UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

SPATIOTEMPORAL GAP-FILLING OF NASA DEEP BLUE AEROSOL OPTICAL DEPTH
OVER CONUS USING THE UNET 3+ ARCHITECTURE

A THESIS

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

MASTER OF SCIENCE

By

Jeffrey Lee

Norman, Oklahoma

2024

SPATIOTEMPORAL GAP-FILLING OF NASA DEEP BLUE AEROSOL
OPTICAL DEPTH
OVER CONUS USING THE UNET 3+ ARCHITECTURE

A THESIS APPROVED FOR THE

SCHOOL OF METEOROLOGY

BY THE COMMITTEE CONSISTING OF

Dr. Marcela Loría-Salazar, Chair

Dr. Amy McGovern

Dr. Feng Xu

Dr. Jason Furtado

Acknowledgements

Firstly, I would like to thank my research advisor Dr. Marcela Loría-Salazar, who has offered me countless hours of instruction, research guidance, and mental support. This research would never have materialized without her steadfast presence.

I would also like to thank the members of the committee, Dr. Amy McGovern, Dr. Feng Xu, and Dr. Jason Furtado, for graciously agreeing to review my work, provide their valuable insights, and accommodate the various delays that have taken place throughout the process.

I would also like to thank Dr. Heather Holmes at the University of Utah and her research group for allowing me access to the CHPC supercomputing resources. Without these resources, training a model of our size would have been an impracticality.

I want to thank the faculty and staff at the University of Oklahoma's School of Meteorology for providing such a wonderful learning environment. I have learned so much about things I had never dreamed of before from the passionate and experienced educators here at SoM. The wonderful staff have also made sure my logistical, administrative, and financial issues were kept at a minimum.

Finally, I would like to thank my friends and family who have stuck with me through the years. You guys are the best.

List of Tables

List of Figures

Table of Contents

Abstract

Due to sensor and algorithmic constraints, satellite aerosol optical depth (AOD) retrievals are spatially incomplete over clouds, deserts, and other bright surfaces. These gaps in satellite AOD datasets represent a significant challenge in characterizing aerosol distributions at a daily temporal resolution. These challenges are essential to overcome due to aerosol impacts on human health, the economy, weather, and climate. The task of filling these gaps in AOD datasets is known as AOD gap-filling. In this study, we propose using a deep learning (DL) architecture called UNet 3+ to perform this task. The model is trained on Deep Blue (DB) AOD retrievals from Terra, Aqua, and S-NPP, MERRA-2 reanalysis AOD, meteorological and land-use variables from NAM, and HMS smoke polygons. Through spatial evaluations against AERONET and DB AOD, we show that such an approach is feasible over CONUS, even in the semi-arid western U.S. where historically, topography, bright surfaces, and snowpacks have made AOD gap-filling a challenging problem. We created spatiotemporal datasets of daily gap-filled DB AOD from 2012-2022 over CONUS at a 12 x 12 km^2 resolution with statistical evaluations of RMSE~0.08 and r~0.84 against collocated AERONET retrievals. This dataset will be a starting point for future aerosol-related studies, such as acute daily $PM_{2.5}$ (particulate matter with an aerodynamic diameter smaller than 2.5 µm) exposure studies. Some potential challenges exist with the suggested approaches and, more generally, in estimating AOD and $PM_{2.5}$. The first is the sampling bias that naturally arises from AOD retrievals, whether from ground- or satellite-based sensors. As no retrievals are performed in cloudy pixels, any model trained on this data is only aware of the dynamics of clear-sky AOD, and we cannot directly validate estimations over cloudy-sky areas (not even with ground truth AERONET). However, because this research aims

to use AOD to estimate $PM_{2.5}$ and hydrological effects tend to lower aerosol concentrations (e.g., washout from precipitation), the methods can be justified as providing an upper-bound (acute) estimate for exposure.

1. Background and Introduction

1.1 Aerosol Definition

Atmospheric aerosols are small airborne solid and liquid particles that range in size from nanometers to micrometers (Pöschl, 2005). These aerosols come in diverse forms (everything from dust, sea salt, ash, smoke, pollen, etc.) and are emitted from various sources, both natural and anthropogenic (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 (Charlson et al., 1992; Lee et al., 2023; Leibensperger et al., 2012; Wilcox et al., 2013). Aerosols may be small but have a significant impact on the Earth. For instance, they can cause health complications when inhaled and alter the weather and climate through their light absorption/scattering properties and cloud interactions.

Unfortunately, the mechanisms behind the effects of aerosols on the climate are complicated and not yet well-characterized . The largest uncertainty for radiative forcings in climate models lies with aerosols (Lee et al., 2023). These uncertainties can be attributed to the difficulty in characterizing aerosol chemical and physical properties, evolution, cloud interactions, and transport (Altaratz et al., 2014; Barbaro et al., 2013; Rosenfeld et al., 2014; Wang et al., 2013; Zhang et al., 2015).

A growing body of research has demonstrated a link between aerosol inhalation and health problems (Anderson et al., 2020; I. M. Kennedy, 2007; Mauderly & Chow, 2008; Pöschl, 2005; Schraufnagel, 2020; Shiraiwa et al., 2017). This is because some aerosols are small enough to pass through the nasal passages and into the lungs, as shown in Figure 1, which can potentially lead to cardiovascular complications (Kim et al., 2015).

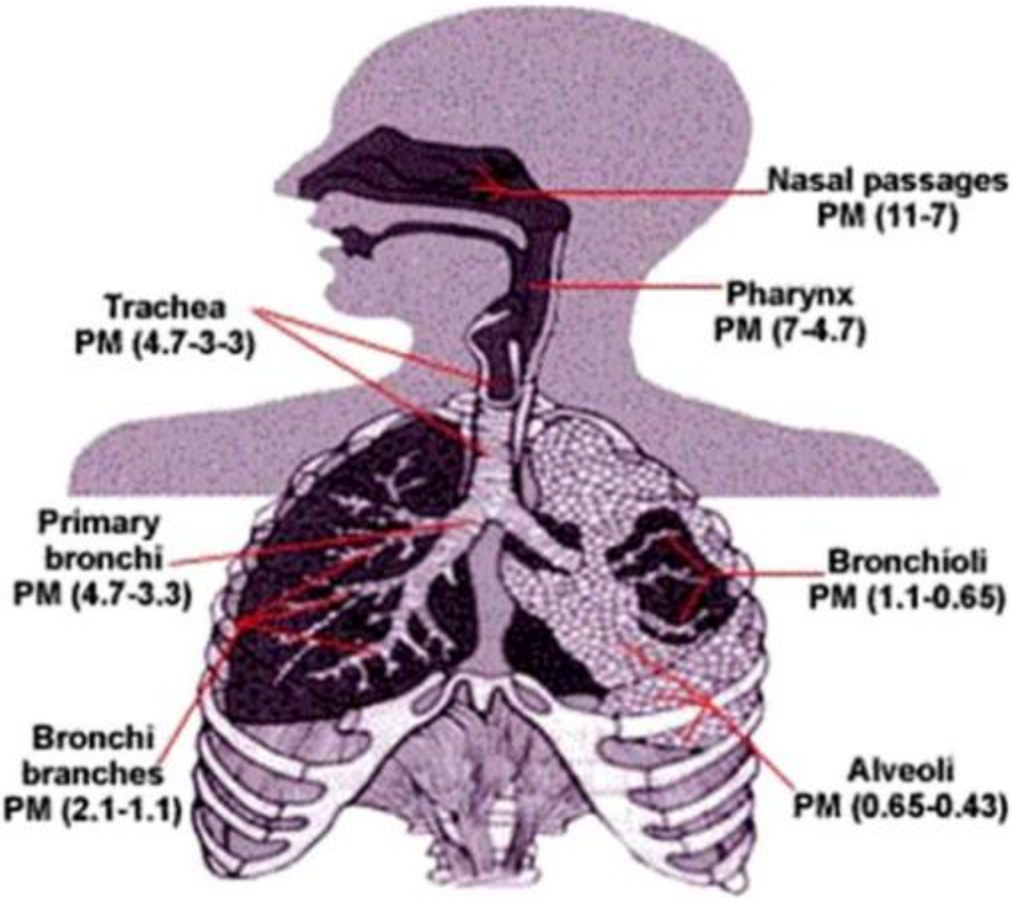

Figure 1: The size of aerosols, or particulate matter (PM), that can penetrate the human body (Kim et al., 2015).

One example of a source of aerosols with high societal impact is wildfires. In North America have been occurring more frequently and are common in the western U.S. during the wildfire season (summer and fall for the northern hemisphere) (Gonzalez-Alonso et al., 2019; Neale & May, 2018; Peterson et al., 2014). An example of a wildfire with the resulting smoke plumes on September 5, 2017, is shown in Figure 2, with the locations of the fire sources indicated using orange dots. These wildfires have devastating effects on the people residing in or around the burn area, destroying homes, businesses, and lives (Kramer et al., 2019). The smoke generated from these fires can be transported upstream and downwind over long distances (i.e., across the entire country), causing damaging health, environmental, and economic damage

(Alman et al., 2016; Lim et al., 2012; Shuman et al., 2022; Wu et al., 2018). Properties of smoke from wildfires vary depending on vegetation type, ecological and geographical conditions, and burn conditions (Chakrabarty et al., 2010; Chen et al., 2007; Chow et al., 2004).

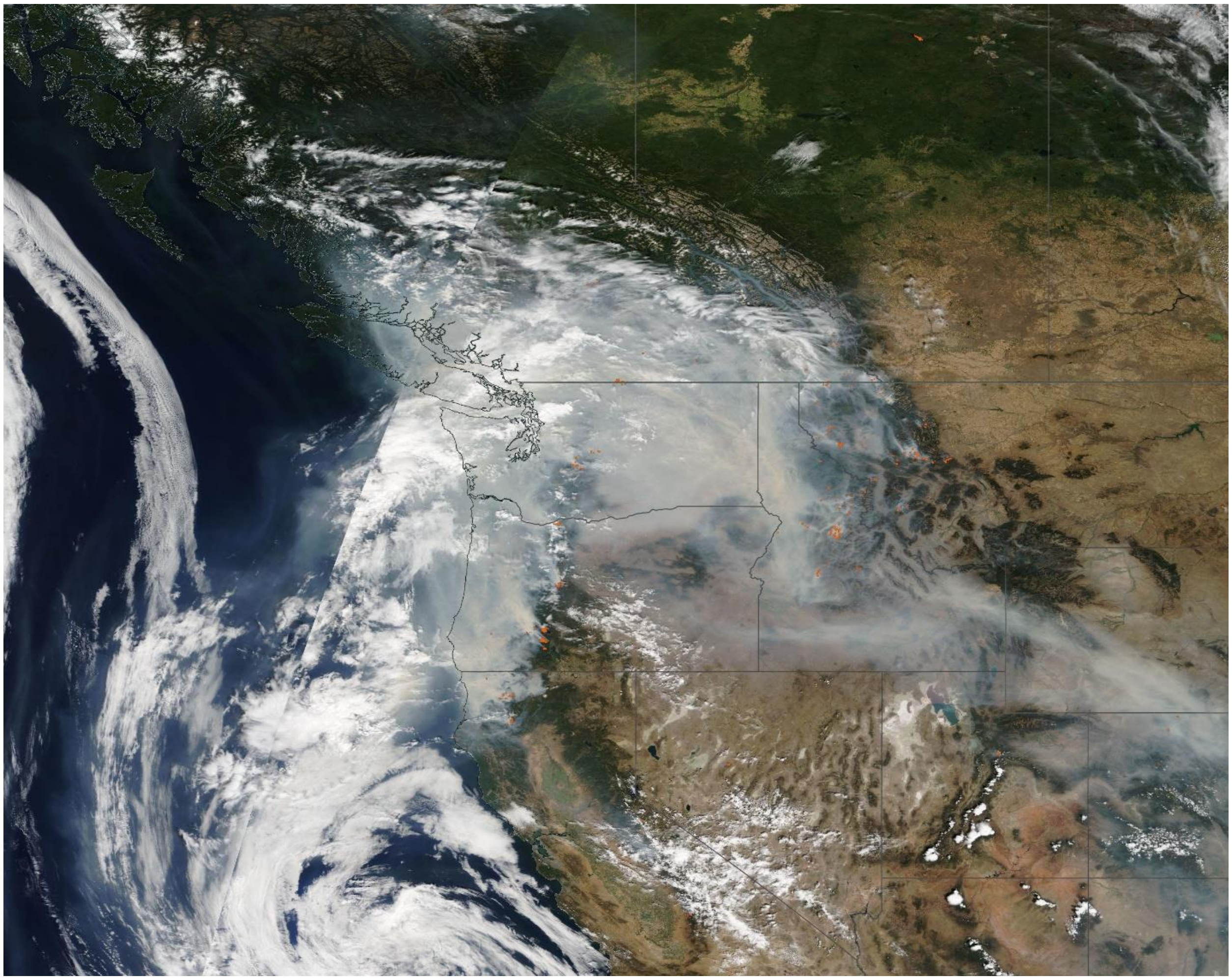

Figure 2: VIIRS visible imagery over the Pacific Northwest on September 5, 2017, showing heavy wildfire smoke. Source: NASA Worldview.

