radiative Power (FRP) retrieved from MODIS Terra and Aqua (Giglio et al., 2016) was used to train the UNet 3+. FRP is given in megawatts (MW) and is often used to characterize fire size, energy, and intensity (Archibald et al., 2013; Ichoku & Kaufman, 2005; Kaufman et al., 1998; Roy & Kumar, 2017; Wooster & Zhang, 2004; Wooster et al., 2003). It is retrieved using the brightness temperatures derived from the MODIS 4 and 11 μm channels (Giglio et al., 2003, 2016; Kaufman et al., 1998). By including FRP, the UNet 3+ can be trained with a notion of fire source regions, typically a small subset of the total area occluded by wildfire smoke. This information can help inform the UNet 3+ of fire events and the resulting AOD in the surrounding regions.

2.4. Hazard Mapping System (HMS)

We further included NOAA's Hazard Mapping System (HMS) smoke polygons product to facilitate the decision-making of the UNet 3+ for smoke recognition. HMS is a framework that takes near-real-time polar (S-NPP, NOAA-20, Terra) and geostationary (GOES-16 and GOES-17) satellite images, passes them through an automated fire and smoke plume detection algorithm, and is quality-controlled by expert image analysts (Brey et al., 2018; McNamara et al., 2004). HMS runs year-round and produces daily, spatial qualitative data sets of smoke plumes over CONUS. These plumes are given as shapefiles and are categorized as light, medium, and heavy depending on the apparent opacity of the smoke in the satellite imagery.

3. Methods

The first step toward building an AOD gap-filling model is to select the model architecture. Of all the classes of modern ML models, UNets (Ronneberger et al., 2015) are well-suited for AOD gap-filling as they have the desirable properties of inputting and outputting spatial images, identifying features at various resolutions, and generally being fast to train. While initially developed for image segmentation UNets can easily be repurposed to output continuous values (i.e., AOD values) instead. Moreover, different variants of UNet have achieved successful results in different applications, from medical image segmentation (Huang et al., 2020; Ronneberger et al., 2015; Zhou et al., 2018), image enhancements such as colourization, inpainting, uncropping, or restoration (Saharia et al., 2022), and meteorology (Appel, 2024; Justin et al., 2023).

Several UNet variants have been developed, such as UNet ++ (Zhou et al., 2018), UNet 3+ (Huang et al., 2020), and TransUNet (J. Chen et al., 2021). We chose UNet 3+ because it demonstrated state-of-the-art performance in its original medical image processing field, overperforming UNet and UNet++ while containing fewer parameters (39.39 and 47.18 M compared to 26.97 M, respectively). Additionally, we chose UNet 3+ over attention-based architectures like vision transformer (ViT) (Dosovitskiy et al., 2021) and UNet variants with attention components such as TransUNet because empirical results suggest that attention-based models require larger training data sets and training time to converge compared to fully convolutional networks like UNet 3+. On the other hand, UNet 3+ has reached state-of-the-art performance on extremely small data sets, for instance, on the minimal liver segmentation data set (<1,000 samples) from the ISBI LiTS 2017 Challenge (Huang et al., 2020). Since the training data set we constructed was much closer in size to this compared to the 300 million sample data set (JFT-300M) used to train the ViT to state-of-the-art performance, UNet 3+ was a clear choice (Dosovitskiy et al., 2021). However, networks with attention components may be accessible once we curate larger data sets. We built our UNet 3+ implementation using the PyTorch DL framework for Python, following the architecture described in Huang et al. (2020) but without the full-scale deep supervision or classification-guided module elements.

Before training the UNet 3+, we regridded all data sources onto a unified Albers Equal Area projection with standard parallels set to 29.5°N and 45.5°N, which minimizes scale distortion over CONUS (Kennedy & Kopp, 2001), and contains 360 rows and 515 columns of 12×12 km² grid cells. This resolution was chosen because preliminary testing suggested that the receptive field of the filters in a UNet 3+ model with the original number of encoder/decoder blocks are too small to generate contextually meaningful predictions. Future investigations will explore increasing resolution with deeper models. Figure 1 shows the spatial domain and projection of our output. Data sets with a temporal resolution finer than 24 hr were sliced/averaged to match the satellite data set overpass times. For instance, we took only the 18:00 UTC NAM data and averaged MERRA-2 data between 14:30 and 22:30 UTC in order to better align with the satellites' local overpass times in the U.S. to obtain daily average values. Additional details regarding this process can be found in Section a in Supporting Information S1.

Before training, AOD values are (natural) log-transformed, which widens the distribution in the low end where most of the samples are concentrated, giving the UNet 3+ a chance to learn from a more dynamic range of outputs. Additionally, all continuous variables are standardized from zero to one, and all categorical variables are turned into one-hot vectors. Missing values are converted to zeros. Zero was chosen as it is not too far removed from valid values post-normalization yet distinct from most other values in the data set. Another fill value like -1 would likely yield very similar results. Additional details regarding data pre-processing can be found in Section b in Supporting Information S1.

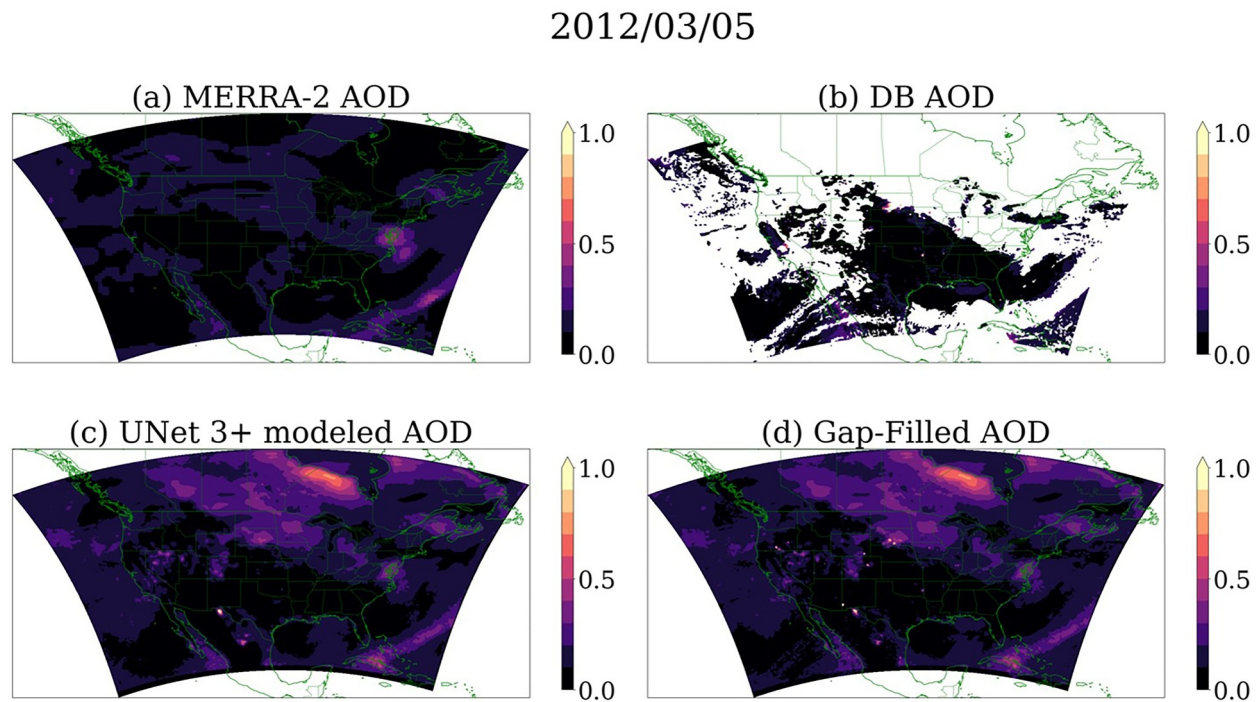


Figure 1. (a) Modern-Era Retrospective analysis for Research and Applications, Version 2 aerosol optical depth (AOD), (b) deep blue (DB) retrievals, (c) UNet 3+ AOD predictions, and (d) gap-filled AOD for 5 March 2012. Gap-filling performed by the gap-filler model trained on only (weighted) satellite DB AOD retrievals.

We also performed a simple bias correction on MERRA-2 AOD before using it in the UNet 3+. Visual inspection, initial statistical testing, and preliminary training suggested that MERRA-2 systematically underestimated high AOD events (>3). The additional MERRA-2 AOD bias correction was performed using a quadratic fit on the log-transformed MERRA-2 AOD data against log-transformed AERONET AOD data. We modeled this as

$$\ln(\text{AOD}_{\text{MERRA-2}} + \beta) = \alpha_1 \ln(\text{AOD}_{\text{AER}} + \beta)^2 + \alpha_2 \ln(\text{AOD}_{\text{AER}} + \beta) + \alpha_3 \quad (1)$$

and found empirically to be $\alpha_1 = 0.0337$, $\alpha_2 = 0.8643$, and $\alpha_3 = -0.3369$. The equation was then solved for AERONET AOD as a function of MERRA-2 AOD, which was then used to bias correct our entire MERRA-2 AOD data set. The raw MERRA-2 AOD values are only used in figures, where any subplot denoted with “MERRA-2” refers to raw instead of bias-corrected values.

Preliminary testing showed that model performance is closely tied to the volume of training data used. Initial tests with training the model on 5 years of data (2013–2017) resulted in a model that struggled to capture high AOD events. We initially attempted to combat this by adding a weight factor in the loss function to penalize mistakes over higher AOD pixels more heavily. However, this led to the generation of high AOD artifacts during the prediction. An example of this is shown in Figure 1, where the model estimated unrealistically high values of AOD in Canada (Figure 1c). These artifacts make it into the final gap-filled AOD product (Figure 1d) because of the lack of DB retrievals over those areas (Figure 1b). Furthermore, MERRA-2 does not suggest such elevated AOD values in these areas (Figure 1a).

Comparing this to the AOD data availability map for winter shown in Figure 5, it seems plausible that the lack of data in these regions either directly or indirectly leads to the generation of these artifacts. The model likely learns to correlate areas of missing aerosol data with high AOD. This behavior is probably learned through cases where extremely optically thick pyrocumulonimbus clouds are partially flagged for removal. Reinforcing this pattern in the model could inadvertently lead to the mistaken association between locations that did not receive an AOD retrieval due to cloud cover or snow reflectivity and those that did not receive an AOD retrieval due to dense smoke plumes.

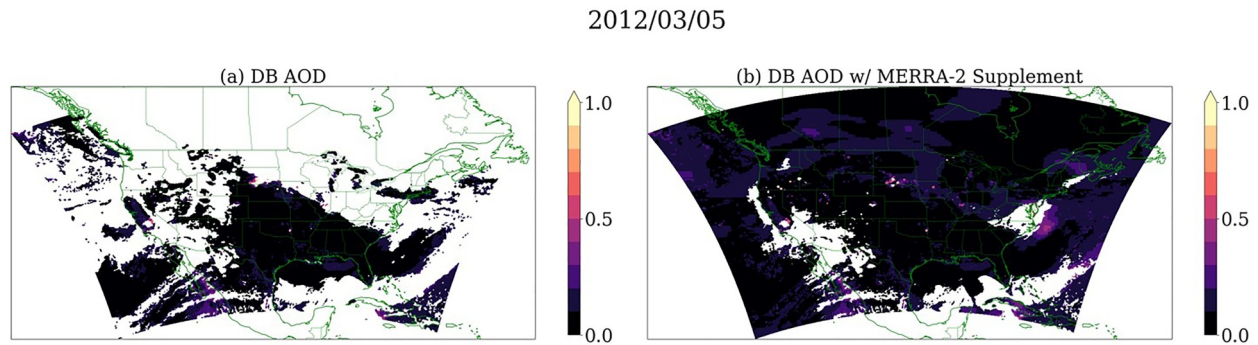


Figure 2. (a) Deep blue aerosol optical depth (AOD) retrievals along with (b) supplemental AOD from Modern-Era Retrospective analysis for Research and Applications, Version 2 for 5 March 2012.