Another condition under which aerosols can have damaging effects is during stable conditions that inhibit pollutant mixing, known as temperature inversions. Temperature inversions are a meteorological phenomenon in which a layer of warm air overlies a layer of colder air in the troposphere, contrasting with the typical temperature gradient, which falls with height in the troposphere (Iacobellis et al., 2009). Temperature inversions can present health risks to those living in the area, particularly in urban cities during the winter. Under these conditions, urban emissions from cars, heating, businesses, etc. (i.e., smog) are trapped near the ground and pose a severe health risk for its inhabitants (Trinh et al., 2019; Wallace et al., 2010).

1.2 Aerosol Health Effects

For aerosol-related health studies, aerosols are referred to as particulate matter (PM) and categorized by their aerodynamic diameter. PM_{10} is defined as PM with an aerodynamic diameter of less than 10 microns, $PM_{2.5}$ is defined as PM with an aerodynamic diameter of less than 2.5 microns, and $PM_{0.1}$ is defined as PM with an aerodynamic diameter of less than 0.1 microns. These delineations were chosen with respect to the ability of various human body parts to filter out particles of these sizes (Figure 1).

There are many different types of PM that can affect human health. For example, wildfire smoke exposure has been shown to have strong associations with general respiratory health problems, overall increased morbidity and increasing evidence of cardiovascular risk. The most common effects include exacerbations of asthma and chronic obstructive pulmonary disease (Reid et al., 2016). Long-term effects also include an increased risk of cancer (Grant & Runkle, 2022). For example, in Indonesia, long-term exposure to $PM_{2.5}$ is estimated to have caused 648 premature deaths per year from 2011 to 2015 (26 deaths per 100,000 people per year) (Uda et al., 2019). In Canada, all-cause mortality due to wildfire $PM_{2.5}$ exposure is estimated to have increased from 570 deaths in 2013 to 2,500 deaths in 2017 before decreasing to 1,400 in 2018 (Matz et al., 2020). Kim et al., (2017) found an average decrease in lung capacity of 20.4 L min^{-1} following a large wildfire in Indonesia in 1997 (Kim et al., 2017). A growing body of research also shows a link between high PM levels and cardiovascular disease, especially from smoke exposure (Brook et al., 2010; Martinelli et al., 2013).

Dust aerosols are another source of concern when it comes to human health risks. In the U.S., dust storms occur frequently in the Great Plains and the Great Basin, where dry climate and high winds form the perfect conditions (Gillette & Hanson, 1989). Specific regions that also experience intense dust effects include the Coachella Valley and Owens Valley in California (Bach et al., 1996); Phoenix, Arizona (Péwé & Science, 1981); and Lubbock, Texas, which has experienced spikes in pneumonia following dust events, a phenomenon dubbed the “Haboob Lung Syndrome” (Panikkath et al., 2013). Many studies have found respiratory clinic visits increased following dust events (Cheng et al., 2008; Chien et al., 2012; Kang et al., 2012; Lee et al., 2013).

1.3 Aerosol Weather and Climate Effects

Aerosols have complex effects on the Earth's radiative balance. These effects are separated into direct and indirect effects. The direct effect is the change in Earth’s radiative budget due to the direct scattering and absorption of incoming shortwave radiation by atmospheric aerosols (Atwater, 1970; Boucher et al., 2013; Charlson & Pilat, 1969; Coakley et al., 1983; McCormick & Ludwig, 1967). The indirect effect is the change in cloud albedo and lifetime that causes affected clouds to reflect more shortwave radiation than they otherwise would (Lohmann & Feichter, 2005; Prather et al., 2008; Ramanathan et al., 2001). Increased aerosols in the atmosphere can lead to an increase in cloud condensation nuclei (CCN), which, assuming the same amount of water vapor in both scenarios, leads to smaller cloud droplets with higher albedo (reflectivity), known as the Twomey effect (Twomey, 1974). Higher albedo causes clouds to reflect more shortwave radiation, and the decrease in cloud droplet effective radius also leads to a decreased chance of rain and an increased cloud lifetime, known as the Albrecht effect (Albrecht, 1989). A brighter, more persistent cloud will again lead to more shortwave radiation reflected and a cooler Earth.

Aerosols also affect extreme weather like tropical cyclones (TC). Rosenfeld et al., (2012) noted how CCN absorbed by a TC can significantly weaken its intensity. This mechanism is accomplished through "redistributing the precipitation and latent heating to more vigorous convection in the storm periphery that cools the low levels and interferes with the inflow of energy to the eyewall, making the eye larger and the maximum winds weaker". This effect was observed through both simulations and observations (Khain et al., 2008; Koren et al., 2005, 2010; Wang, 2005).

The impact of aerosols on vertical heating profiles is not just limited to extreme weather events. For instance, aerosols have been shown to impact the height of the planetary boundary layer (PBL) (Zhang et al., 2012). Quan et al., (2012) used various remote sensing instruments to measure the vertical profile of aerosols and the PBL height. They found that aerosol concentration and PBL height were anticorrelated during clear and haze conditions. The results from this study showed that the average maximum PBL heights were approximately 1.08 and 1.7 km while the averaged aerosol concentrations were 52 and 17 μgm^{-3} under haze and clear sky conditions, respectively. A positive feedback loop in which haze conditions (i.e., high aerosol concentrations), significantly reduced heat flux and repressed the PBL, which in turn usually led to higher aerosol concentrations. This phenomenon would lead to high concentrations of surface air pollution, particularly in megacities where urban emissions are high on a daily basis due to traffic and factories.

1.4 Sensing Aerosols

There are numerous techniques to monitor the aerosols in our atmosphere. However, this research focused on two types of measurements: ground-based-instrument networks and satellite remote sensings. We gathered retrievals of aerosol optical depth (AOD) from these two sources.

1.4.1 Aerosol Optical Depth (AOD)

AOD is directly related to the transmittance t at wavelength λ through a column of the atmosphere between points s_1 and s_2, where transmittance is simply the ratio of outgoing $\mathrm{I}_\lambda(s_2)$ and incident $\mathrm{I}_\lambda(\mathrm{s}_1)$ radiant intensity (Petty, 2006)

$$t_\lambda(s_1, s_2) = \frac{I_\lambda(s_2)}{I_\lambda(s_1)} = e^{-\tau_\lambda(s_1,s_2)}. \qquad \text{(Eq. 1)}$$

AOD quantifies the amount of light at a particular wavelength that is attenuated by aerosols after passing through a column of air and therefore acts as a proxy measurement of the quantity of aerosols within that column.

1.4.2 Retrieving AOD Using Ground-Based Instruments

One method of retrieving AOD from the ground is through the use of sunphotometers, devices that can measure irradiance at different wavelengths while pointing at the sun. While the solar irradiance at the top of the atmosphere $I_\lambda(s_1)$ cannot be directly measured by these instruments because they are ground-based, it can be inferred using a technique called Langley extrapolation. Langley extrapolation assumes a plane-parallel atmosphere with minimal temporal and spatial disturbances between two or more sequential observations (e.g. throughout the morning). These measurements at different solar zenith angles allow the user to extrapolate the solar irradiance at the top of the atmosphere. Once this value is obtained, AOD can be calculated using Eq. 1 (Shaw, 1983). These AOD retrievals are point-based, only measuring the column of atmosphere above the instrument.

The wavelengths at which these retrievals are made are chosen to avoid the molecular atmospheric absorption bands. Doing so ensures the optical depth retrieved results from aerosol-induced extinction and not atmospheric gases (e.g., O_2, O_3, N_2, and CO_2) and water vapor (Angstrom, 1929; Shaw, 1983).

1.4.3 Retrieving AOD Using Satellites

Another method of retrieving AOD is through radiometers mounted on satellites. Radiometers are similar to photometers, but instead measure radiant flux. These measurements also enable a retrieval of AOD but through a more complicated retrieval algorithm due to the added complexity from light reflecting off the Earth and its atmosphere. This method allows for area-averaged retrievals of AOD and has become the primary source for monitoring loadings of aerosols at the global scale (Wei et al., 2020). Satellite-retrieved AOD datasets are complementary to sunphotometer AOD datasets, as the former can cover areas that are inaccessible for sunphotometer stations and the latter can be used to calibrate the retrieval algorithms of the former (Wei et al., 2020).

Examples of such radiometers include the Moderate Resolution Imaging Spectroradiometer (MODIS) onboard Terra and Aqua satellites and the Visible Infrared Imaging

Radiometer Suite (VIIRS) onboard Suomi-NPP satellite. The three major retrieval algorithms used for retrieving AOD are Dark Target (DT) (Chu et al., 2002; Kaufman et al., 1997; Remer et al., 2002, 2005; Tanré et al., 1997a, 1999), Deep Blue (DB) (Hsu et al., 2004, 2013; Jeong et al., 2011; Sayer et al., 2013; Sayer, Hsu, Bettenhausen, Ahmad, et al., 2012; Sayer, Hsu, Bettenhausen, Jeong, et al., 2012), and Multi-Angle Implementation of Atmospheric Correction (MAIAC) (Lyapustin et al., 2012, 2018; Lyapustin, Martonchik, et al., 2011; Lyapustin, Wang, et al., 2011; Lyapustin & Wang, 2008).

DT is the original aerosol retrieval algorithm developed for the MODIS sensors and was implemented on board Terra and Aqua (Kaufman et al., 1997). The algorithm works by organizing radiance data into 10km boxes, removing distortion effects (e.g., gas absorption, angular effects) from the satellite signal, establishing a land/ocean mask, separating the aerosol signal from noise (clouds, surface inhomogeneities, etc.), inferring the AOD from the cloud-free signal. As the name suggests, the DT algorithm is intended for use over dark surfaces, mainly vegetated land areas and the ocean. This is an issue for areas like the western U.S., which has many more bright surface regions such as deserts, snowpack, and salt pans compared to the eastern U.S. (Loría-Salazar et al., 2016) .The DT algorithm was not selected as the primary AOD training input of the gap-filled AOD model because the challenges presented accurately retrieve AOD on bright surfaces.

Hsu et al., (2004) introduced the DB algorithm (Hsu, 2017; Hsu et al., 2004, 2019; Hsu, 2016). This algorithm retrieved AOD over bright surfaces by leveraging blue spectral wavelengths (<500 nm) which render these surfaces much darker than in the red and infrared part of the spectrum, but much darker in the blue spectral region (<500 nm) (Figure 3). The DB algorithm also includes a high-resolution surface reflectance lookup table and can retrieve AOD over bright surfaces with higher accuracy than the DT algorithm (Hsu et al., 2004). This is useful in providing a more spatially complete distribution of AOD and informing of different forms of aerosols and their emission sources. Namely, for desert areas, this means dust aerosols.

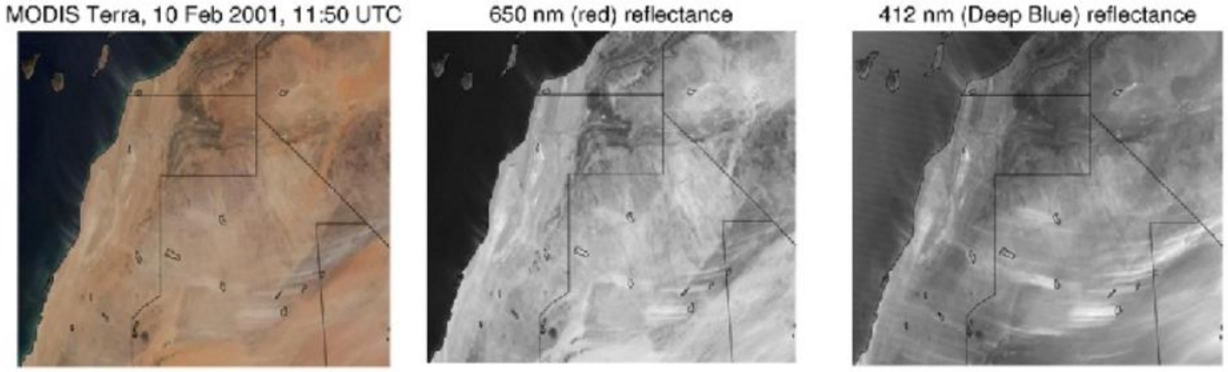

Figure 3: An image over northern Africa showing the presence of dust aerosols and the effectiveness of the Deep Blue wavelength at isolating these aerosols. Source: MODIS Terra.

MAIAC is another AOD retrieval algorithm developed by Alexei Lyapustin and Yujie Wang in 2008 (Lyapustin & Wang, 2008) that takes an entirely different approach for cloud masking, aerosol retrievals, and atmospheric correction, considering previous retrievals and treating these parameters as time series. It also has an extremely high resolution of 1 km, compared to the 3, 6, and 10 km resolution of DT and DB. Unfortunately, MAIAC Collection 6, which was the latest collection at the time of the gap-filled model input selection, was not chosen as MAIAC AOD underestimated high AOD loadings such as smoke in comparison to ground-truth AERONET AOD (Loría-Salazar et al., 2021).

1.5 Estimating $PM_{2.5}$ Using AOD

Satellite retrievals of AOD have commonly been used as a spatial predictor of daily surface-level $PM_{2.5}$. These attempts have included statistical, chemical transport, and machine learning models (Chang et al., 2014; Chen et al., 2021; Holmes et al., 2014; Hu et al., 2014, 2014; Hu et al., 2014). These studies have shown promising results. For instance, Hu et al., (2014) showed that it is possible to estimate surface $PM_{2.5}$ at 1 km resolution in the southeastern United States using MAIAC AOD as the primary estimator and a two-stage model that combines a linear effects model with a geographically weighted regression. They achieved an R^2~0.67, a mean prediction error MPE~2.54 μgm^{-3}, and a root mean squared prediction error RMSE~3.88 μgm^{-3} in cross-validation. However, Hu et al., (2014) noted, “a limitation of the developed approach is the lack of a method to fill the gaps in areas where AOD is not retrieved”.

Chang et al., (2014) introduced a statistical downsampling method that he used to evaluate MODIS DT AOD as a predictor for daily surface $PM_{2.5}$. Chang’s study was conducted over a region in the southeastern United States from 2003-2005. Approximately 43% of the data from AOD grid cells collocated with $PM_{2.5}$ monitoring stations was missing in this study. Various permutations of datasets were tested as inputs, namely MODIS AOD, land use variables like elevation and proximity to roads, and daily meteorological variables. Using 10-fold cross-validation, Chang et al., (2014) found that the model trained on all inputs and the model trained on just meteorological variables and AOD performed the best, both achieving an RMSE~3.71 μgm^{-3} and an R^2~0.71. The lowest performing model was trained on just meteorological and land use variables, with an RMSE~ 4.03 μgm^{-3} and an R^2~0.72. Chang et al., (2014) concluded that satellite AOD can be considered a covariate for predicting daily surface $PM_{2.5}$ concentrations, but also notes that a significant caveat lies in dealing with missing AOD data, where $PM_{2.5}$ predictions cannot be obtained.

Chen et al., (2021) compared the performance of AOD-based and non-AOD-based daily surface $PM_{2.5}$ random forest models. The region of interest in this study was Guangdong, China, where satellite AOD data was extremely sparse (around 80% missing) over the study period of 2016-2018. The results showed that the random forest trained with AOD as a predictor barely outperformed the random forest trained without it. In cross-validation for the three years, the AOD model scored R^2 values of ~0.83, ~0.82, and ~0.80, while the non-AOD model scored ~0.82, ~0.81, and ~0.78. Similarly, for RMSE, the AOD model scored 8.2 μgm^{-3}, 9.2 μgm^{-3}, and 9.0 μgm^{-3}, and the non-AOD model scored 8.4 μgm^{-3}, 9.5 μgm^{-3}, 9.4 μgm^{-3}. This suggests that in developing models that estimate daily surface-level $PM_{2.5}$, sparsity can render current remote-sensed AOD data useless as a predictor.

1.6 AOD Gap-Filling Problem

The previous section suggests that AOD could be valuable as a predictor for surface $PM_{2.5}$. However, the issue of data sparsity must first be overcome. This sparsity exists because AOD cannot be retrieved from satellite sensors when the underlying image is too reflective (Loría-Salazar et al., 2016; Loría-Salazar et al., 2021; Lyapustin et al., 2018; Peterson et al., 2014, 2017) or when the aerosols are concentrated too close to the surface (Silcox et al., 2012). These reflective surfaces include clouds, snow cover, salt pans, and deserts. One approach to dealing with these gaps is filling them in before using satellite AOD datasets in downstream tasks. **This approach, commonly referred to as AOD gap-filling, will be this thesis's primary focus.**

AOD gap-filling has been the focus of many efforts which have yielded varying degrees of success. Zhang et al., (2022) summarized these as data fusion-based (linear regression, Bayesian, Bayesian maximum entropy, spatial statistical data fusion, maximum likelihood estimate, quantile regression (QR), QR neural network, QR forest, and tensor completion), collaborative variable-based (random forest, multiple imputation, and normalized difference vegetation index-based weighted linear regression), interpolation-based (inverse distance weighting and Kriging variations), and hybrid (linear regression and ordinary Kriging, spatial-temporal hybrid fusion considering aerosol variation mitigation, and two-step model) (Zhang et al., 2022). An important point to note is that most of these models have been developed for China, leaving the U.S. with no specialized solutions.

The Bayesian method used by Singh et al., (2017) focused on combining multiple AOD datasets collected from various instruments along with error distributions from AERONET in order to form a more complete and accurate daily "merged" AOD dataset in India (Singh et al., 2017). This method results in high correlation r~0.89-0.93 and RMSE~0.08-0.13 with AERONET compared to the original input datasets. Unfortunately, gaps still exist in the merged dataset where no retrievals in the input datasets exist. Tang et al., (2016) introduced a Bayesian maximum entropy approach to create a more spatially complete daily AOD dataset over China. This method also merges various satellite AOD products and is able to produce a dataset that is 95.2% complete compared to MODIS (22.9%) and Sea-viewing Wide Field-of-view Sensor (20.2%). A correlation of r~0.75 and RMSE~0.29 was achieved against AERONET. Jiang et al., (2021) developed a random forest model to produce hourly gap-filled AOD datasets for China using MAIAC and AHI (Advanced Himawari Imager, a satellite-mounted multispectral imager also capable of retrieving AOD) AOD as well as meteorological and geographical variables. They achieved a correlation of r~0.60 and RMSE~0.20. Bai et al. (2022) created a long-term (2000-2020) high-resolution (1 km) gap-free AOD dataset over China using tensor completion. They used a diverse dataset composed of AOD from multiple instruments/models, meteorology, air quality measurements, and land-use variables into a tensor-flow-based data fusion method.

Their method results in a correlation of r~0.91 and RMSE~0.21. Wei et al., (2023) developed a 4-Dimensional Space-Time Extra-Trees (4D-STET) model which they used to generate global daily gap-filled AOD datasets using Terra and Aqua MAIAC AOD retrievals along with modeled AOD, meteorological variables, and land-use variables. Their method achieved a correlation of r~0.73 with AERONET.

The core research question of the work presented in this thesis is **whether we can develop an AOD gap-filling model specialized for the continental U.S. (CONUS) using Deep Learning (DL) techniques**. DL is a subfield within machine learning (ML), which studies computer systems that can automatically improve through experience (Jordan & Mitchell, 2015). Methods within DL are characterized by the use of extremely large, or “deep”, artificial neural networks (ANN), hence the name (LeCun et al., 2015). ANN’s are computer models loosely inspired by the network of neurons within the brain (Zou et al., 2009). These neurons can receive signals from other neurons via chemicals called neurotransmitters and will send neurotransmitters upon receiving the right neurotransmitters. As demonstrated by the intelligence of complex life-forms, a sufficiently large collection of neurons is capable of many non-trivial tasks. In much the same way, neurons in an ANN send and receive signals to each other depending on the signal they receive. The neurons in an ANN are assembled in a stack of layers, where signals flow sequentially through the layers. These networks are usually trained on samples drawn from the data distribution of interest using some variant of gradient descent, similar to how one would train a regression model.

DL has seen an increase in use in recent years, both in research and commercially. It has been applied to a wide range of tasks, from materials science (Choudhary et al., 2022) to drug discovery (Ma et al., 2015) to natural language processing (Collobert et al., 2011). Perhaps most relevant to this thesis, however, is that researchers have found success in developing DL models for computer vision, the study of machine learning for image-based tasks (Chai et al., 2021). This is promising because any spatial datasets can be viewed as images that can then be fed into such models.

In this thesis, I aim to demonstrate how a DL computer vision model, the UNet 3+, can outperform the more traditional methods mentioned above. The UNet 3+ was initially developed for medical image processing (Huang et al., 2020). It was chosen for this project for a few reasons. One is because its outputs have the same spatial dimensionality as its inputs. This configuration makes for easy operational use as one pass through the input data will produce a spatially complete AOD dataset. Another reason is that it has been empirically shown to converge given a relatively small sample size compared to other state-of-the-art models in computer vision like the vision transformer (Dosovitskiy et al., 2021). This is an important consideration considering the relatively small satellite AOD dataset (one image a day for eleven years). A back-of-the-envelope calculation shows that if a satellite has only been operational for around ten years, then only around 3,650 days of data will be available, a paltry amount compared to the 300 million sample dataset used to train the vision transformer to its best-in-class status. **The goal of this thesis is to test the viability of UNet 3+ as an architecture for AOD gap-filling. Our hypothesis is that the UNet 3+ architecture, combined with DB AOD and supplemental datasets, will improve spatiotemporal coverage of AOD missing data due to cloud cover or bright surfaces (e.g., semi-arid or arid deserts). The primary outcome of this research is the creation of a spatiotemporal dataset of gap-filled AOD with an eye towards of using the end product (gap-filled AOD) as a spatial predictor of $PM_{2.5}$ in future poor air quality exposure studies.**

2. Instrumentation and Datasets

Table 1 summarizes the instruments, models, variables, and spatial-temporal resolution of the datasets used in this investigation. This section will explain the instrumentation, models, and datasets used to train the UNet 3+ gap-filler AOD.