Our approach to mitigate this issue was two-pronged. First, we expanded our study period to 2012–2023. Second, we created a composite of both DB AOD retrievals and MERRA-2 AOD estimates to be used as the training target instead of just the DB AOD retrievals (Figure 2). The first layer of this composite is comprised of the DB AOD retrievals (Figure 2a), as these are preferred as training targets over MERRA-2 AOD whenever available. The second layer is comprised of MERRA-2 AOD, but only for the grid cells where there is no DB AOD retrieval available, and monthly AOD data availability is less than 40% (Figure 2b). The monthly AOD data availability is calculated by taking all DB AOD retrievals over the 2012–2023 study period, grouping them by grid cell and month, and calculating as a percentage of how many days a particular grid cell saw an AOD retrieval within a given month. The threshold of 40% was chosen loosely based on results from the preliminary study, which suggested that areas that received data at least 50%–60% of the time would yield few artifacts. The threshold of 40% covers enough area to prevent the model from generating high AOD artifacts while not completely saturating the training data with MERRA-2 AOD labels. However, we do acknowledge that using MERRA-2 AOD as a training target is not ideal, as the low spatial resolution can obfuscate localized spikes in AOD (e.g., smoke plumes, dust storms), impacting the ability to accurately estimate exposure assessments. As such, a focus for future investigations will be aggregating enough high-resolution AOD data so that MERRA-2 AOD does not need to be used as a surrogate training target.

Even with MERRA-2 AOD supplementing the target, there are still missing pixels. This is a challenge for supervised learning because training labels must be defined to tune the model weights. Our proposed solution for this challenge is simply to train using the available pixels (i.e., just the pixels with defined AOD values), in the backpropagation while ignoring the undefined ones. A simple illustration of this concept on a multilayer perceptron can be seen in Figure 3. This is accomplished by setting the target weight matrix to zero for pixels with

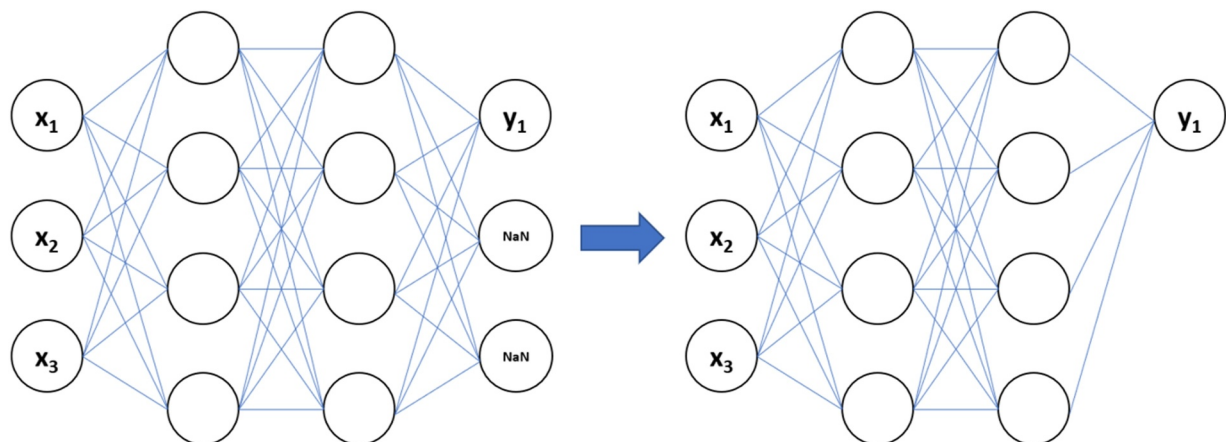


Figure 3. The effect of selective backpropagation on a simple multilayer perceptron model.

Table 1
Input Variables, Native Resolutions, and Lags Provided to the UNet 3+

Input variable	Resolution	Lag (days)
NASA MODIS DB best estimate AOD	$10 \times 10 \text{ km}^2$	2, 1
VIIRS DB best estimate AOD	$6 \times 6 \text{ km}^2$	2, 1
VIIRS DB aerosol type, land, and ocean algorithm flags	$6 \times 6 \text{ km}^2$	2, 1
NAM 12 km reanalysis—P, T, U, V, orography, vegetation, PBLH	$12 \times 12 \text{ km}^2$	2, 1, 0
MERRA-2 reanalysis AOD	$0.5^\circ \times 0.625^\circ$	2, 1, 0
HMS smoke product	$\sim 1 \text{ km}$	2, 1, 0
MODIS FRP	$\sim 1 \text{ km}$	2, 1, 0

undefined AOD, and effectively backpropagates the loss through only the defined pixels because the undefined pixels do not contribute any loss to the loss function.

The UNet 3+ was trained on inputs consisting of 2 days of AOD values (Lags 2, 1) along with meteorology and land-use variables, and targets consisting of the third day's AOD values (Lag 0). Input variables were selected based on their common usage in aerosol models, although we did not perform any variable importance or pruning procedures for this study. All input variables used, and their respective lags, are listed in Table 1. Conceptually, the AOD, aerosol, and smoke/fire data provided at multiple lags can provide the UNet 3+ with the context needed to formulate predictions for the target day (Lag 0), much like the initial conditions for a numerical first principles model. The meteorological variables for all days (Lags 2, 1, 0), were included to enable the UNet 3+ to learn the dynamics of the various aerosol data over time.

The loss function is a composite consisting of a linear combination of the weighted sum of squared errors, which we will refer to as the pixel loss, and the multi-scale structural similarity index (MS-SSIM), which we convert into a patch loss. The former optimizes the accuracy of each pixel, while the latter maintains coherence in the overall structure of the constructed AOD map (Huang et al., 2020). The pixel loss is calculated as

$$l_{\text{pixel}} = \sum_{p \in P} (\hat{y}_p - y_p)^2 \quad (2)$$

where P is the set of all AOD pixels, $\hat{y}_p$ is the UNet 3+ prediction at pixel p , and y_p is the target DB or MERRA-2 AOD value at pixel p . MS-SSIM considers various metrics roughly corresponding to the luminance, contrast, and covariance of predicted and actual values within several differently sized sliding windows (Wang et al., 2003). In other words, MS-SSIM is a holistic metric of how coherent an image is compared to the original. It was developed initially for image quality assessment but was converted into a loss term in the original UNet 3+ as well (Huang et al., 2020). Since MS-SSIM ranges from 0 (worst score) to 1 (perfect score) and loss functions need to increase as performance worsens, MS-SSIM is converted into patch loss simply by subtracting it from one. These two losses are linearly combined, that is

$$l = \alpha l_{\text{pixel}} + \beta l_{\text{patch}} \quad (3)$$

to produce the final loss function that the UNet 3+ is trained on.

We use the Adam optimizer to tune the model weights during training (Kingma & Ba, 2017). The loss function parameters are $\alpha = 1e-5$, $\beta = 1e-2$. The learning rate is $1e-5$. The weight decay is $1e-5$. These hyperparameter settings were tuned by hand, and we note that future investigations may likely achieve better performance through a more rigorous selection process. The Adam parameters are left as the defaults recommended by Kingma and Ba (2017).

After generating AOD predictions from UNet 3+, the final gap-filled output is constructed by taking the average of DB AOD retrievals and UNet 3+ predictions. We find that averaging in this way provides slightly better validation metrics against AERONET and provides smoother transitions between areas with and without DB retrievals, compared to using UNet 3+ predictions only where DB AOD is not retrieved. For pixels with no valid

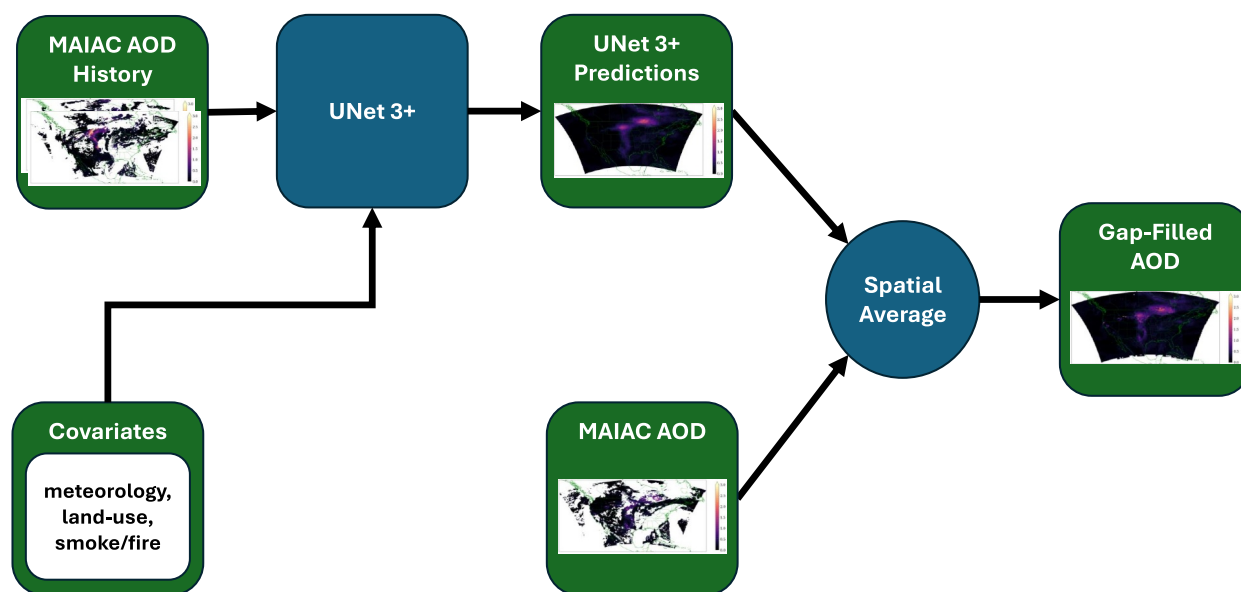


Figure 4. An overview of the gap-filling process introduced in this paper.

AOD retrieval, the gap-filled output is just the UNet 3+ prediction. In the following results, UNet 3+ AOD refers to the “raw” UNet 3+ AOD predictions, while gap-filled AOD refers to the DB AOD and UNet 3+ prediction average. An overview of this process can be found in Figure 4.

To train and evaluate the UNet 3+ performance, we employed a cross-validation scheme with 12 folds (1-fold per year) split into 10 training folds and 2 validation folds. We trained a total of 12 models (1 per rotation) and provided both training and validation statistics in the following section. Details on the statistics used can be found in Section c in Supporting Information S1.

4. Results and Discussion

4.1. Quantitative Results

4.1.1. Increase in AOD Satellite Data

Figure 5 shows the percentage of data available from DB AOD retrievals from MODIS and VIIRS from 2012 to 2023 in the CONTinental United States (CONUS) categorized by seasons. The low AOD retrieval rates signify clouds and/or snow impacting multiple northern U.S. regions during the winter (Figure 5a) and spring (Figure 5b), bright surfaces in the western U.S. (e.g., salt pans) impeding retrievals of DB AOD during the summer (Figure 5c) and fall (Figure 5d), and persistent cloud cover in the southeastern and eastern U.S. during the summer (Figure 5c).

The gap-filling of DB AOD from 2012 to 2023 led to a significant data increase. For this period, we collected a total of 268,296,005 DB AOD data points post-regridding. After gap-filling, the number of AOD data points increased to 752,740,704, or a ~281% increase in data availability. The number of points (N) also reflects this increase in data in the following evaluation against AERONET.

4.1.2. Evaluation Against Ground-Truth AERONET

The performance metrics of MERRA-2 AOD, satellite DB AOD, UNet 3+ AOD, and gap-filled AOD against AERONET AOD can be found in Table 2. The statistics for MERRA-2 and DB AOD are compiled over the entire 2012–2023 period, whereas the statistics for UNet 3+ AOD predictions and gap-filled AOD are separated into training and validation statistics averaged over all cross-validation folds.

UNet 3+ AOD achieves a reasonable mean validation RMSE at 0.10, compared to 0.09 for MERRA-2 and 0.11 for DB. The gap-filled AOD improves the mean validation RMSE to 0.09. This suggests that the gap-filled AOD

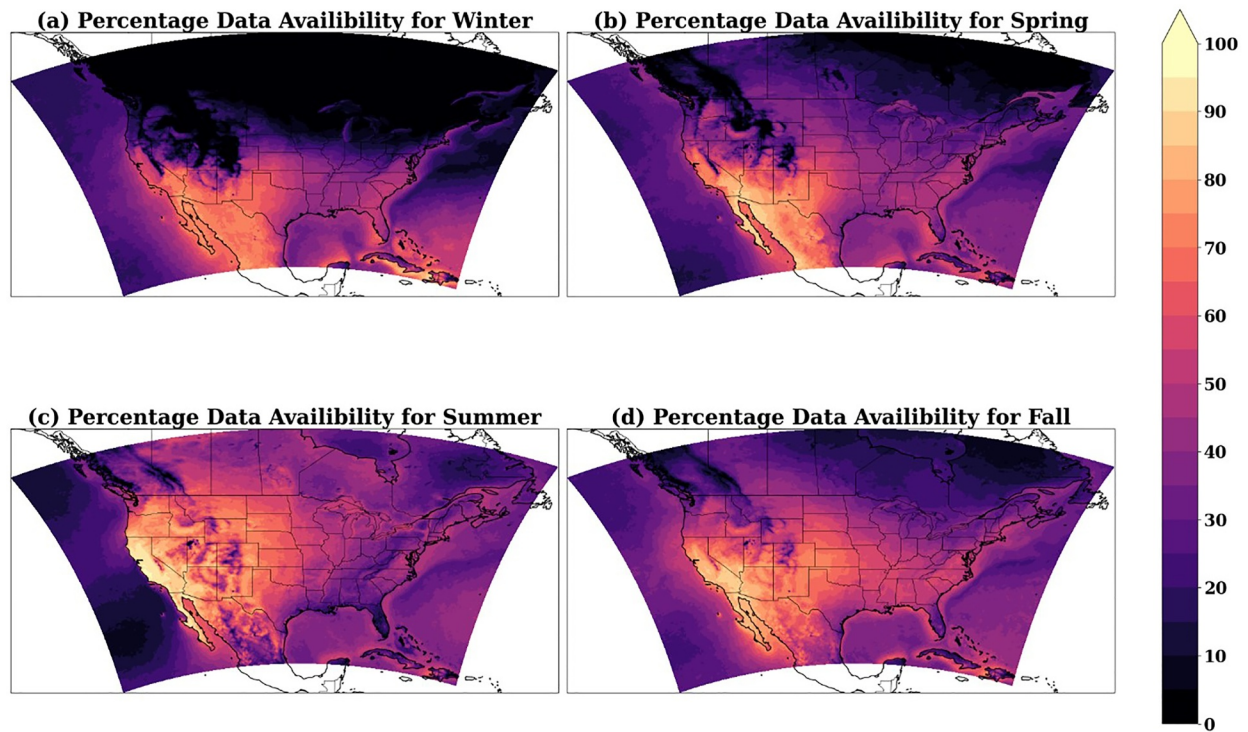


Figure 5. The percentage of days with a successful deep blue Best Estimate aerosol optical depth retrieval from 2012 to 2023 categorized by season.

	RMSE	r	NMB	N	Median	Mean	Std
MERRA-2	0.09	0.81	17.5%	165,443	0.09	0.12	0.12
Smoke	0.14	0.82	9.4%	25,983	0.13	0.19	0.20
Non-smoke	0.08	0.79	20.3%	139,460	0.09	0.11	0.10
DB	0.11	0.86	21.2%	130,580	0.08	0.13	0.20
Smoke	0.17	0.89	21.3%	20,426	0.14	0.22	0.34
Non-smoke	0.09	0.82	21.2%	110,154	0.08	0.11	0.15
UNet 3+ (train)	0.11	0.76	9.2%	137,885	0.07	0.11	0.16
Smoke	0.19	0.74	11.8%	21,659	0.12	0.19	0.26
Non-smoke	0.09	0.76	9.1%	116,225	0.07	0.10	0.13
UNet 3+ (val)	0.10	0.76	8.2%	27,577	0.07	0.11	0.15
Smoke	0.16	0.74	10.0%	4,332	0.11	0.18	0.22
Non-smoke	0.09	0.75	7.5%	23,245	0.07	0.10	0.12
Gap-filled (train)	0.09	0.86	16.6%	137,885	0.08	0.12	0.18
Smoke	0.15	0.88	19.6%	21,659	0.13	0.20	0.30
Non-smoke	0.08	0.83	15.6%	116,225	0.08	0.11	0.14
Gap-filled (val)	0.09	0.84	16.3%	27,577	0.08	0.12	0.17
Smoke	0.13	0.87	19.3%	4,332	0.13	0.19	0.26
Non-smoke	0.08	0.81	15.3%	23,245	0.08	0.11	0.13

is on par with MERRA-2 and DB AOD regarding accuracy. The best-performing cross-validation fold generates gap-filled AOD that achieves a validation RMSE of 0.067, outperforming both MERRA-2 and DB. When comparing training and validation RMSE for both UNet 3+ predictions and gap-filled AOD, consistent values suggest that our model generalizes well even to data it has not been exposed to before. After partitioning our data into smoke ($AE_{440-800} \geq 1.7$) and non-smoke ($AE_{440-800} < 1.7$) events based on AERONET AE values, we can see that smoke events are still significantly more difficult to characterize compared to non-smoke events. Namely, RMSE for smoke events for MERRA-2 (0.14), DB (0.17), UNet 3+ validation (0.16), and gap-filled validation AOD (0.13) are all higher than their counterparts for non-smoke events: MERRA-2 (0.08), DB (0.09), UNet 3+ validation (0.09), and gap-filled validation AOD (0.08).

Regarding correlation values, UNet 3+ scores slightly lower (0.76) than MERRA-2 (0.81) and DB (0.86). The best-performing cross-validation fold for UNet 3+ has a correlation value of $r \sim 0.83$, putting it on par with MERRA-2 performance. The gap-filled AOD further improves on UNet 3+ predictions, with a correlation of $r \sim 0.84$ on average and a maximum of $r \sim 0.89$ in validation.

UNet 3+ can generate AOD predictions with lower NMB ($\sim 8.2\%$) compared to MERRA-2 ($\sim 17.5\%$) and DB ($\sim 21.2\%$). However, upon averaging with DB to generate the gap-filled AOD, NMB unfortunately increases to $\sim 16.3\%$. This is somewhat to be expected, as DB has the highest NMB against AERONET. This is still better than the MERRA-2 NMB, although much of the bias from MERRA-2 comes from non-smoke events ($\sim 20.3\%$) as opposed to smoke events ($\sim 9.4\%$). This is the opposite for the gap-filled AOD, which has increased bias during smoke events ($\sim 19.3\%$) versus non-smoke events ($\sim 15.3\%$).

The corresponding scatter density plots are shown in Figure 6. MERRA-2 (Figure 6a) underestimates most AOD values larger than one ($AOD > 1$) but is reasonably accurate for lower values of AOD ($AOD < 1$) with $RMSE \sim 0.09$, $r \sim 0.81$, and $NMB \sim 17.5\%$. On the other hand, DB (Figure 6b) overestimates most AOD values larger than 1. There is a cluster of points where DB suggested high AOD ($AOD > 0.5$) despite collocated AERONET retrievals at ~ 0.1 . This could potentially be due to smoke plumes within a DB retrieval pixel but not directly over an AERONET station, as the DB pixel resolution is larger than the point of observation from AERONET, or surface or cloud masking artifacts. However, further investigation is required to confirm this. The statistical evaluation for DB shows $RMSE \sim 0.11$, $r \sim 0.86$, and $NMB \sim 21.2\%$. UNet 3+ (Figure 6c) has performance metrics ($RMSE \sim 0.10$, $r \sim 0.80$, and $NMB \sim 6.7\%$) similar to the previous two comparisons. One improvement is that it lacks the cluster of overestimations for low AOD values mentioned previously for DB, suggesting that there is a systematic mischaracterization of certain low AOD events by DB that UNet 3+ can avoid. The gap-filled AOD (Figure 6d) shows performance metrics of $RMSE \sim 0.09$, $r \sim 0.87$, and $NMB \sim 15.5\%$, combining the high correlation from DB with the low RMSE from MERRA-2 and proper characterization by UNet 3+. Additional evaluations of smoke (Figure S4 in Supporting Information S1) and non-smoke (Figure S5 in Supporting Information S1) cases can be found in Section d in Supporting Information S1.

Figure 7 shows residuals as a function of AERONET AOD values for (a) MERRA-2, (b) DB, (c) UNet 3+, and (d) gap-filled AOD, with the blue line showing the mean residual and red area denoting the central ± 1 standard deviation, that is, 68.2% . We can see that each AOD data set has relatively low error and high certainty up until $AOD \sim 0.5$. For AOD values higher than this, we see that MERRA-2 progressively underestimates AOD. DB, on the other hand, begins to progressively overestimate AOD, peaking at $AOD \sim 1.9$ before gradually returning to 0 mean error $AOD \sim 3$. UNet 3+ AOD is interesting in that it remains relatively low error up until around $AOD \sim 1.7$, seemingly finding a perfect middle-ground between MERRA-2 underestimation and DB overestimation. However, this also appears to have come at the expense of a higher uncertainty. Beyond $AOD \sim 1.7$, UNet 3+ begins to underestimate AOD. The gap-filled AOD roughly follows the same residual curve and uncertainty as DB AOD. There is slightly less overestimation between $0 < AOD < 2.5$, but the gap-filled appears to underestimate between $2.5 < AOD < 3$.

4.1.3. Spatial Evaluation

A spatial evaluation between UNet 3+ and DB AOD against AERONET AOD is shown in Figure 5; the figures show results from spatial RMSE and r between AERONET and UNet 3+/DB for all stations that had at least 1,000 collocated points throughout the study period. From overall evaluations (Figure 8), UNet 3+ performs slightly better in RMSE throughout eastern CONUS and the Midwest than DB, for example, $RMSE \sim 0.12$ in Illinois and Texas for DB compared to $RMSE \sim 0.08$ for UNet 3+. The main improvement is arguably the