Model/Instrument	**Variables**	**Spatial Resolution**	**Temporal Resolution**
NAM 12km	P_{surf}, P_{msl}, T_{surf}, T_{2m}, DWT_{2m}, orography, RH_{2m}, U_{10m}, V_{10m}, land mask, vegetation, PBLH	12x12 km^2	6 hours
Terra/Aqua MODIS	DB AOD_{550nm}	10x10 km^2	daily
	FRP	1x1 km^2	daily
S-NPP VIIRS	DB AOD_{550nm}, aerosol type, algorithm flag	6x6 km^2	daily
MERRA-2	AOD_{550nm}	50x62.5 km^2	hourly
HMS	smoke density	~1x1 km^2	daily
AERONET	AOD_{675nm}, $AEE_{440\text{-}675nm}$	point location	<1 hour

Table 1: List of all data sources used in the training and evaluation of the UNet 3+ gap-filler.

2.1 Instrumentation

2.1.1 Cimel CE-318

The Cimel CE-318 is a photometer that can take both daytime and nighttime iradiance measurements. These devices can measure direct (Direct-Sun scan) and sky (Almucantar and Principal Plane scans) irradiance using the sun as the light source under non-cloudy conditions (Dubovik et al., 2000; Dubovik et al., 2006; Holben et al., 1998). NASA operates a global network of Cimel-CE 318 sunphotometers called the AErosol RObotic NETwork (AERONET).

2.1.2 Terra/Aqua Moderate Resolution Imaging Spectroradiometer (MODIS)

Terra and Aqua are two NASA polar-orbiting satellites equipped with the twin instruments Moderate Resolution Imaging Spectroradiometer (MODIS), sensors used for earth and climate measurements. Terra was launched in 1999, and Aqua was launched shortly after in 2002. MODIS was designed to resolve the spatial and temporal distribution of aerosols, clouds, and other weather features (e.g., albedo) globally. It has 36 spectral channels ranging from 0.41 μm to 15 μm at spatial resolutions ranging from 250 m to 1 km at nadir, allowing for aerosol and cloud characterization (Ackerman et al., 1998; Gao et al., 2002; Martins et al., 2002). MODIS's spectral range enabled it to make AOD retrievals (Tanré et al., 1996, 1997b).

Polar-orbiting satellites capture images of each part of the Earth at the same local time daily (Figure 4). Since their orbit is relatively closer to the Earth than geostationary satellites, polar-orbiting satellites can acquire daily high-resolution images. Terra has a morning overpass time of 10:30 a.m., and Aqua has an afternoon overpass time of 1:30 p.m.

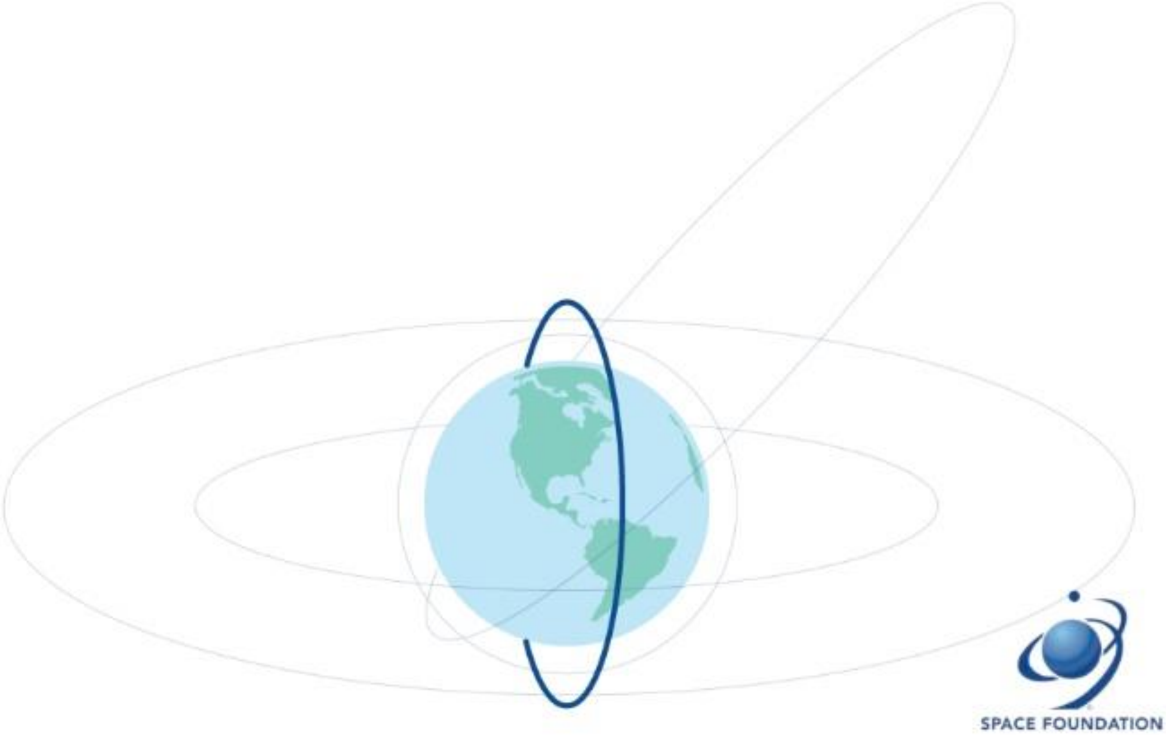

Figure 4: Depiction of the orbit of a polar-orbiting satellite. Source: Space Foundation.

2.1.3 Suomi-NPP Visible Infrared Imaging Radiometer Suite (VIIRS)

The Suomi National Polar-orbiting Partnership (S-NPP) is a weather satellite launched in partnership between NASA and NOAA in 2011. S-NPP carries an instrument suite called the Visible Infrared Imaging Radiometer Suite (VIIRS), which combines features of MODIS with those of the Advanced Very High-Resolution Radiometer (AVHRR) (the previous generation of radiometer equipped with NOAA polar-orbiting satellites) and the Operation Linescan System (an instrument used by the Defense Meteorological Satellite Program (DMSP) to monitor the global distribution of clouds and cloud top temperatures) (“DMSP Operational Linescan System

(OLS),” 2022; Liang, 2018) . VIIRS has 22 spectral channels, ranging from 0.412 μm to 11.45 μm, at 375 m and 750 m spatial resolution at nadir (Liang, 2018). VIIRS also has a swath width of 3,000 km, which is 700 km greater than that of MODIS and eliminates gaps in coverage by S-NPP (Hillger et al., 2014). S-NPP also has the same 1:30 p.m. overpass time as the A-train (constellation of satellites that overpass 1:30 p.m. such as Aqua satellite) (Stephens et al., 2002).

2.2 Datasets

2.2.1 Weather Variables from North American Model (NAM) 12 km

We incorporated weather data from the North American Mesoscale Forecast System (NAM) 12 km (Janjic et al., 2005) to provide the gap-filled model the building blocks for deciphering the nature of aerosol transport and evolution. The NAM datasets were selected because 1) a weather model would offer spatially complete data on a regularly spaced grid, unlike data collected from balloon soundings or weather stations, and 2) a mesoscale model would be gridded at roughly the resolution as that of satellite products.

The NAM 12 km model is a regional weather forecast model developed by the National Oceanic and Atmospheric Administration (NOAA) and run by the National Centers for Environmental Prediction (NCEP) (Information (NCEI), n.d.). It offers weather parameters, including pressure, temperature, humidity, planetary boundary layer height (PBLH), and land-use parameters such as orography and vegetation. NAM uses a Lambert Conformal grid spaced at 12x12 km^2 that covers CONUS (Figure 5), is run four times daily at 00:00, 06:00, 12:00, and 18:00 UTC, and has been operational since March 2004 (Information (NCEI), n.d.).

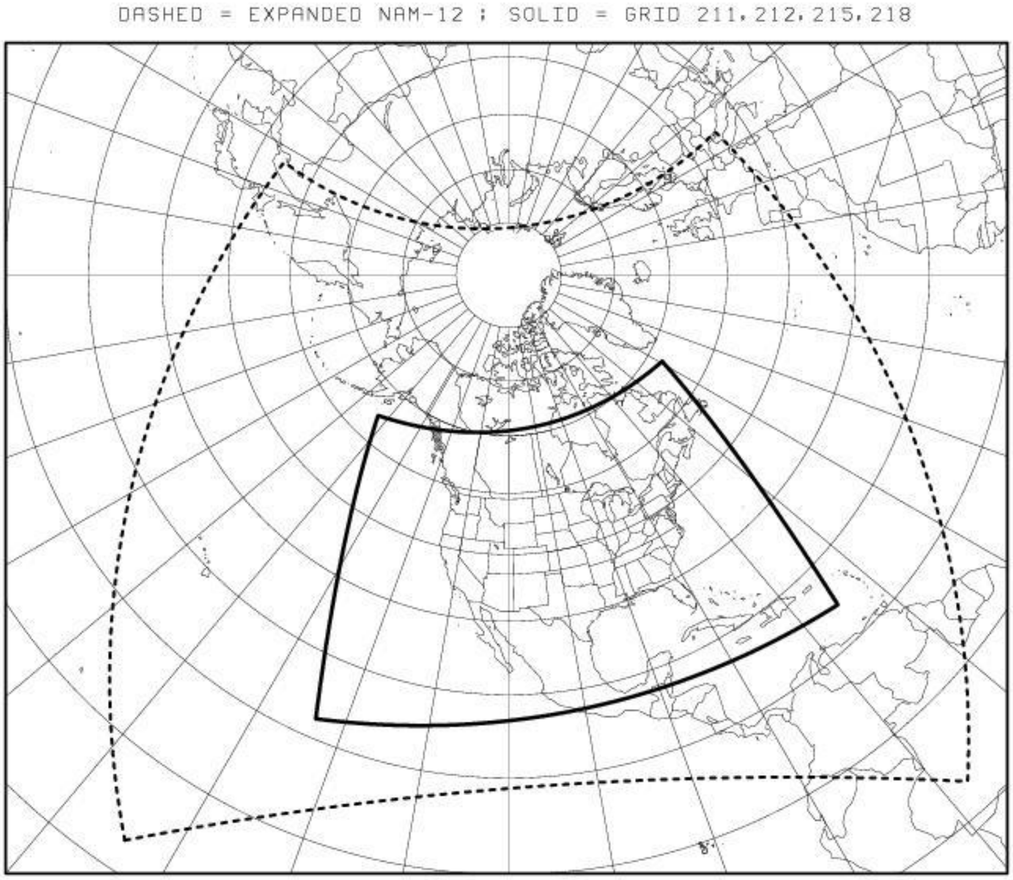

Figure 5: A map of the NAM domain (solid) as well as the parent domain (dashed). We used data gridded in the solid region. Source: ("North American Mesoscale Forecast System," 2020).

We selected surface pressure (P_0), pressure at mean sea level (MSLP), surface temperature (T_0), 2-meter temperature (2-T), 2-meter dewpoint temperature (2-DewT), 2-meter relative humidity (2-RH), 10-meter wind speeds (10-WS), PBLH, orography, and vegetation. We aligned the time of the NAM variables with satellite overpasses and averaged consecutive 18:00 and 00:00 UTC datasets, which range from 1:00 p.m. PST to 4:00 p.m. EST.

2.2.3 Hazard Mapping System (HMS)

We also included NOAA’s Hazard Mapping System (HMS) smoke polygons product further to facilitate the model's decision-making 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 observations, passes them through automated fire and smoke plume detection and is finalized by quality-controlled by expert image analysts (Brey et al., 2018; McNamara et al., 2004). HMS runs 24x7x365 and produces daily spatial datasets of smoke plumes over CONUS. These plumes are given as shape files and are categorized as light, medium, and heavy depending on the apparent opacity of the smoke in the satellite imagery.