Total 2012-2023

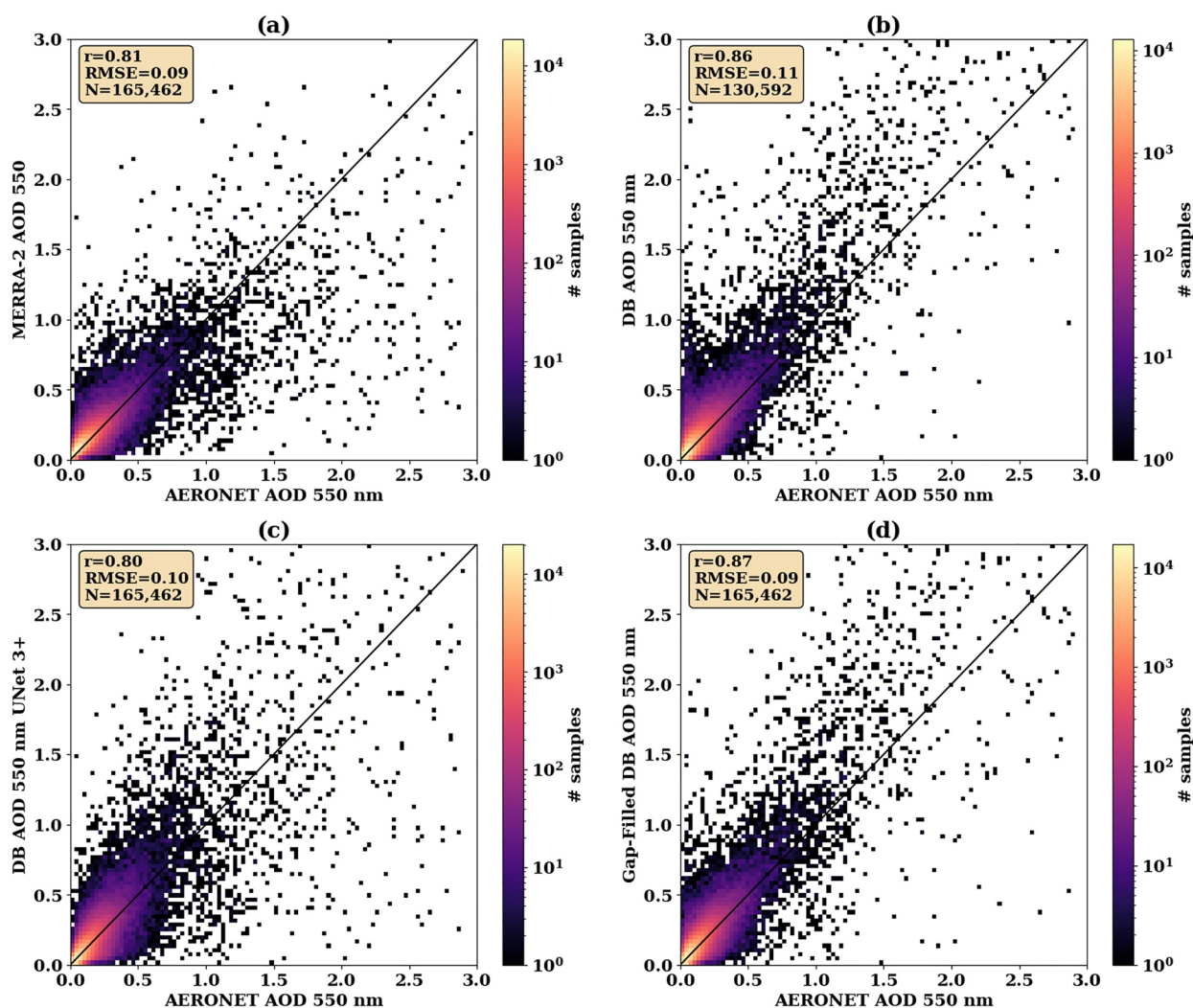


Figure 6. Scatter plots of (a) Modern-Era Retrospective analysis for Research and Applications, Version 2 aerosol optical depth (AOD), (b) deep blue AOD, (c) UNet 3+ AOD, and (d) gap-filled AOD against AERONET AOD at 550 nm. The color scale represents the density of samples, and the black line represents the 1:1 line.

increased coverage since RMSE is already relatively low in this region. Note the increased number of station points in the northern half of the map due to the gap-filling.

UNet 3+ is challenged the most in the Pacific Northwest. RMSEs in this region are increased to 0.15 and above, compared to the mostly <0.08 values present throughout the rest of CONUS. However, DB AOD retrievals also show similar difficulty in properly characterizing AOD loadings in this region ($RMSE > 0.15$). This is likely due to the complex smoke loading and heterogeneous vertical transport, involving fresh smoke from local sources and aged smoke transported from long range. VIIRS and MODIS spectral channels might be challenged in characterizing the loading of aged smoke, as these are better detected in the UV band (Ahn et al., 2014). This challenge could be resolved with the recently launched NASA Plankton, Aerosol, Cloud, ocean Ecosystem (PACE) mission equipped with the hyperspectral Ocean Color Instrument (OCI) (ultraviolet—shortwave infrared). Because of the MODIS and VIIRS spectral bands, aerosol retrievals need to make strong assumptions about aerosol absorption and height. Due to the spectral band capabilities of OCI, it will be possible to retrieve aerosol loading with comparatively more relaxed assumptions about these (Remer et al., 2019). Figures S6–S9 in Supporting Information S1 shows a similar spatial statistical analysis categorized by season.

2012-2023 Residuals

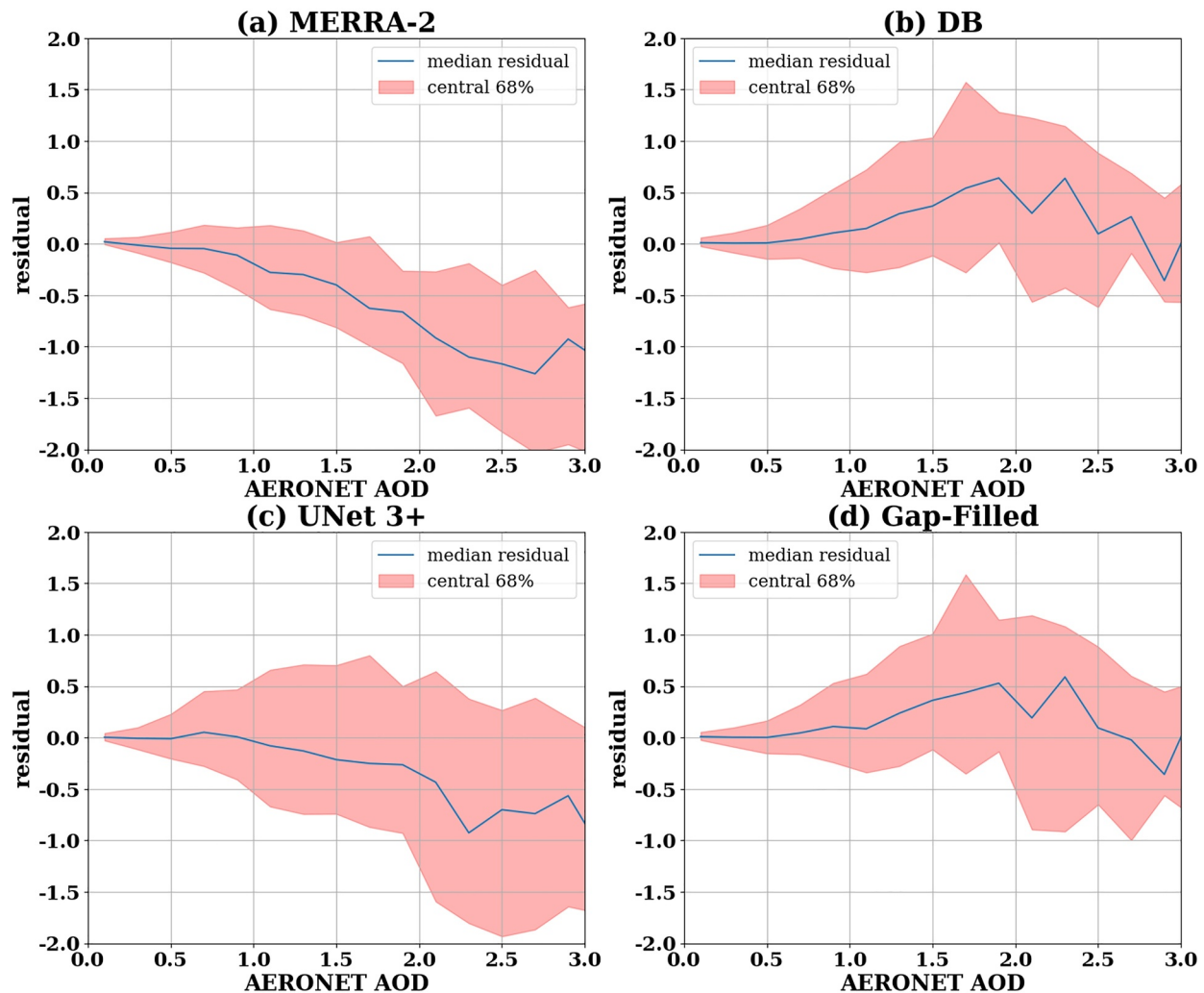


Figure 7. Residuals of AERONET aerosol optical depth (AOD) against (a) Modern-Era Retrospective analysis for Research and Applications, Version 2, (b) deep blue, (c), UNet 3+, and (d) gap-filled AOD. Blue denotes the average residual and red denotes 1 standard deviation.

Figure 9 shows the composite of the daily differences between UNet 3+ predictions and DB AOD retrievals. Over CONUS, UNet 3+, and DB AOD are very closely aligned, outside of a few select areas in California, Oregon, Utah, and Florida, where DB has retrieved much higher AOD values on average compared to UNet 3+. In Canada, there seems to be a similar disagreement between UNet 3+ and DB, particularly in British Columbia, where a historically AOD data-sparse region with complex aerosol loadings, including wildfire smoke, has made estimating aerosol loading particularly difficult. However, we do note that this comparison between UNet 3+ and DB AOD suffers from sampling bias as in many areas, particularly in Canada, where the DB retrieval rate is very low (refer to Figure 5). Section d and Figure S10 in Supporting Information S1 shows the composite difference categorized by season.

4.2. Qualitative Results

Due to the relatively sparse coverage of AERONET stations throughout CONUS and surrounding regions, it is also important to examine if UNet 3+ predictions are plausible when considering seasonality, location, and satellite imagery. This section presents a wildfire case study that provides insight into the model's strengths and weaknesses. A second case study can be found in Figure S11 in Supporting Information S1.

2012-2023 UNet 3+ and Satellite DB AOD vs. AERONET statistics

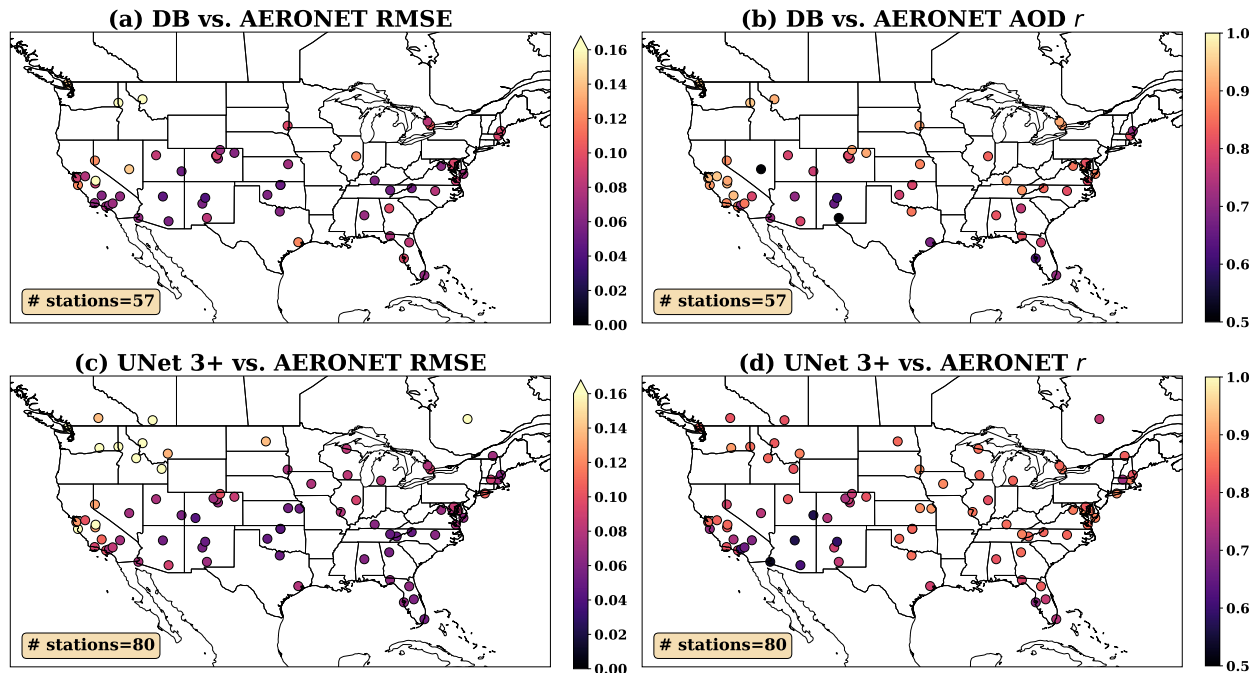


Figure 8. Spatial statistical evaluation of (a) Deep Blue RMSE, (b) Deep Blue r , (c) UNet 3+ RMSE, and (d) UNet 3+ r against AERONET stations throughout CONUS with at least 1,000 collocated observations.

On 30 June 2015, multiple wildfires were experienced throughout Canada, primarily in Saskatchewan and Alberta (circled in red), as well as throughout the U.S. Midwest (circled in blue) (Figure 10e). Parts of these smoke plumes were impeded by cloud cover, leading to gaps in DB AOD retrievals (Figure 10b). Figure 10a shows that MERRA-2 generally retrieves high AOD (>3) around the fire source locations. However, MERRA-2 underestimates the loading and advection for both the Canadian and U.S. wildfires, as seen from the steep drop-off to AOD values around ~ 1.5 surrounding the source regions. UNet 3+ can capture the loading and spread of wildfire smoke (Figure 10c). For the Canadian fires, the nature of the northern (non-cloud-occluded) part of the smoke plume is captured by UNet 3+. However, similar to MERRA-2, the loading is not as high as suggested by satellite retrievals. UNet 3+ predicts AOD ~ 3 for a larger area than MERRA-2, which is more in line with the DB retrievals of AOD > 3 for the entire plume. UNet 3+ suggests that the southern part of the plume (cloud-occluded) is also very dense, which we could not evaluate for DB due to missing retrievals, but the predictions seem reasonable, contextualized by the available satellite retrievals in the gap-filled AOD. Similarly, for the smoke plume over the Midwest, UNet 3+ has more reasonably elevated AOD values of around 2.5–3 over Iowa and Missouri compared to MERRA-2, which has values of around 1.5 for most areas and 3 for just the source region. The same observation also extends to Tennessee and Kentucky, where MERRA-2 has values of AOD from ~ 1.5 to ~ 2 , but DB AOD retrievals are ~ 2.5 . For the gap-filled AOD (Figure 10d), we get the high-resolution smoke plume boundaries where available. While the r and RMSE for gap-filled AOD ($r = 0.95$, RMSE = 0.18) lies between that of UNet 3+ ($r = 0.90$, RMSE = 0.22) and DB ($r = 0.97$, RMSE = 0.12) AOD, the NMB for gap-filled AOD (19.2%) is higher than both UNet 3+ (13.3%) and DB (12.8%) AOD. However, the gap-filled AOD still performs better than MERRA-2 AOD ($r = 0.91$, RMSE = 0.27, NMB = 30.4%) in all three metrics.

Section d in Supporting Information S1 (Figure S14 in Supporting Information S1) provides additional statistics and analysis by season, region, and training rotation.

4.3. Comparison With Existing Methods

Table 3 shows the statistical performance of previous gap-filling models from the literature compared to this study. From Table 3, the gap-filled product from UNet 3+ performs either better or on par with a selection of

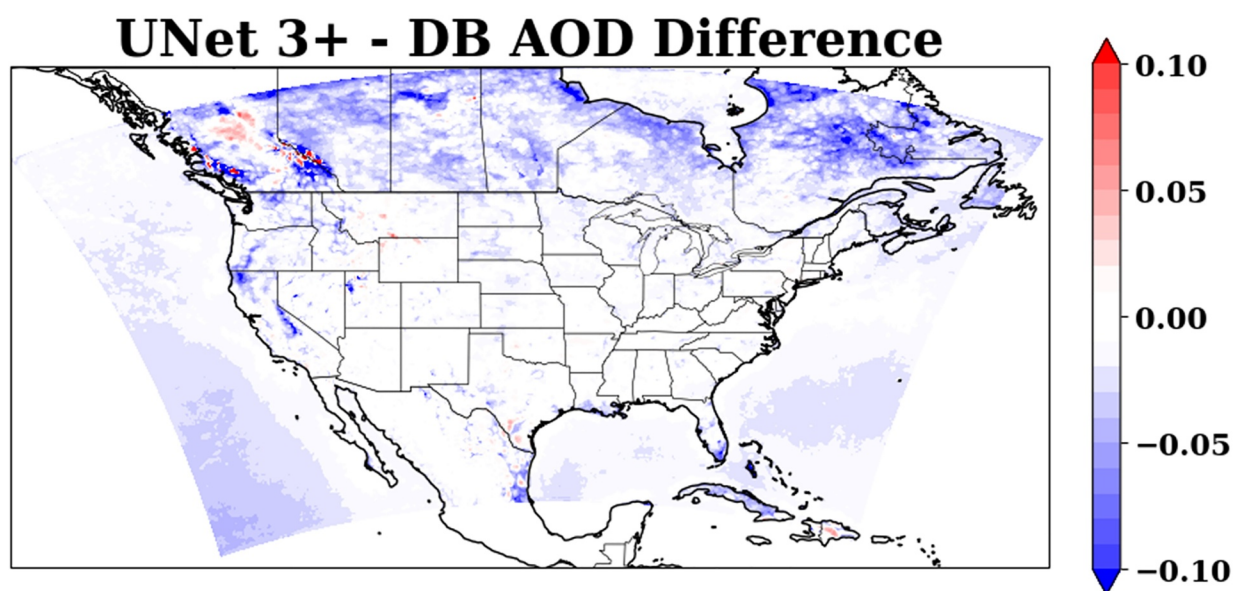


Figure 9. Composite difference between UNet 3+ aerosol optical depth (AOD) and satellite deep blue AOD.

modern gap-filling methods when evaluated against AERONET for RMSE and r . UNet 3+ gap-filling scores for RMSE and r are among the best, and while some methods like the Bayesian method, multi-stage spatiotemporal fitting, and tensor completion have attained better r values, this appears to come at the cost of higher RMSE. Our gap-filling model seems to be very competitive compared to these previous methods, scoring both a low RMSE and a high r .

The vast majority of AOD gap-filling efforts have been focused on Asia, specifically China, as indicated by Table 3. As such, the metrics provided in the table can only serve as a rough point of comparison since the data sources over North America and China have different distributions, biases, aerosol sources, and other characteristics. For instance, Bai et al. (2022) found that Aqua and Terra MODIS MAIAC retrievals over China had an RMSE ~ 0.13 – 0.14 and $r \sim 0.95$, and SNPP VIIRS MAIAC retrievals had an RMSE ~ 0.22 and $r \sim 0.80$. This contrasts with the RMSE ~ 0.09 and $r \sim 0.82$ for the satellite AOD composite used for our study. The approach we present in this study holds value because to the best of our knowledge, it is the first AOD gap-filling model to achieve performance metrics over CONUS comparable to those developed in historically much more studied regions.

5. Summary

In this investigation, we created a spatially complete $12 \times 12 \text{ km}^2$ daily AOD data set over the Continental United States (CONUS) from 2012 to 2023, with an eye toward future air quality studies. Considering the low resolution of reanalysis AOD data sets, like that from MERRA-2, and the data frequently missing from satellite-retrieved AOD data sets, our data set satisfies the demand for a high-resolution and spatially complete data set. We achieved this using DL, specifically by employing a UNet 3+ model to process disparate data sets, including DB AOD and FRP from polar-orbiting satellites (i.e., Terra, Aqua, Suomi-NPP), AOD from the MERRA-2 reanalysis, meteorological and land-use variables from the NAM, and smoke polygons from the HMS, to produce predictions of AOD. We evaluated the model against AOD from NASA Aerosol Robotic Network (AERONET) and found that our gap-filled data set is more performant than MERRA-2 AOD and DB AOD, while again being higher resolution than MERRA-2 and spatially complete unlike DB retrievals. In terms of performance statistics, our gap-filling model seems very competitive compared to earlier AOD gap-filling studies, although the majority of these have been conducted over parts of Asia. Some additional considerations are:

- Training the model to converge at a useful solution (i.e., one that predicts an accurate and dynamic range of AOD values from extremely clean (~ 0) to intense wildfire smoke (> 3) is contingent on data pre-processing steps; namely log-transformation of the AOD inputs, data normalization, and training target augmentation.

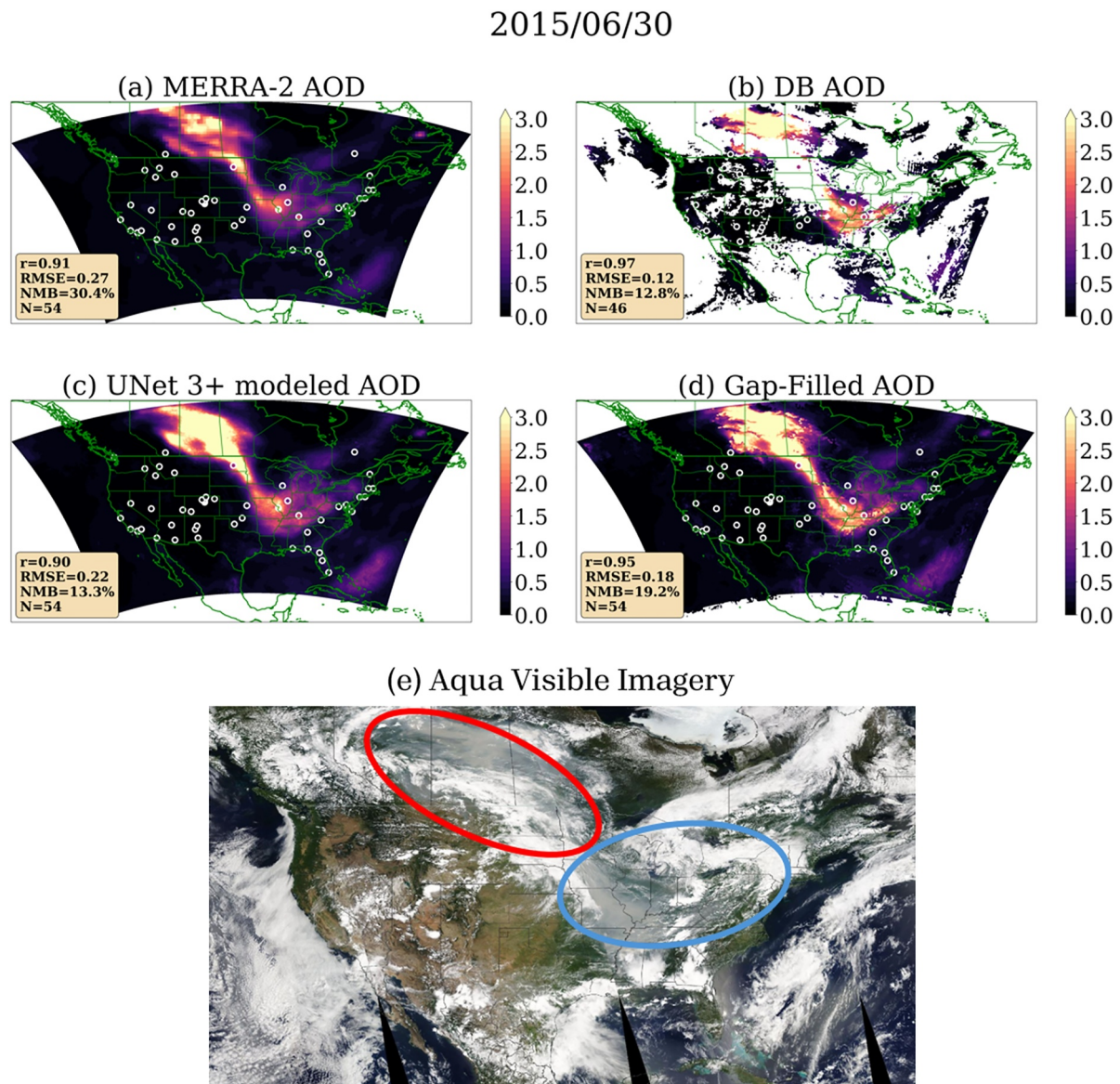


Figure 10. (a) Modern-Era Retrospective analysis for Research and Applications, Version 2 aerosol optical depth, (b) deep blue, (c) UNet 3+, (d) gap-filled, and (e) along with satellite visible imagery from Aqua on 30 June 2015. White circles in panels (a–d) denote AERONET stations with measurements taken on that day. In panel (e), the Canadian smoke plume is circled in red, and the U.S. Midwest plume is circled in blue.