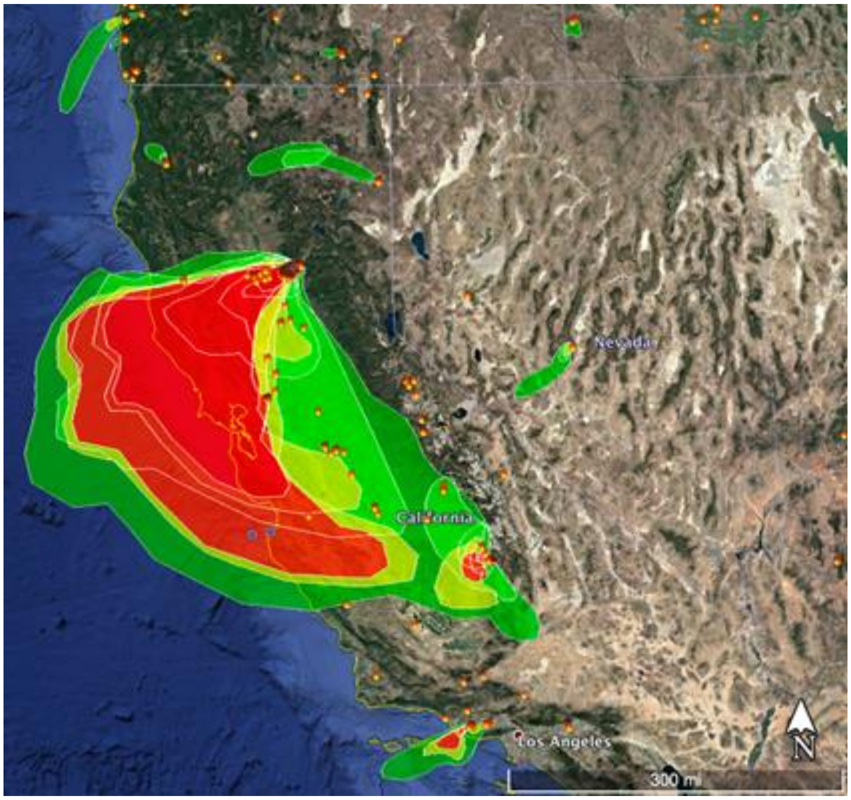

Figure 6: A sample map of the HMS smoke product for November 8, 2018, showing a large smoke plume from the Camp Fire. Green, yellow, and red denote light, medium, and heavy smoke densities, respectively. Source: NOAA HMS.

2.2.2 Modern-Era Retrospective Analysis for Research and Applications, version 2 (MERRA-2)

The Modern-Era Retrospective Analysis for Research and Applications, version 2 (MERRA-2), is a global reanalysis product developed by NASA’s Global Modeling and Assimilation Office (GMAO). 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). It has been noted that such data assimilation has prospects for air quality studies(Bocquet et al., 2015). Most of MERRA-2’s aerosol assimilation data comes from MODIS Terra and Aqua, with additional inputs from AERONET.

From the MERRA-2 dataset, an hourly time-averaged single-level assimilation aerosol diagnostics dataset, we used the reanalysis AOD at 550 nm product that is spatially complete and offered at the native resolution of 0.5° x 0.625°.

2.2.4 AErosol Robotic NETwork (AERONET)

AERONET, is a global federation of ground-based remote sensing aerosol networks. Each site contains a Cimel CE-318 Photometer which retrieves AOD hourly. These AOD retrievals are highly accurate but are only point measurements compared to the area-averaged AOD retrievals from satellite instruments and reanalysis. The AOD retrievals from AERONET are highly accurate (Holben et al., 1998), but do not reveal much about the spatial distribution of AOD since they are only 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 dataset. As such, we used AERONET AOD data to evaluate the gap-filling model but not as input.

We use AOD_{675} and $AE_{440-675}$ from AERONET Version 3 Level 2 (quality assurance) between 10 a.m. and 2 p.m. local time. Version 3 offers a number of improvements over Version

2, including automatic cloud screening and instrument anomaly quality controls (Giles et al., 2019) as well as a higher retention rate of high AOD values (Bhattacharjee et al., 2023).

2.2.5 Satellite Remote Sensing of AOD

Collection 6.1 (C6.1) was used for Terra and Aqua. DB was chosen over DT due to its increased accuracy over bright, reflective surfaces, common in the western U.S., and better characterization of AOD during smoke (Hsu et al., 2004; Loría-Salazar et al., 2021). Collection 6.1 (C6.1) offers many improvements over 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 (Wei et al., 2019). We used the DB AOD at 550 nm from C6.1, which is given at a resolution of 10 km x 10 km^2.

For VIIRS, the Version 1 dataset was used. This collection was chosen over the more recently released Version 2 because Version 1 DB AOD is algorithmically identical to the Terra and Aqua C6.1 DB AOD (Lee et al., 2016). This allows for a simple average to be taken over the data collected from the three satellites, which will be used as a training target and discussed later. In addition to the AOD retrievals, the model was given three supplemental aerosol variables collected by VIIRS. The DB AOD data from Version 1 is given at a 6 x 6 km^2 resolution. The first is aerosol type, a categorical variable with nine levels: no retrieval, dust, smoke, high altitude smoke, pyrocumulonimbus 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 model with further information to track aerosol transport and evolution.

2.2.6 Fire Radiative Power (FRP)

Fire radiative power (FRP) retrieved from MODIS Terra and Aqua (Giglio et al., 2016) was used to train the model. FRP is the radiative power from wild and prescribed fires, it is given in megawatts (MW) and can be used to characterize many fire properties (Archibald et al., 2013; Ichoku & Kaufman, 2005; Kaufman et al., 1998; Roy & Kumar, 2017; Wooster et al., 2003; Wooster & Zhang, 2004). It is retrieved using the brightness temperatures derived from the MODIS 4 µm and 11 µm channels (Giglio et al., 2003, 2016; Kaufman et al., 1998). By including FRP, the model can be trained with a notion of fire source regions, which are typically just a small subset of the total area occluded by wildfire smoke. This information can help to inform the model on the spread of the fire/smoke and the resulting AOD in the surrounding regions.

3. Methods

3.1 Model Selection

To estimate AOD, a model must be aware of complex, nonlinear interactions between weather, emissions, and aerosol types. DL is particularly well-equipped for dealing with these types of problems because DL models contain an extremely large number of parameters and non-linearities. Furthermore, while it has been shown that neural networks with two layers, suitable activation functions, and sufficient width are universal function approximators (Barron, 1994; Cybenko, 1989; Funahashi, 1989; Hornik et al., 1989), scaling models to be deeper rather than wider is more efficient in terms of total parameter count (Lu et al., 2017). These insights have led to the development of numerous different deep networks designed for a variety of difficult tasks throughout the past decade or so.

The subfield of DL most likely relevant to AOD gap-filling is computer vision. Computer vision concerns the development of algorithms that can perform image-related tasks. These tasks range from simpler ones like colorization and classification (identifying the subject of an image), to more difficult ones like generation (hallucinating new images from the data distribution), inpainting (proposing contents for a missing interior portion of an image), and object detection (localizing and identifying objects within an image) (Goodfellow et al., 2014; Lecun et al., 1998; Redmon et al., 2016; Saharia et al., 2022). For most deep learning models within computer vision, a key component is the convolution. Convolutions are spatially local operators that output a linear combination of pixels within a patch. Neural networks built using convolutions are called convolutional neural networks (CNN’s) (O’Shea & Nash, 2015). CNN’s are a natural fit for computer vision tasks because convolutions are most appropriate for settings that satisfy the locality assumption, i.e., objects (pixels in this case) are only directly influenced by its immediate surroundings (neighboring pixels) (Albawi et al., 2017). Problems within atmospheric science that work with regularly gridded data can be easily casted as computer vision problems, because data points can be interpreted as pixels of an image and different input datasets can be viewed as channels of the image (in the same way red, green, and blue channels compose an RGB image, for instance).

One of the earliest DL computer vision models is the Visual Geometry Group (VGG), originally developed for image classification (Simonyan & Zisserman, 2015). It passes an input image through 16-19 convolutional layers and outputs a probability distribution over the different image classes. For use in AOD gap-filling, VGG can be adapted to output continuous values instead of a distribution. However, since the output of this adapted VGG would be a 1x1 pixel, the operational use of VGG for gap-filling would be inefficient. This is because in gap-filling two neighboring missing pixels, VGG would have to be run over two inputs that are mostly identical. This presents efficiency issues that can be avoided by using a different type of model architecture.

In contrast, another class of CNN’s, the UNet (Ronneberger et al., 2015), seems particularly well-suited to this setting. While they were originally developed for image segmentation, much like VGG, UNets can easily be repurposed to output continuous values instead. However, unlike VGG, UNets have the same spatial dimensions for both its inputs and outputs. 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 colorization/inpainting/uncropping/restoration (Saharia et al., 2022), and meteorology (Appel, 2024; Justin et al., 2023). Given its applicability and track record, we chose to use a UNet for this research.

There have been several UNet variants that have been developed since its inception, such as UNet ++ (Zhou et al., 2018), UNet 3+ (Huang et al., 2020), TransUNet (Chen et al., 2021), and more. We chose the UNet 3+ variant for this research because it had demonstrated state-of-the-art performance in its original medical image processing field, beating out UNet and UNet++ while containing fewer parameters (39.39M and 47.18M compared to 26.97M, respectively). We chose UNet 3+ over attention-based architectures like ViT (Dosovitskiy et al., 2021) and UNet variants with attention components like TransUNet because empirical results suggest that attention-based models require larger training datasets and training time to reach convergence compared to fully convolutional networks like UNet 3+. On the other hand, UNet 3+ has reached state-of-the-art performance on extremely small datasets, for instance on the very small liver segmentation dataset (<1000 samples) from the ISBI LiTS 2017 Challenge (Huang et al., 2020). Since the training dataset we constructed was much closer in size to this compared to the 300 million sample dataset (JFT-300M) used to train the ViT to state-of-the-art performance, UNet 3+ was the clear choice (Dosovitskiy et al., 2021). However, networks with attention components may be accessible in the future once we curate larger datasets.

3.2 Spatial/Temporal Projection

3.2.1 Projection and Grid

Before training the UNet 3+, we regridded all data sources onto a unified input grid. This step is necessary because none of the data sources are collocated. For instance, modeled outputs like those from NAM or MERRA-2 are given in different spatial resolutions and on different mapping projections. The same holds true for satellite-derived products such as AOD and HMS smoke as the satellites orbit over slightly different tracks each day, resulting in different latitude/longitude coordinates for retrievals. Information on spatial and temporal resolution for all datasets used for this research is displayed in Table 1.

We selected an Albers Equal Area (AEA) projection with the standard parallels set to 29.5°N and 45.5°N. This projection minimizes scale distortion over the lower 48 states, which is capped at 1.25 percent according to the Understanding Map Projections guidebook by ESRI (M. Kennedy & Kopp, 2001). The input grid selected contains 360 rows and 515 columns of 12x12 km^2 squares. This setup maximizes area over CONUS. (Figure 7).

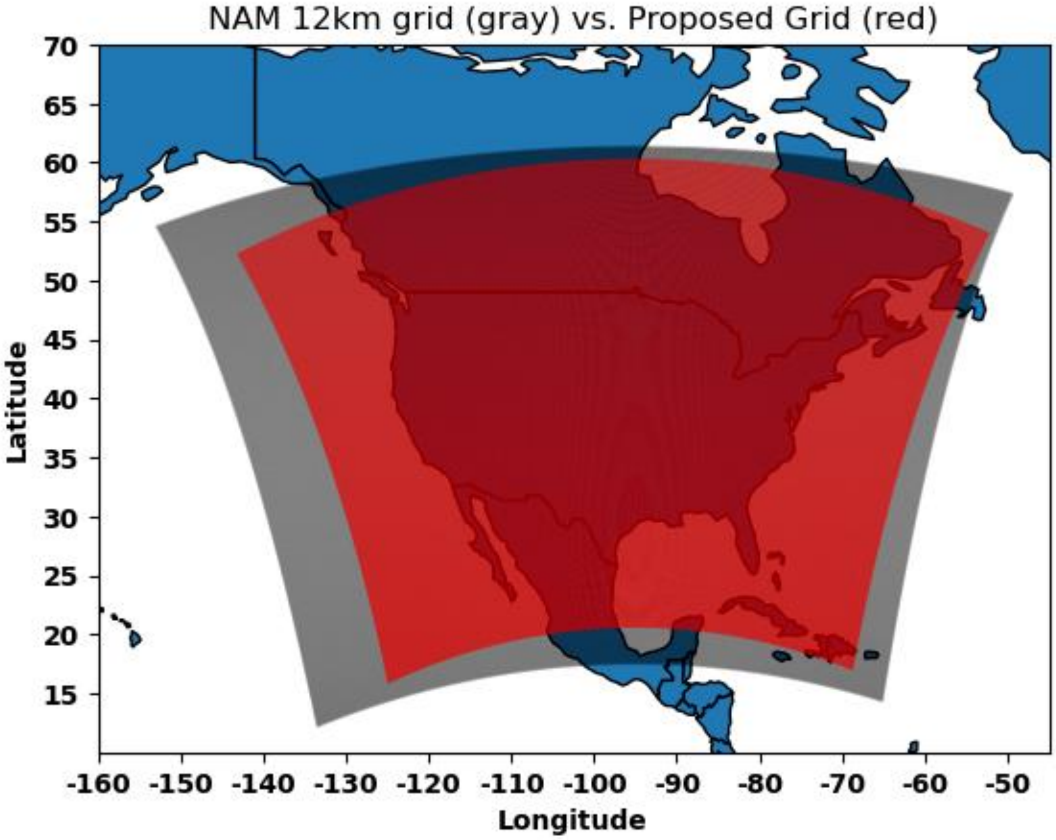

Figure 7: The proposed grid in red and the NAM 12km grid in gray. Both have 12x12 km^2 resolution but are defined on different projections.

3.2.2 Temporal Selection

Much like with the spatial heterogeneity in the input data, the time of day measurements are made is similarly misaligned between the various models and instruments (Table 2).

Source	Temporal Resolution
NAM 12km	0000, 0600, 1200, 1800 UTC daily
MERRA-2	hourly
HMS	daily

Terra	10:30 a.m. local time daily
Aqua/S-NPP	1:30 p.m. local time daily

Table 2: Temporal resolution of various input data.

Since the AOD retrievals from S-NPP, Terra, and Aqua were the primary training targets for our model, we selected from each dataset the daily data points that fell around 10:30 a.m. local time to 1:30 p.m. local time across CONUS. This corresponds to 18:30-21:30 UTC in PST and 15:30-18:30 UTC in EST.

At the start of the project, we worked only with the western U.S. and S-NPP, and thus decided to create a daily NAM dataset by averaging the 18:00 and 00:00 UTC datasets for each day. Since expanding to CONUS, we have not adjusted this averaging to account for the additional time zones and kept the same scheme. Further investigation into model performance and NAM averaging will be done in future work (see Section 4.3).

The 2-dimensional aerosol dataset from MERRA-2 that we included is hourly, but the output selected corresponds to 19:30 UTC. This was again motivated by the preliminary constraints of the western U.S. and S-NPP. HMS data is already in a daily format and focused on fires and smoke over CONUS, so no additional modifications were needed.

To create a singular DB AOD target for the model to train against, we averaged DB AOD retrievals from S-NPP, Terra, and Aqua. In this case, pixels with no retrieval for a particular satellite were not included in the averaging, and pixels with no retrievals from any satellites were assigned NaN. The goal of creating this DB AOD composite was to provide a more coherent and cohesive signal to the model, instead of three separate lossier signals.

3.2.3 Regridding

Regridding was performed by first creating buffers around the data points of interest. These buffers were sized roughly following the documented resolutions of their respective data. Next, the buffered data was overlaid on the grid, and the intersections between the two were calculated (Figure 9). Finally, the intersections were grouped by their respective grid cells and their relative areas were compared to determine the weights. For continuous data variables, the value assigned to a grid cell after regridding is equal to the weighted average of all contributing intersections. The value associated with the intersection with the highest weight is chosen for categorical data variables. To illustrate this, Figure 8 shows a particular grid cell along with all of the NAM data buffers that intersect it. The two NAM buffers on the right would have high weights since they have significant intersections with the grid cell. However, the two buffers on the left would have much smaller contributions. The value in the lower right buffer would be chosen for categorical variables since it has the largest intersection.

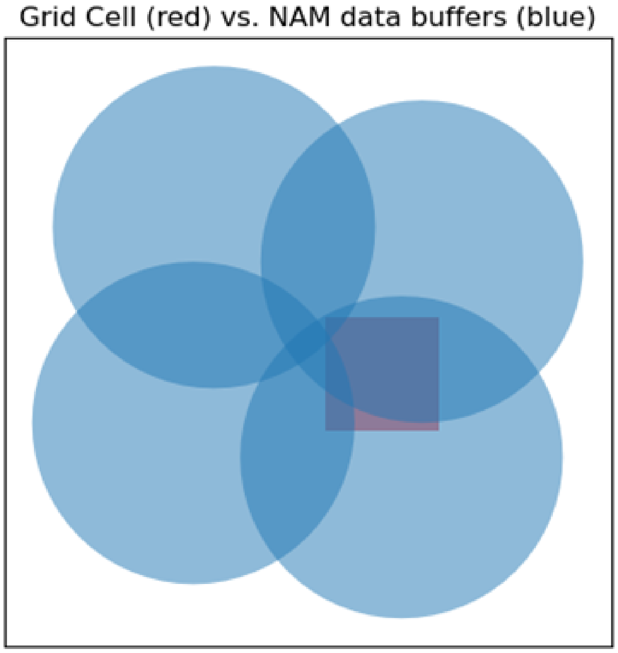

Figure 8: An example of how NAM data buffers (blue) may overlie a grid cell (red).

Expressed mathematically, if the intersections between the top left, top right, bottom left, and bottom right data buffers with the grid cell were a_1, a_2, a_3, and a_4, respectively, the value for a continuous NAM variable x assigned to the grid cell would be

$$\sum_{\mathrm{i}=1}^{4} \frac{\mathrm{a_i x_i}}{\mathrm{a_{total}}} \qquad \text{(Eq. 2)}$$

where x_i is the value of that variable within data buffer i and a_{total} is the sum of all a_i. The value for a categorical NAM variable c assigned to the grid cell would be

$$\mathrm{c}_{\mathrm{argmax(a_i)}} = \mathrm{c}_2 \qquad \text{(Eq. 3)}$$

where c_i is the value of the c within data buffer i.

3.3 Data Preprocessing

3.3.1 Log Transform

A significant challenge with training a model directly on AOD data is the highly imbalanced nature of the dataset. (Thankfully,) the vast majority of AOD samples collected from AERONET and satellites are heavily skewed towards the lower values (AOD < 0.1). However, from a modeling perspective, this presents a very severe class imbalance problem. The class imbalance problem is where the training dataset contains more of a particular sample type than others. This can be a problem even if this skewed training distribution is representative of the "true" data distribution. In working with a model trained with some form of gradient descent and mean squared error (MSE) loss, most gradient descent steps will optimize for the more frequent sample classes and can end up "drowning out" the infrequent sample classes.

Preliminary testing with training the gap-filling model directly on AOD retrievals resulted in a model that only output low AOD values. One method to mitigate this issue is to add greater distinctions between the different types of samples by taking the log transform of the data before feeding it into the model. The log transform, in this case, widens the distribution in the low-end where most of the samples are concentrated, giving the model a chance to learn a more dynamic range of outputs (Figure 9).

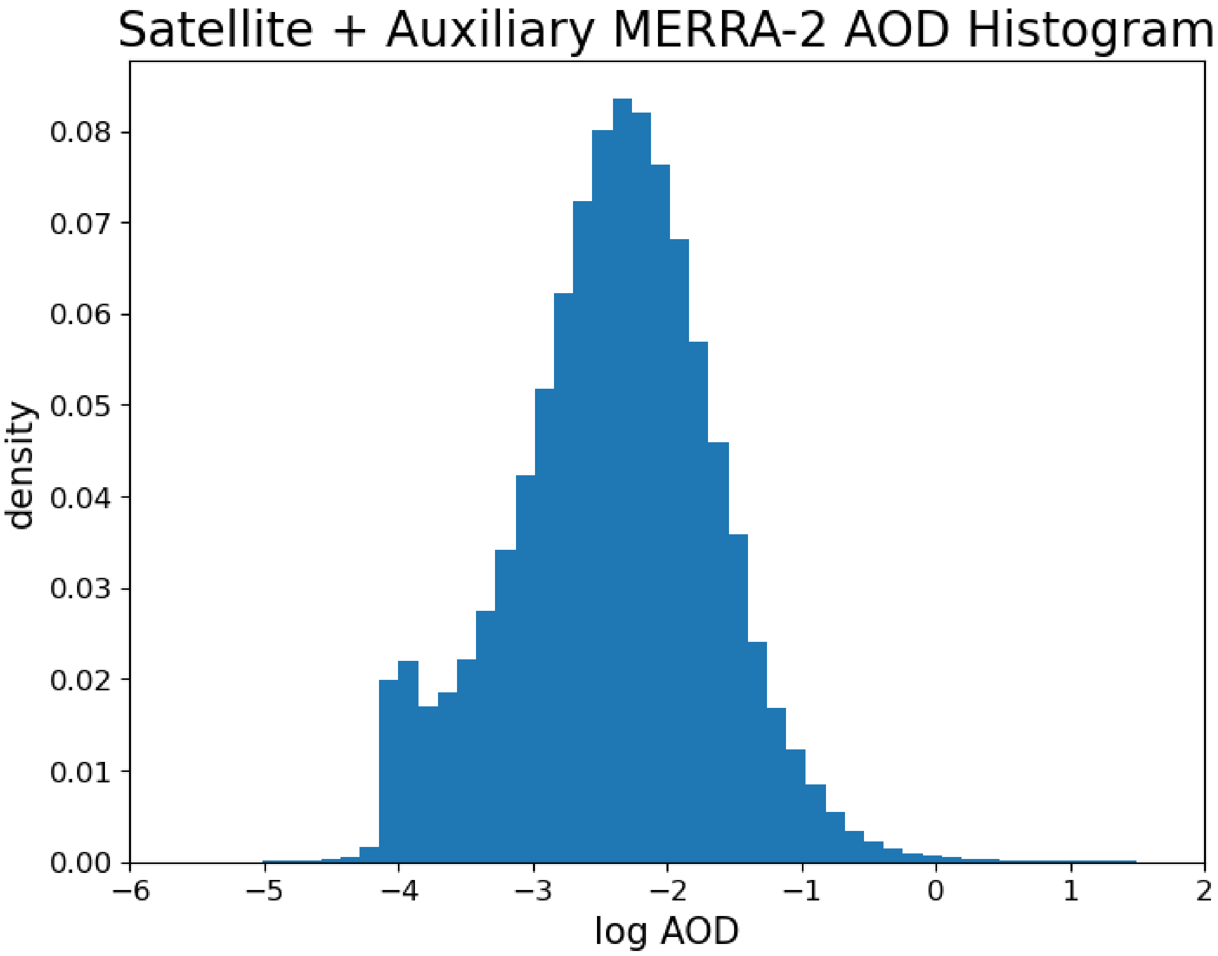

Figure 9: A histogram of all log AOD values from satellite and some auxiliary MERRA-2 values (discussed in section 3.3.5).

Hence, all AOD data that is given to the model (i.e., satellite retrievals and MERRA-2 AOD) are put through a log transform first.