- The UNet 3+ infers aerosol loading and transport through gridded, modeled meteorological variables, and DB AOD data where available. When DB AOD is too sparse to provide enough context, the UNet 3+ relies more on MERRA-2 AOD.
- The additional data sets chosen as inputs to the UNet 3+ enable the recognition of the first day of smoke from fires (HMS and FRP), transport of aerosols (NAM), increased aerosol loading due to temperature inversions (MERRA-2 AOD), and low aerosol loadings (MERRA-2 AOD).
- The model struggles most in regions with extremely low DB data coverage (<30%). Notable areas with challenges are the Pacific Northwest, Canada, and Florida. These regions also presented poor statistical evaluation between DB and AERONET AOD. Results suggest that further investigations are needed to add additional auxiliary data sets to improve model performance.
- A statistical evaluation of UNet 3+ gap-filled AOD with respect to ground-truth AERONET AOD retrievals shows a root mean square error RMSE ~ 0.09 and a Pearson correlation $r \sim 0.87$. This performance is

Table 3
Various Aerosol Optical Depth (AOD) Gap-Filling Models and Their Performance Against AERONET AOD

Method	RMSE	r	Region	Source
Bayesian (still has gaps)	0.08–0.13	0.89–0.93	India	M. K. Singh et al. (2017)
Bayesian maximum entropy (still has gaps)	0.29	0.75	China	Tang et al. (2016)
Random forest	0.20	0.60	China	Jiang et al. (2021)
Multi-stage spatiotemporal fitting	0.24–0.27	0.84–0.88	China	Zhang et al. (2022)
Tensor completion	0.21	0.91	China	Bai et al. (2022)
4D-STET	Did not report	0.73	Global (land)	Wei et al. (2023)
Gap-filled	0.09	0.87	CONUS	This study

- comparable to DB, which scores an RMSE ~ 0.11 and $r \sim 0.86$ in the same evaluation. While the performance is comparable, the data coverage increases by 281% compared to DB retrieved from Aqua, Terra, and S-NPP.
- The UNet 3+ performance indicates that the DL approach allows for generating a spatially complete daily AOD data set over CONUS, even in the semi-arid western U.S., which has historically been a challenging region for this task (Loría-Salazar et al., 2016, 2021).
 - The computational requirements for the UNet 3+ gap-filling model are low compared to numerical and reanalysis first principles transport models of the atmosphere. Using an HPC node with 2 Nvidia A100 GPUs, it took ~ 3 hr to train the model on 12 years of data to convergence and less than a second to generate AOD estimations over CONUS for 1 day.
 - These results support the feasibility of using DL to gap-fill AOD over CONUS. The advantages of using the DL approach are the flexibility of fusing diverse data from disparate sources, extracting high-level features from these data without imposing any assumptions on the problem, and the fast operational runtime. Leveraging these benefits can lead to a better understanding of aerosol distributions at a much finer temporal resolution than possible and enable future aerosol-focused health, economic, and weather studies.
 - A significant focus for future work is how best to incorporate new data from the novel PACE mission. This has capabilities beyond those of MODIS and VIIRS, including observations in the UV and O₂ A and B bands (similar to, e.g., OMI or EPIC). This enables the continuity of heritage algorithm types such as DB. Given these improvements, assimilating PACE data into the existing gap-filling model trained on current DB data would be ideal for improving model performance and ensuring the continuity of the NASA aerosol mission.
 - While our gap-filling approach has shown promising results, several limitations remain. First, the model does not explicitly account for cloud interactions, which limits its suitability for applications related to radiative forcing. Second, because we currently use MERRA-2 as a surrogate training target, the model may underestimate AOD in some cases, particularly where MERRA-2's coarse resolution fails to capture localized AOD spikes. Future work will explore alternative training targets to address this issue. Third, comprehensive validation of the gap-filled AOD data set remains challenging due to the sparse distribution of AERONET stations (a limitation not unique to this study but still significant). Finally, while our current resolution of 12×12 km² improves upon many existing products, it may still be too coarse for certain applications. Increasing spatial resolution will be a key focus in future development.

Data Availability Statement

The MODIS and VIIRS data used in this study is publicly available at <https://ladsweb.modaps.eosdis.nasa.gov/search/>. The data sets can be found by entering the Aerosol tab and selecting MOD04_L2 (Terra MODIS DB), MYD04_L2 (Aqua MODIS DB), and AERDB_L2_VIIRS_SNPP (SNPP VIIRS DB). The MERRA-2 data used in this study is publicly available at <https://search.earthdata.nasa.gov/search?q=M2T1NXAER>. The NAM data used in this study is publicly available at <https://www.ncei.noaa.gov/products/weather-climate-models/north-american-mesoscale>. The reanalysis metadata is found under the NAM Analyses section, along with several links to access the data. The HMS data used in this study is publicly available at <https://www.ospo.noaa.gov/products/land/hms.html#data>. We used the smoke product downloaded as shapefiles. The generated data set is available at Zenodo (J. Lee & Loria Salazar, 2025). The code is available on GitHub (<https://github.com/JeffSMLee/aod-gapfiller>).

Access is permitted under the Creative Commons Attribution license. Please contact the corresponding author for details.

Acknowledgments

Funding for H.A. Holmes was supported by the National Science Foundation (NSF) CAREER Chemical, Bioengineering, Environmental, and Transport Systems (CBET) (2048423) and the NIH National Institute of Environmental Health Sciences (NIEHS) (R01ES029528). The support and resources from the Center for High-Performance Computing at the University of Utah (<https://www.chpc.utah.edu>) are gratefully acknowledged. The AERONET team, Pls, and site managers are thanked for the ongoing generation and stewardship of these data records. Deep blue retrieval algorithm development has been supported through multiple NASA ROSES funding opportunities. This material is partially based upon work supported by the National Aeronautics and Space Administration under Grant 80NSSCM0029 issued through Oklahoma NASA EPSCoR.

References

- Ahn, C., Torres, O., & Jethva, H. (2014). Assessment of OMI near-UV aerosol optical depth over land. *Journal of Geophysical Research: Atmospheres*, 119(5), 2457–2473. <https://doi.org/10.1002/2013JD020188>
- Amato, F., Cassee, F. R., Denier van der Gon, H. A. C., Gehrig, R., Gustafsson, M., Hafner, W., et al. (2014). Urban air quality: The challenge of traffic non-exhaust emissions. *Journal of Hazardous Materials*, 275, 31–36. <https://doi.org/10.1016/j.jhazmat.2014.04.053>
- Andreae, M. O., & Rosenfeld, D. (2008). Aerosol–cloud–precipitation interactions. Part 1. The nature and sources of cloud-active aerosols. *Earth-Science Reviews*, 89(1), 13–41. <https://doi.org/10.1016/j.earscirev.2008.03.001>
- Ångström, A., & Ångström, A. (1929). On the atmospheric transmission of sun radiation and on dust in the air. *Geografiska Annaler*, 11, 156–166. <https://doi.org/10.2307/519399>
- Appel, M. (2024). Efficient data-driven gap filling of satellite image time series using deep neural networks with partial convolutions. *Artificial Intelligence for the Earth Systems*, 1(aop), 220055. <https://doi.org/10.1175/AIES-D-22-0055.1>
- Archibald, S., Lehmann, C. E. R., Gómez-Dans, J. L., & Bradstock, R. A. (2013). Defining pyromes and global syndromes of fire regimes. *Proceedings of the National Academy of Sciences of the United States of America*, 110(16), 6442–6447. <https://doi.org/10.1073/pnas.1211466110>
- Bach, A. J., Brazel, A. J., & Lancaster, N. (1996). Temporal and spatial aspects of blowing dust in the Mojave and Colorado deserts of Southern California, 1973–1994. *Physical Geography*, 17(4), 329–353. <https://doi.org/10.1080/02723646.1996.10642589>
- Bai, K., Li, K., Ma, M., Li, K., Li, Z., Guo, J., et al. (2022). LGHAP: The long-term gap-free high-resolution air pollutant concentration dataset, derived via tensor-flow-based multimodal data fusion. *Earth System Science Data*, 14(2), 907–927. <https://doi.org/10.5194/essd-14-907-2022>
- Bhattacharjee, P. S., Zhang, L., Baker, B., Pan, L., Montuoro, R., Grell, G. A., & McQueen, J. T. (2023). Evaluation of aerosol optical depth forecasts from NOAA's global aerosol forecast model (GEFS-Aerosols). *Weather and Forecasting*, 38(2), 225–249. <https://doi.org/10.1175/WAF-D-22-0083.1>
- Bond, T. C., Bhardwaj, E., Dong, R., Jogani, R., Jung, S., Roden, C., et al. (2007). Historical emissions of black and organic carbon aerosol from energy-related combustion, 1850–2000. *Global Biogeochemical Cycles*, 21(2), GB2018. <https://doi.org/10.1029/2006GB002840>
- Brey, S. J., Ruminski, M., Atwood, S. A., & Fischer, E. V. (2018). Connecting smoke plumes to sources using Hazard Mapping System (HMS) smoke and fire location data over North America. *Atmospheric Chemistry and Physics*, 18(3), 1745–1761. <https://doi.org/10.5194/acp-18-1745-2018>
- Brook, R. D., Rajagopalan, S., Pope, C. A., Brook, J. R., Bhatnagar, A., Diez-Roux, A. V., et al. (2010). Particulate matter air pollution and cardiovascular disease: an update to the scientific statement from the American Heart Association. *Circulation*, 121(21), 2331–2378. <https://doi.org/10.1161/CIR.0b013e3181dbce1>
- Buchard, V., da Silva, A. M., Colarco, P. R., Darmenov, A., Randles, C. A., Govindaraju, R., et al. (2015). Using the OMI aerosol index and absorption aerosol optical depth to evaluate the NASA MERRA aerosol reanalysis. *Atmospheric Chemistry and Physics*, 15(10), 5743–5760. <https://doi.org/10.5194/acp-15-5743-2015>
- Buckeridge, D. L., Glazier, R., Harvey, B. J., Escobar, M., Amrhein, C., & Frank, J. (2002). Effect of motor vehicle emissions on respiratory health in an urban area. *Environmental Health Perspectives*, 110(3), 293–300. <https://doi.org/10.1289/ehp.02110293>
- Burke, M., Driscoll, A., Heft-Neal, S., Xue, J., Burney, J., & Wara, M. (2021). The changing risk and burden of wildfire in the United States. *Proceedings of the National Academy of Sciences of the United States of America*, 118(2), e2011048118. <https://doi.org/10.1073/pnas.2011048118>
- Chang, H. H., Hu, X., & Liu, Y. (2014). Calibrating MODIS aerosol optical depth for predicting daily PM_{2.5} concentrations via statistical downscaling. *Journal of Exposure Science and Environmental Epidemiology*, 24(4), 398–404. <https://doi.org/10.1038/jes.2013.90>
- Charlson, R. J., Schwartz, S. E., Hales, J. M., Cess, R. D., Coakley, J. A., Hansen, J. E., & Hofmann, D. J. (1992). Climate forcing by anthropogenic aerosols. *Science*, 255(5043), 423–430. <https://doi.org/10.1126/science.255.5043.423>
- Chen, G., Chen, J., Dong, G., Yang, B., Liu, Y., Lu, T., et al. (2021). Improving satellite-based estimation of surface ozone across China during 2008–2019 using iterative random forest model and high-resolution grid meteorological data. *Sustainable Cities and Society*, 69, 102807. <https://doi.org/10.1016/j.scs.2021.102807>
- Chen, J., Lu, Y., Yu, Q., Luo, X., Adeli, E., Wang, Y., et al. (2021). TransUNet: Transformers make strong encoders for medical image segmentation. *arXiv*. <https://doi.org/10.48550/arXiv.2102.04306>
- Cheng, M.-F., Ho, S.-C., Chiu, H.-F., Wu, T.-N., Chen, P.-S., & Yang, C.-Y. (2008). Consequences of exposure to Asian dust storm events on daily pneumonia hospital admissions in Taipei, Taiwan. *Journal of Toxicology and Environmental Health, Part A*, 71(19), 1295–1299. <https://doi.org/10.1080/15287390802114808>
- Chien, L.-C., Yang, C.-H., & Yu, H.-L. (2012). Estimated effects of Asian dust storms on spatiotemporal distributions of clinic visits for respiratory diseases in Taipei children (Taiwan). *Environmental Health Perspectives*, 120(8), 1215–1220. <https://doi.org/10.1289/ehp.1104417>
- Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., et al. (2021). An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv*. <https://doi.org/10.48550/arXiv.2010.11929>
- Dubovik, O., Sinyuk, A., Lapyonok, T., Holben, B. N., Mishchenko, M., Yang, P., et al. (2006). Application of spheroid models to account for aerosol particle nonsphericity in remote sensing of desert dust. *Journal of Geophysical Research*, 111(D11), D11208. <https://doi.org/10.1029/2005JD006619>
- Dubovik, O., Smirnov, A., Holben, B. N., King, M. D., Kaufman, Y. J., Eck, T. F., & Slutsker, I. (2000). Accuracy assessments of aerosol optical properties retrieved from Aerosol Robotic Network (AERONET) Sun and sky radiance measurements. *Journal of Geophysical Research*, 105(D8), 9791–9806. <https://doi.org/10.1029/2000JD900040>
- Eck, T. F., Holben, B. N., Reid, J. S., O'Neill, N. T., Schafer, J. S., Dubovik, O., et al. (2003). High aerosol optical depth biomass burning events: A comparison of optical properties for different source regions. *Geophysical Research Letters*, 30(20), 2035. <https://doi.org/10.1029/2003GL017861>
- Gelaro, R., McCarty, W., Suárez, M. J., Todling, R., Molod, A., Takacs, L., et al. (2017). The Modern-Era Retrospective Analysis for Research and Applications, version 2 (MERRA-2). *Journal of Climate*, 30(14), 5419–5454. <https://doi.org/10.1175/JCLI-D-16-0758.1>
- Giglio, L., Descloitres, J., Justice, C. O., & Kaufman, Y. J. (2003). An enhanced contextual fire detection algorithm for MODIS. *Remote Sensing of Environment*, 87(2), 273–282. [https://doi.org/10.1016/S0034-4257\(03\)00184-6](https://doi.org/10.1016/S0034-4257(03)00184-6)