3.3.2 Normalization

Another common challenge with training machine learning models on a diverse set of input data is sometimes different elements in the data have different expressive ranges, which can lead to convergence issues because it is difficult to accurately calculate the gradient at each step of the gradient descent algorithm if all the inputs going into the model have different scales. Therefore, a necessary preprocessing step is the normalization of the dataset prior to inputting it into the model (Jo, 2019; Singh & Singh, 2020). In our case, we chose to normalize to a range of 0 to 1 (Ali & Faraj, 2014) Normalizing to a range of 0 to 1 is accomplished using the equation

$$x_{\text{norm}} = \frac{x_{\text{true}} - x_{\text{min}}}{x_{\text{max}} - x_{\text{min}}}. \qquad \text{(Eq. 4)}$$

All continuous data, both for inputs and training targets, are normalized in this way before being given to the model. These consist of FRP, MERRA-2 AOD, satellite DB AOD, P_{surf}, P_{msl}, T_{surf}, $T_{2\text{m}}$, $DWT_{2\text{m}}$, orography, $RH_{2\text{m}}$, $U_{10\text{m}}$, $V_{10\text{m}}$, vegetation, and PBLH. For the AOD values, this occurs after the log transform, using maximums and minimums of the log AOD values.

3.3.3 One-Hot Encoding

The remaining variables (i.e., smoke density, aerosol type, and DB algorithm flag for land/ocean) are categorical variables and cannot be meaningfully normalized. However, as these variables are all stored as integer values from a predefined set (i.e., 1, 2, and 3 for light, medium, and heavy smoke, respectively), these assignments are arbitrary and do not make much sense when interpreted numerically (with the possible exception of smoke density since there is a natural ordering to its values). Instead, the classical approach to reformatting these variables for a more straightforward interpretation by the model is to use a one-hot encoding. This works by expanding a categorical variable with N categories into an N-length vector containing only zeros except at the location associated with the value for the variable. If we have a variable x which is of class c_i out of $c_1 \dots c_n$ classes, this is formalized by a one-hot vector

$$a_x = \langle 0,0,\dots 1,\dots,0 \rangle \qquad \text{(Eq. 5)}$$

Where $a_{x,i} = 1$ and $a_{x,j \neq i} = 0$.

For example, if we have a particular smoke density reading of heavy, and our one-hot encoding assignment is none, light, medium, and heavy from left to right, then the one-hot vector is $a_x = \langle 0,0,0,1 \rangle$. This method benefits neural networks because it directly presents the view that certain properties are "switched on or off" instead of forcing the model to learn these properties. This is particularly true for variables like aerosol type or algorithm flag, which do not have any natural ordering.

3.3.4 Missing Values

All NaNs in the data are converted to zeros before getting fed into the model. This is a common solution for dealing with missing values in deep learning, where the model is expected to extract useful features from noisy and lossy data sources.

3.3.5 MERRA-2 Augmentation

When training the model using purely DB retrievals, it can be noted from the resulting predictions that most (high-AOD) artifacts occur in areas that do not often get DB retrievals due to cloud cover or bright surfaces. An example of this is shown in Figure 10, where the model estimated unrealistically high values of AOD across the Pacific Northwest and southern Canada (Figure 10.a). These artifacts make it into the final gap-filled AOD product (Figure 10.c) because of the lack of DB retrievals over those areas (Figure 10.b). Furthermore, MERRA-2 suggests no elevated AOD values in these areas (Figure 10.d).

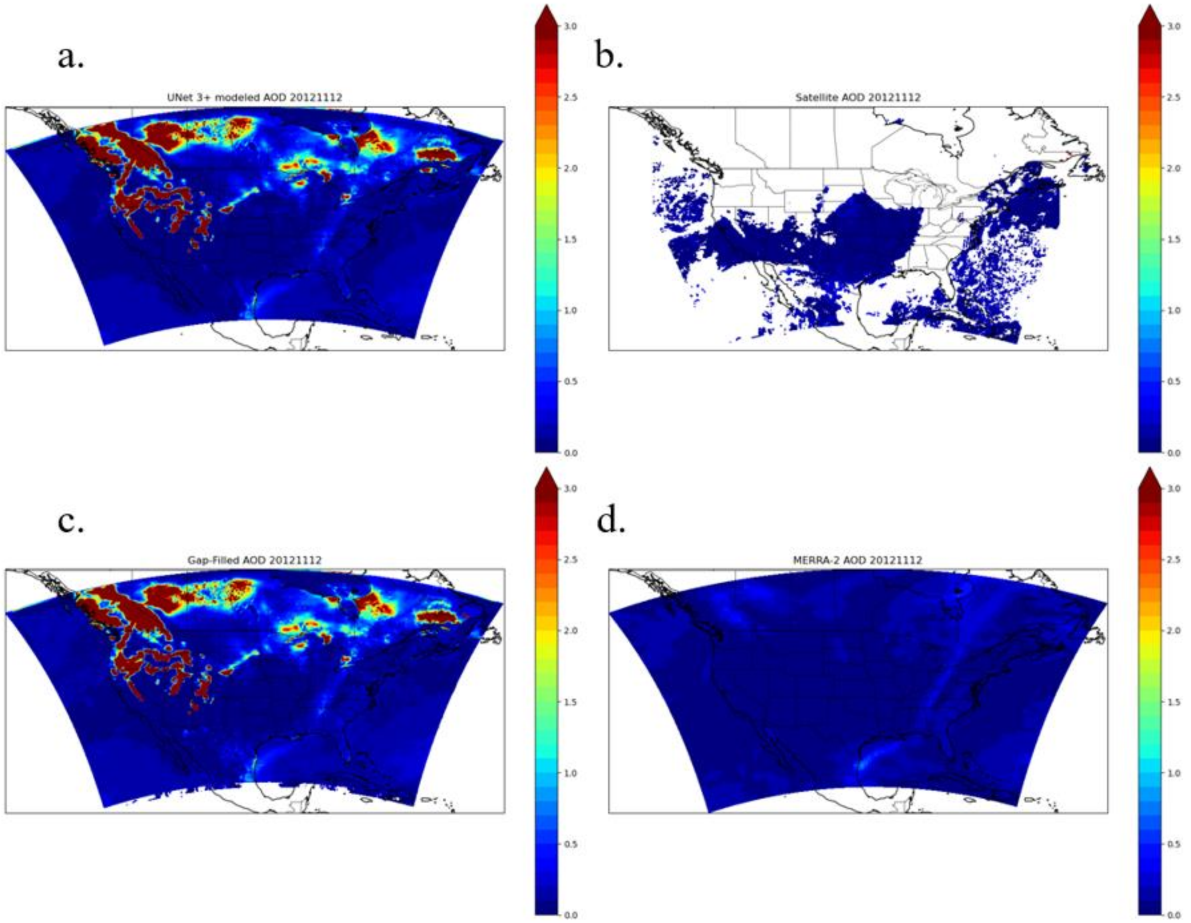

Figure 10: (a) UNet 3+ AOD predictions, (b) DB retrievals, (c) gap-filled AOD, and (d) MERRA-2 AOD for December 11, 2012. Gap-filling performed by the gap-filler model trained on only (weighted) satellite DB AOD retrievals.

Comparing this to the AOD data availability map for winter shown in Figure 11.a, 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 pyro cumulonimbus 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.

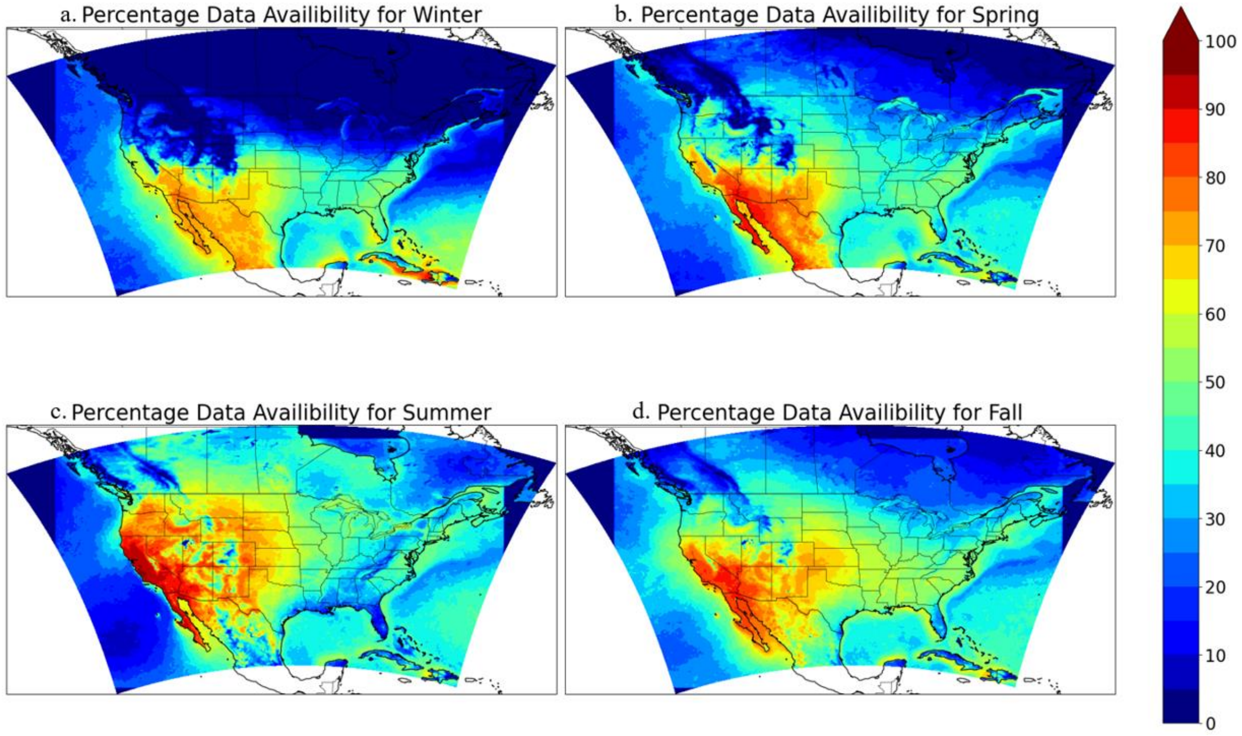

Figure 11: The percentage of days with a successful DB AOD retrieval from 2012-2022 categorized by season.

A composite of both DB AOD retrievals and MERRA-2 AOD estimates was created and used as the training target to mitigate this issue (Figure 12). The first layer of this composite is comprised of the DB AOD retrievals (Figure 12.a), 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 12.b). The monthly AOD data availability is calculated by taking all DB AOD retrievals over the 2012-2022 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.

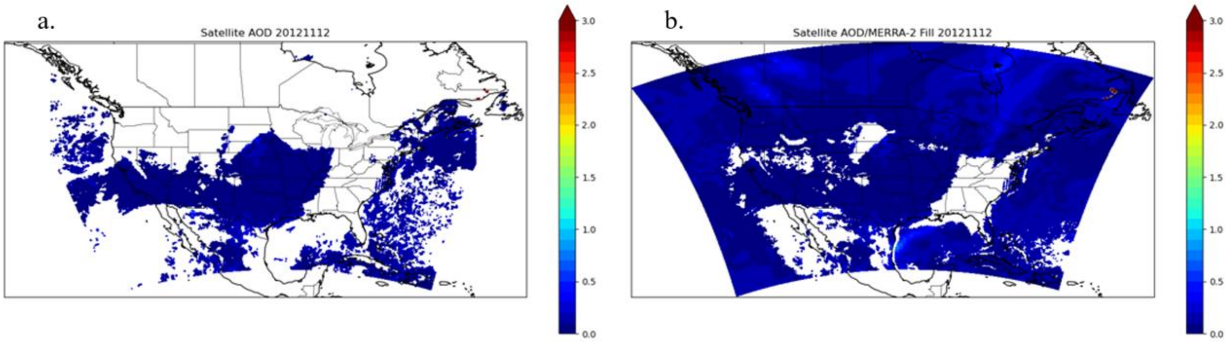

Figure 12: (a) DB AOD retrievals along with (b) supplemental AOD from MERRA-2 for November 12, 2012.

3.4 Model Training

3.4.1 Model Inputs

The gap-filled AOD model was input with the two previous days' AOD values along with meteorology, land-use, etc. variables to inform transport, evolution, and deposition patterns of aerosols into the third (target) day. All input variables used and their respective lags are listed in Table 3.