- Giglio, L., Schroeder, W., & Justice, C. O. (2016). The collection 6 MODIS active fire detection algorithm and fire products. *Remote Sensing of Environment*, 178, 31–41. <https://doi.org/10.1016/j.rse.2016.02.054>
- Giles, D. M., Sinyuk, A., Sorokin, M. G., Schafer, J. S., Smirnov, A., Slutsker, I., et al. (2019). 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(1), 169–209. <https://doi.org/10.5194/amt-12-169-2019>
- Gillette, D. A., & Hanson, K. J. (1989). Spatial and temporal variability of dust production caused by wind erosion in the United States. *Journal of Geophysical Research*, 94(D2), 2197–2206. <https://doi.org/10.1029/JD094iD02p02197>
- Grant, E., & Runkle, J. D. (2022). Long-term health effects of wildfire exposure: A scoping review. *The Journal of Climate Change and Health*, 6, 100110. <https://doi.org/10.1016/j.joclim.2021.100110>
- Gupta, P., Christopher, S. A., Wang, J., Gehrig, R., Lee, Y., & Kumar, N. (2006). Satellite remote sensing of particulate matter and air quality assessment over global cities. *Atmospheric Environment*, 40(30), 5880–5892. <https://doi.org/10.1016/j.atmosenv.2006.03.016>
- Guttikunda, S. K., & Kopakka, R. V. (2014). Source emissions and health impacts of urban air pollution in Hyderabad, India. *Air Quality, Atmosphere & Health*, 7(2), 195–207. <https://doi.org/10.1007/s11869-013-0221-z>
- Herrera-Molina, E., Gill, T. E., Ibarra-Mejia, G., Jeon, S., & Ardon-Dryer, K. (2024). Associations between dust exposure and hospitalizations in a dust-prone city, Lubbock, TX, USA. *Air Quality, Atmosphere & Health*, 17(5), 1091–1105. <https://doi.org/10.1007/s11869-023-01489-9>
- Holben, B. N., Eck, T. F., Slutsker, I., Tanré, D., Buis, J. P., Setzer, A., et al. (1998). AERONET—A federated instrument network and data archive for aerosol characterization. *Remote Sensing of Environment*, 66(1), 1–16. [https://doi.org/10.1016/S0034-4257\(98\)00031-5](https://doi.org/10.1016/S0034-4257(98)00031-5)
- Hsu, N. C., Lee, J., Sayer, A. M., Kim, W., Bettenhausen, C., & Tsay, S.-C. (2019). VIIRS deep blue aerosol products over land: Extending the EOS long-term aerosol data records. *Journal of Geophysical Research: Atmospheres*, 124(7), 4026–4053. <https://doi.org/10.1029/2018JD029688>
- Hsu, N. C., Tsay, S.-C., King, M. D., & Herman, J. R. (2004). Aerosol properties over bright-reflecting source regions. *IEEE Transactions on Geoscience and Remote Sensing*, 42(3), 557–569. <https://doi.org/10.1109/TGRS.2004.824067>
- Hu, X., Waller, L. A., Lyapustin, A., Wang, Y., Al-Hamdan, M. Z., Crosson, W. L., et al. (2014). Estimating ground-level PM_{2.5} concentrations in the southeastern United States using MAIAC AOD retrievals and a two-stage model. *Remote Sensing of Environment*, 140, 220–232. <https://doi.org/10.1016/j.rse.2013.08.032>
- Huang, H., Lin, L., Tong, R., Hu, H., Zhang, Q., Iwamoto, Y., et al. (2020). UNet 3+: A full-scale connected UNet for medical image segmentation. In *ICASSP 2020 - 2020 IEEE international conference on acoustics, speech and signal processing (ICASSP)* (pp. 1055–1059). IEEE. <https://doi.org/10.1109/ICASSP40776.2020.9053405>
- Ichoku, C., & Kaufman, Y. J. (2005). A method to derive smoke emission rates from MODIS fire radiative energy measurements. *IEEE Transactions on Geoscience and Remote Sensing*, 43(11), 2636–2649. <https://doi.org/10.1109/TGRS.2005.857328>
- Jacobson, M. C., Hansson, H.-C., Noone, K. J., & Charlson, R. J. (2000). Organic atmospheric aerosols: Review and state of the science. *Reviews of Geophysics*, 38(2), 267–294. <https://doi.org/10.1029/1998RG000045>
- Janjic, Z. I. (2003). A nonhydrostatic model based on a new approach. *Meteorology and Atmospheric Physics*, 82(1), 271–285. <https://doi.org/10.1007/s00703-001-0587-6>
- Janjic, Z. I., Black, T., Pyle, M., Chuang, H., Rogers, E., & DiMego, G. (2005). The NCEP WRF core. In *20th conference on weather analysis and forecasting/16th conference on numerical weather prediction, Seattle, Washington* (pp. 10–25).
- Jiang, T., Chen, B., Nie, Z., Ren, Z., Xu, B., & Tang, S. (2021). Estimation of hourly full-coverage PM_{2.5} concentrations at 1-km resolution in China using a two-stage random forest model. *Atmospheric Research*, 248, 105146. <https://doi.org/10.1016/j.atmosres.2020.105146>
- Justin, A. D., Willingham, C., McGovern, A., & Allen, J. T. (2023). Toward operational real-time identification of frontal boundaries using machine learning. *Artificial Intelligence for the Earth Systems*, 2(3). <https://doi.org/10.1175/AIES-D-22-0052.1>
- Kang, Y., Man, Y. B., Cheung, K. C., & Wong, M. H. (2012). Risk assessment of human exposure to bioaccessible phthalate esters via indoor dust around the Pearl River Delta. *Environmental Science & Technology*, 46(15), 8422–8430. <https://doi.org/10.1021/es300379v>
- Kaufman, Y. J., Justice, C. O., Flynn, L. P., Kendall, J. D., Prins, E. M., Giglio, L., et al. (1998). Potential global fire monitoring from EOS-MODIS. *Journal of Geophysical Research*, 103(D24), 32215–32238. <https://doi.org/10.1029/98JD01644>
- Kaufman, Y. J., Tanré, D., Remer, L. A., Vermote, E. F., Chu, A., & Holben, B. N. (1997). Operational remote sensing of tropospheric aerosol over land from EOS moderate resolution imaging spectroradiometer. *Journal of Geophysical Research*, 102(D14), 17051–17067. <https://doi.org/10.1029/96JD03988>
- Kennedy, M., & Kopp, S. (2001). *Understanding map projections* (8th ed.). ESRI Press.
- Kim, K.-H., Kabir, E., & Kabir, S. (2015). A review on the human health impact of airborne particulate matter. *Environment International*, 74, 136–143. <https://doi.org/10.1016/j.envint.2014.10.005>
- Kim, Y., Knowles, S., Manley, J., & Radoias, V. (2017). Long-run health consequences of air pollution: Evidence from Indonesia's forest fires of 1997. *Economics and Human Biology*, 26, 186–198. <https://doi.org/10.1016/j.ehb.2017.03.006>
- Kingma, D. P., & Ba, J. (2017). Adam: A method for stochastic optimization. *arXiv*. <https://doi.org/10.48550/arXiv.1412.6980>
- Lee, H., Calvin, K., Dasgupta, D., Krinner, G., Mukherji, A., Thorne, P., et al. (2023). *Climate change 2023: Synthesis report. Contribution of working groups I, II and III to the sixth assessment report of the intergovernmental panel on climate change (report)*. The Australian National University. Retrieved from <https://openresearch-repository.anu.edu.au/handle/1885/299630>
- Lee, H., Kim, H., Honda, Y., Lim, Y.-H., & Yi, S. (2013). Effect of Asian dust storms on daily mortality in seven metropolitan cities of Korea. *Atmospheric Environment*, 79, 510–517. <https://doi.org/10.1016/j.atmosenv.2013.06.046>
- Lee, J., & Loria Salazar, S. M. (2025). Gapfilled DB AOD [Dataset]. *Zenodo*. Retrieved from <https://zenodo.org/records/15570470>
- Leibensperger, E. M., Mickley, L. J., Jacob, D. J., Chen, W.-T., Seinfeld, J. H., Nenes, A., et al. (2012). Climatic effects of 1950–2050 changes in US anthropogenic aerosols – Part 1: Aerosol trends and radiative forcing. *Atmospheric Chemistry and Physics*, 12(7), 3333–3348. <https://doi.org/10.5194/acp-12-3333-2012>
- Loria-Salazar, S. M., Holmes, H. A., Arnott, W. P., Barnard, J. C., & Moosmüller, H. (2016). Evaluation of MODIS columnar aerosol retrievals using AERONET in semi-arid Nevada and California, U.S.A., during the summer of 2012. *Atmospheric Environment*, 144, 345–360. <https://doi.org/10.1016/j.atmosenv.2016.08.070>
- Loria-Salazar, S. M., Sayer, A. M., Barnes, J., Huang, J., Flynn, C., Lareau, N., et al. (2021). Evaluation of novel NASA MODerate Resolution Imaging Spectroradiometer and Visible Infrared Imaging Radiometer Suite aerosol products and assessment of smoke height boundary layer ratio during extreme smoke events in the western USA. *Journal of Geophysical Research: Atmospheres*, 126(11), e2020JD034180. <https://doi.org/10.1029/2020JD034180>
- Lv, B., Hu, Y., Chang, H. H., Russell, A. G., & Bai, Y. (2016). Improving the accuracy of daily PM_{2.5} distributions derived from the fusion of ground-level measurements with aerosol optical depth observations, a case study in North China. *Environmental Science & Technology*, 50(9), 4752–4759. <https://doi.org/10.1021/acs.est.5b05940>

- Lyapustin, A., Wang, Y., Korkin, S., & Huang, D. (2018). MODIS collection 6 MAIAC algorithm. *Atmospheric Measurement Techniques*, 11(10), 5741–5765. <https://doi.org/10.5194/amt-11-5741-2018>
- Martinelli, N., Olivieri, O., & Girelli, D. (2013). Air particulate matter and cardiovascular disease: A narrative review. *European Journal of Internal Medicine*, 24(4), 295–302. <https://doi.org/10.1016/j.ejim.2013.04.001>
- Matz, C. J., Egyed, M., Xi, G., Racine, J., Pavlovic, R., Rittmaster, R., et al. (2020). Health impact analysis of PM_{2.5} from wildfire smoke in Canada (2013–2015, 2017–2018). *Science of the Total Environment*, 725, 138506. <https://doi.org/10.1016/j.scitotenv.2020.138506>
- McNamara, D., Stephens, G., & Ruminski, M. (2004). *The hazard mapping system (HMS) - NOAA'S multi-sensor fire and smoke detection program using environmental satellites*. NOAA/NESDIS.
- Panikkath, R., Jumper, C. A., & Mulkey, Z. (2013). Multilobar lung infiltrates after exposure to dust storm: The Haboob Lung Syndrome. *The American Journal of Medicine*, 126(2), e5–e7. <https://doi.org/10.1016/j.amjmed.2012.08.012>
- Peterson, D. A., Fromm, M. D., Solbrig, J. E., Hyer, E. J., Surratt, M. L., & Campbell, J. R. (2017). Detection and inventory of intense pyroconvection in western North America using GOES-15 daytime infrared data. *Journal of Applied Meteorology and Climatology*, 56(2), 471–493. <https://doi.org/10.1175/JAMC-D-16-0226.1>
- Peterson, D. A., Hyer, E. J., Campbell, J. R., Fromm, M. D., Hair, J. W., Butler, C. F., & Fenn, M. A. (2014). The 2013 rim fire: Implications for predicting extreme fire spread, pyroconvection, and smoke emissions. *Bulletin of the American Meteorological Society*, 96(2), 229–247. <https://doi.org/10.1175/BAMS-D-14-00060.1>
- Péwé, T. L., & American Association for the Advancement of Science. (1981). *Desert dust: Origin, characteristics, and effect on man*. Geological Society of America.
- Pöschl, U. (2005). Atmospheric aerosols: Composition, transformation, climate and health effects. *Angewandte Chemie International Edition*, 44(46), 7520–7540. <https://doi.org/10.1002/anie.200501122>
- Randles, C. A., Silva, A. M. D., Buchard, V., Colarco, P. R., Darnenov, A., Govindaraju, R., et al. (2017). The MERRA-2 aerosol reanalysis, 1980 onward. Part I: System description and data assimilation evaluation. *Journal of Climate*, 30(17), 6823–6850. <https://doi.org/10.1175/JCLI-D-16-0609.1>
- Reid, C. E., Brauer, M., Johnston, F. H., Jerrett, M., Balmes, J. R., & Elliott, C. T. (2016). Critical review of health impacts of wildfire smoke exposure. *Environmental Health Perspectives*, 124(9), 1334–1343. <https://doi.org/10.1289/ehp.1409277>
- Reid, J. S., Koppmann, R., Eck, T. F., & Eleuterio, D. P. (2005). A review of biomass burning emissions part II: Intensive physical properties of biomass burning particles. *Atmospheric Chemistry and Physics*, 5(3), 799–825. <https://doi.org/10.5194/acp-5-799-2005>
- Remer, L. A., Davis, A. B., Mattoo, S., Levy, R. C., Kalashnikova, O. V., Coddington, O., et al. (2019). Retrieving aerosol characteristics from the PACE mission, Part 1: Ocean color instrument. *Frontiers in Earth Science*, 7. <https://doi.org/10.3389/feart.2019.00152>
- 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>
- Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional networks for biomedical image segmentation. *arXiv*. <https://doi.org/10.48550/arXiv.1505.04597>
- Roy, D. P., & Kumar, S. S. (2017). Multi-year MODIS active fire type classification over the Brazilian Tropical Moist Forest Biome. *International Journal of Digital Earth*, 10(1), 54–84. <https://doi.org/10.1080/17538947.2016.1208686>
- Russell, P. B., Bergstrom, R. W., Shinozuka, Y., Clarke, A. D., DeCarlo, P. F., Jimenez, J. L., et al. (2010). Absorption angstrom exponent in AERONET and related data as an indicator of aerosol composition. *Atmospheric Chemistry and Physics*, 10(3), 1155–1169. <https://doi.org/10.5194/acp-10-1155-2010>
- Saharia, C., Chan, W., Chang, H., Lee, C. A., Ho, J., Salimans, T., et al. (2022). Palette: Image-to-image diffusion models. *arXiv*, 1–10. <https://doi.org/10.48550/arXiv.2111.05826>
- Sayer, A. M., Hsu, N. C., Lee, J., Kim, W. V., Dubovik, O., Dutcher, S. T., et al. (2018). Validation of SOAR VIIRS over-water aerosol retrievals and context within the global satellite aerosol data record. *Journal of Geophysical Research: Atmospheres*, 123(23), 13496–13526. <https://doi.org/10.1029/2018JD029465>
- Sayer, A. M., Hsu, N. C., Lee, J., Kim, W. V., & Dutcher, S. T. (2019). Validation, stability, and consistency of MODIS collection 6.1 and VIIRS version 1 deep blue aerosol data over land. *Journal of Geophysical Research: Atmospheres*, 124(8), 4658–4688. <https://doi.org/10.1029/2018JD029598>
- Singh, M. K., Gautam, R., & Venkatachalam, P. (2017). Bayesian merging of MISR and MODIS aerosol optical depth products using error distributions from AERONET. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 10(12), 5186–5200. <https://doi.org/10.1109/JSTARS.2017.2734331>
- Singh, R., Singh, V., Gautam, A. S., Gautam, S., Sharma, M., Soni, P. S., et al. (2023). Temporal and spatial variations of satellite-based aerosol optical depths, angstrom exponent, single scattering albedo, and ultraviolet-aerosol index over five polluted and less-polluted cities of northern India: Impact of urbanization and climate change. *Aerosol Science and Engineering*, 7(1), 131–149. <https://doi.org/10.1007/s41810-022-00168-z>
- Sinyuk, A., Holben, B. N., Eck, T. F., Giles, D. M., Slutsker, I., Korkin, S., et al. (2020). The AERONET version 3 aerosol retrieval algorithm, associated uncertainties and comparisons to version 2. *Atmospheric Measurement Techniques*, 13(6), 3375–3411. <https://doi.org/10.5194/amt-13-3375-2020>
- Tang, Q., Bo, Y., & Zhu, Y. (2016). Spatiotemporal fusion of multiple-satellite aerosol optical depth (AOD) products using Bayesian maximum entropy method. *Journal of Geophysical Research: Atmospheres*, 121(8), 4034–4048. <https://doi.org/10.1002/2015JD024571>
- Uda, S. K., Hein, L., & Atmoko, D. (2019). Assessing the health impacts of peatland fires: A case study for Central Kalimantan, Indonesia. *Environmental Science and Pollution Research*, 26(30), 31315–31327. <https://doi.org/10.1007/s11356-019-06264-x>
- van Donkelaar, A., Martin, R. V., Brauer, M., Hsu, N. C., Kahn, R. A., Levy, R. C., et al. (2016). Global estimates of fine particulate matter using a combined geophysical-statistical method with information from satellites, models, and monitors. *Environmental Science & Technology*, 50(7), 3762–3772. <https://doi.org/10.1021/acs.est.5b05833>
- Wang, Z., Simoncelli, E. P., & Bovik, A. C. (2003). Multiscale structural similarity for image quality assessment. In *The thirty-seventh Asilomar conference on signals, systems & computers* (Vol. 2, pp. 1398–1402). <https://doi.org/10.1109/ACSSC.2003.1292216>
- Wei, J., Huang, W., Li, Z., Xue, W., Peng, Y., Sun, L., & Cribb, M. (2019). Estimating 1-km-resolution PM_{2.5} concentrations across China using the space-time random forest approach. *Remote Sensing of Environment*, 231, 111221. <https://doi.org/10.1016/j.rse.2019.111221>
- Wei, J., Li, Z., Lyapustin, A., Wang, J., Dubovik, O., Schwartz, J., et al. (2023). First close insight into global daily gapless 1 km PM_{2.5} pollution, variability, and health impact. *Nature Communications*, 14(1), 8349. <https://doi.org/10.1038/s41467-023-43862-3>
- Wilcox, L. J., Highwood, E. J., & Dunstone, N. J. (2013). The influence of anthropogenic aerosol on multi-decadal variations of historical global climate. *Environmental Research Letters*, 8(2), 024033. <https://doi.org/10.1088/1748-9326/8/2/024033>

- Wooster, M. J., & Zhang, Y. H. (2004). Boreal forest fires burn less intensely in Russia than in North America. *Geophysical Research Letters*, 31(20), L20505. <https://doi.org/10.1029/2004GL020805>
- Wooster, M. J., Zhukov, B., & Oertel, D. (2003). Fire radiative energy for quantitative study of biomass burning: Derivation from the BIRD experimental satellite and comparison to MODIS fire products. *Remote Sensing of Environment*, 86(1), 83–107. [https://doi.org/10.1016/S0034-4257\(03\)00070-1](https://doi.org/10.1016/S0034-4257(03)00070-1)
- Wu, Y., Gan, C., Gross, B., Moshary, F., & Ahmed, S. (2009). Aerosol plume observations by the ground-based lidar, sunphotometer and satellite: Cases analysis. *Proceedings of SPIE - The International Society for Optical Engineering*, 7479, 74790U. <https://doi.org/10.1117/12.830639>
- Xiao, L., Lang, Y., & Christakos, G. (2018). High-resolution spatiotemporal mapping of PM_{2.5} concentrations at Mainland China using a combined BME-GWR technique. *Atmospheric Environment*, 173, 295–305. <https://doi.org/10.1016/j.atmosenv.2017.10.062>
- Zhang, T., Zhou, Y., Zhao, K., Zhu, Z., Asrar, G. R., & Zhao, X. (2022). Gap-filling MODIS daily aerosol optical depth products by developing a spatiotemporal fitting algorithm. *GIScience and Remote Sensing*, 59(1), 762–781. <https://doi.org/10.1080/15481603.2022.2060596>
- Zhou, Z., Rahman Siddiquee, M. M., Tajbakhsh, N., & Liang, J. (2018). UNet++: A nested U-Net architecture for medical image segmentation. In D. Stoyanov, Z. Taylor, G. Carneiro, T. Syeda-Mahmood, A. Martel, L. Maier-Hein, et al. (Eds.), *Deep learning in medical image analysis and multimodal learning for clinical decision support* (pp. 3–11). Springer International Publishing. https://doi.org/10.1007/978-3-030-00889-5_1

References From the Supporting Information

- Barnes, W. L., & Salomonson, V. V. (1992). MODIS: A global imaging spectroradiometer for the Earth Observing System. In *Optical technologies for aerospace sensing: A critical review* (Vol. 10269, pp. 280–302). SPIE. <https://doi.org/10.1117/12.161578>
- Giglio, L. (2007). Characterization of the tropical diurnal fire cycle using VIRS and MODIS observations. *Remote Sensing of Environment*, 108(4), 407–421. <https://doi.org/10.1016/j.rse.2006.11.018>
- Hillger, D., Seaman, C., Liang, C., Miller, S., Lindsey, D., & Kopp, T. (2014). Suomi NPP VIIRS Imagery evaluation. *Journal of Geophysical Research: Atmospheres*, 119(11), 6440–6455. <https://doi.org/10.1002/2013JD021170>
- Hsu, N. C. (2017). Changes to MODIS deep blue aerosol products between collection 6 and collection 6.1.
- Hsu, N. C., Jeong, M.-J., Bettenhausen, C., Sayer, A. M., Hansell, R., Seftor, C. S., et al. (2013). Enhanced deep blue aerosol retrieval algorithm: The second generation. *Journal of Geophysical Research: Atmospheres*, 118(16), 9296–9315. <https://doi.org/10.1002/jgrd.50712>
- Jo, J.-M. (2019). Effectiveness of normalization pre-processing of big data to the machine learning performance. *The Journal of the Korea Institute of Electronic Communication Sciences*, 14(3), 547–552. <https://doi.org/10.13067/JKIECS.2019.14.3.547>
- Muhammad Ali, P. J., & Faraj, R. H. (2014). Data Normalization and Standardization: A Technical Report. In *Machine learning technical reports* (pp. 1–6).
- Sayer, A. M. (2020). How long is too long? Variogram analysis of AERONET data to aid aerosol validation and intercomparison studies. *Earth and Space Science*, 7(9), e2020EA001290. <https://doi.org/10.1029/2020EA001290>
- Sayer, A. M., Hsu, N. C., Bettenhausen, C., & Jeong, M.-J. (2013). Validation and uncertainty estimates for MODIS collection 6 “Deep Blue” aerosol data. *Journal of Geophysical Research: Atmospheres*, 118(14), 7864–7872. <https://doi.org/10.1002/jgrd.50600>
- Singh, D., & Singh, B. (2020). Investigating the impact of data normalization on classification performance. *Applied Soft Computing*, 97, 105524. <https://doi.org/10.1016/j.asoc.2019.105524>