Input Variable	Resolution	Lag (days)
NASA MODIS DB Best Estimate AOD	10 x 10 km^2	2, 1
VIIRS DB Best Estimate AOD	6 x 6 km^2	2, 1
VIIRS DB aerosol type, land and ocean algorithm flags	6 x 6 km^2	2, 1

NAM 12km reanalysis – P, T, U, V, orography, vegetation, PBLH	12 x 12 km^2	2, 1, 0
MERRA-2 reanalysis AOD	50 x 62.5 km^2	2, 1, 0
HMS Smoke Product	~1 km	2, 1, 0
MODIS FRP	~1 km	2, 1, 0

Table 3: Input variables, native resolutions, and lags provided to the model.

3.4.2 Target Weighting

To further counteract the class imbalance problem, we modified the loss function to punish the model more for mistakes on high AOD pixels than low AOD pixels by adding a large multiplier to the squared error between high AOD samples and their respective predictions. This artificially inflates the MSE between high AOD samples and their predictions, while keeping the MSE for low AOD samples the same. Doing so will make the loss for high AOD samples so high that they cannot be ignored during training, despite only accounting for a small portion of the training dataset. In practice, weighting AOD samples according to user-defined thresholds has the effect of boosting model sensitivity to those thresholds relative to the assigned weights. This method allows for added user control over model characteristics. The weights chosen are listed below in Table 4.

AOD	**Weight**
$0 < \tau \leq 0.3$	1
$0.3 < \tau \leq 1$	3
$1 < \tau \leq 2$	5
$2 < \tau \leq 3$	10

3 < τ	15

Table 4: Target AOD weights.

3.4.3 Selective Backpropagation

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 (MLP) can be seen in Figure 13. This is accomplished by setting the target weight matrix to zero for pixels with 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.

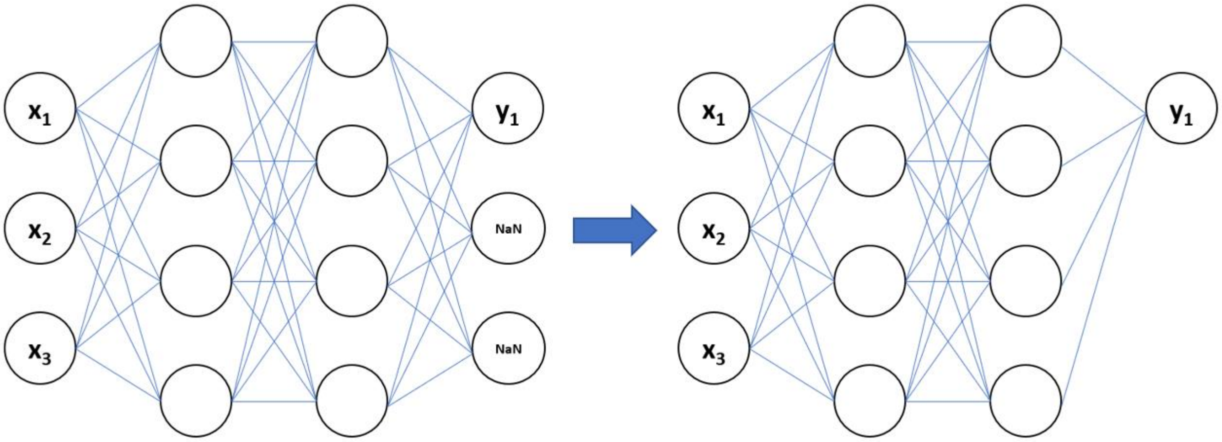

Figure 13: The effect of selective backpropagation on a simple MLP model.

3.4.4 Loss Function and Optimizer

The loss function is a composite one, consisting of a linear combination of the weighted sum of squared errors (SSE), which I will refer to as the pixel loss, and the multi-scale structural similarity index (MS-SSIM), which I convert into a patch loss. The prior 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

$$l_{\text{pixel}} = \sum_{p \in P} w_p \left(\widehat{y_p} - y_p\right)^2 \qquad \text{(Eq. 6)}$$

where P is the set of all AOD pixels, $\widehat{y_p}$ is the model prediction at pixel p, y_p is the actual AOD value at pixel p, and w_p is the training weight associated with pixel p.

MS-SSIM considers various metrics that roughly correspond to the luminance, contrast, and covariance of predicted and actual values within several differently sized sliding windows. It was developed by Zhou Wang, Eero P. Simoncelli, and Alan C. Bovik initially for image quality assessment , but was converted into a loss term in the original UNet 3+ paper 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 the patch loss simply by subtracting it from one.

These two losses are combined as a linear combination, i.e.

$$l = \alpha l_{\text{pixel}} + \beta l_{\text{patch}} \qquad \text{(Eq. 7)}$$

This loss function is the one the model is trained on.

We use the Adam optimizer to tune the model weights during training. Adam is a momentum-based gradient descent algorithm developed by Kingma and Ba (2017) that efficiently computes adaptive learning rates for different model parameters (Kingma & Ba, 2017). Momentum-based optimizers are preferable to vanilla gradient descent in this research, where memory limitations necessitate small batch sizes, and the gradient descent path is likely very erratic. Momentum-based learning schedulers have been shown to smooth this trajectory and speed up convergence in many problems.

3.4.5 Training Hyperparameters

The loss function parameters are $\alpha = 1e - 5, \beta = 1e - 2$. The learning rate is $1e - 5$. The weight decay is $1e - 5$. The Adam parameters are left as the defaults recommended by the original paper(Kingma & Ba, 2017). The batch size was set to either 1 or 2 depending on the availability of GPU’s (2 for A100 and 1 for 3090).

The complete dataset was split into 90-10 training/validation datasets. Validation was performed at the end of every epoch to guard against overfitting.

3.4.6 Input/Output Dimensionality

Considering the grid size of 360x515, the large model size, and limited GPU memory, we chose to make the inputs 352x352x90, essentially a 352x352 map for each of the 90 input variables. The 352 comes from the architectural requirement that the input must be divisible by 2^5 in each of the spatial dimensions. Samples are drawn from the northernmost 352 pixels, with one sample from the westernmost 352 pixels and another from the easternmost 352 pixels. This results in an overlapping region in the middle. The output is 352x352x1, i.e., the target AOD map.

3.5 Metrics

We use a few basic statistical metrics to evaluate the performance of the gap-filling model. The first is root mean square error (RMSE), which signifies the average distance between predictions $\hat{y}_i$ and targets y_i and is given by

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\hat{y}_i - y_i)^2}. \qquad \text{(Eq. 8)}$$

The Pearson correlation coefficient r indicates to what degree the predictions and target vary with each other, with 1 being perfectly correlated and -1 being perfectly inversely correlated. It is given by

$$r = \frac{cov(\hat{y}, y)}{\sigma_{\hat{y}}\sigma_y}. \qquad \text{(Eq. 9)}$$

The coefficient of determination R^2 indicates the fraction of variance explained by the model. It is given by

$$R^2 = 1 - \frac{SS_{\text{res}}}{SS_{\text{tot}}} \qquad \text{(Eq. 10)}$$

where $SS_{\text{res}} = \sum_{i=1}^{N}(\hat{y}_i - y_i)^2$ is the residual sum of squares and $SS_{\text{tot}} = \sum_{i=1}^{N}(y_i - \bar{y})^2$is the total sum of squares. R^2 is equal to 1 when the model perfectly predicts all values, 0 when the model always predicts the mean of the dataset, and can extend into the negatives if it performs worse.

The normalized mean bias indicates the average bias of predicted values as a fraction of the average true value, and is given by

$$NMB = \frac{\sum_{i=1}^{N}(\hat{y}_i - y_i)}{\sum_{i=1}^{N} y_i}. \qquad \text{(Eq. 11)}$$

This value will be positive if the model tends to overestimate and negative otherwise.

To compute these metrics, outputs from the model are collocated with AERONET stations within a 25 km radius, as per (Hsu et al., 2019; Sayer, 2020). Comparisons are made using UNet 3+ predictions and the gap-filled product as well as satellite DB AOD retrievals and MERRA-2 AOD as a point of comparison.

AERONET AOD was interpolated to 550 nm to match DB AOD retrievals according to Eck et al., (1999)

$$AOD_{550nm} \approx AOD_{675nm} \frac{675^{AE_{440-675nm}}}{550^{AE_{440-675nm}}}. \qquad \text{(Eq. 11)}$$

We use visual inspections to evaluate the model qualitatively and comparing them to satellite visible imagery. This is an imperfect mode of analysis due to the fact that many of the obstacles stopping AOD retrieval algorithms also prevent visual analysis (i.e., cloud cover), but certain desirable and undesirable features can be noted by a satellite, aerosol, and atmospheric composition expert such as my advisor Dr. Loría-Salazar.

4. Results and Discussion

4.1 Quantitative Results

4.1.1 Comparison With 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 5. The RMSE~0.09 and r~0.79 indicate that UNet 3+ and gap-filled AOD values generally do not deviate from AERONET AOD values and follow the same overall trend. The R^2~0.34 suggests that UNet 3+ AOD struggles to capture as much of the variance in AERONET AOD as MERRA-2 (R^2~0.56) or DB (R^2~0.41). However, the gap-filled AOD does not have as much of this problem, as its R^2~0.51 is on par with that of MERRA-2 and satellite. The R^2 values are difficult to interpret due to the extreme skewness of the AOD distribution. Due to this skewness, the majority of variance in the data comes in the form of small fluctuations around (low) baseline AOD values. AERONET AOD, for example, has a mean of 0.096 and a standard deviation of 0.11. It is challenging to capture a large percentage of these minor fluctuations, as evidenced by the low MERRA-2 and DB R^2 values. However, characterizing these small fluctuations in certain contexts, such as air quality studies, has a low impact on results as long as the overall AOD loading is characterized correctly (i.e., low/medium/high aerosol loadings). The NMB shows that all AOD products have a positive bias (~ 20%) against AERONET. Superimposing DB AOD on top of UNet 3+ AOD to create the gap-filled AOD increases NMB from 19.1% to 21.4%, despite satellite DB AOD generally having a lower NMB of 17.3%. This suggests that UNet 3+ tends to overestimate

AOD, particularly in areas with DB retrieval gaps. Finally, gap-filling the satellite DB AOD retrievals results in an additional ~30,000 collocated observations with AERONET over the 11-year period.

	RMSE	**r**	**R^2**	**NMB**	**N**
MERRA-2	0.07	0.78	0.56	19.5%	138,709
DB	0.09	0.82	0.41	17.3%	108,725
UNet 3+	0.09	0.79	0.34	19.1%	138,709
Gap-Filled	0.08	0.84	0.51	21.4%	138,709

Table 5: Statistics of MERRA-2 AOD, DB AOD, UNet 3+ AOD, and gap-filled AOD against AERONET AOD

The corresponding scatterplots are shown in Figure 14. MERRA-2 (Figure 14.a) appears to underestimate most AOD values larger than 1 but is reasonably accurate for lower values of AOD (RMSE~0.07, r~0.78, NMB~19.5%). DB (Figure 14.b) on the other hand appear to overestimate most AOD values larger than 1. There also appears to be a cluster of points where DB suggested high AOD (> 0.5) despite collocated AERONET retrievals at around 0.1. This is likely due to smoke plumes within a DB retrieval pixel but not directly over an AERONET station, but further investigation is required to confirm this. The statistical evaluation for DB shows RMSE~0.09, r~0.82, and NMB~17.3%. UNet 3+ (Figure 14.c) has performance metrics (RMSE~0.09, r~0.79, NMB~19.1%) similar to the previous two comparisons. However, unlike MERRA-2, it appears to overestimate at higher values of AOD (>0.3). Also, 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 14.d) shows some of the best performance metrics yet (RMSE~0.08, r~0.84, NMB~21.4%), combining the high correlation from DB with the low RMSE from MERRA-2 and proper characterization by UNet 3+.

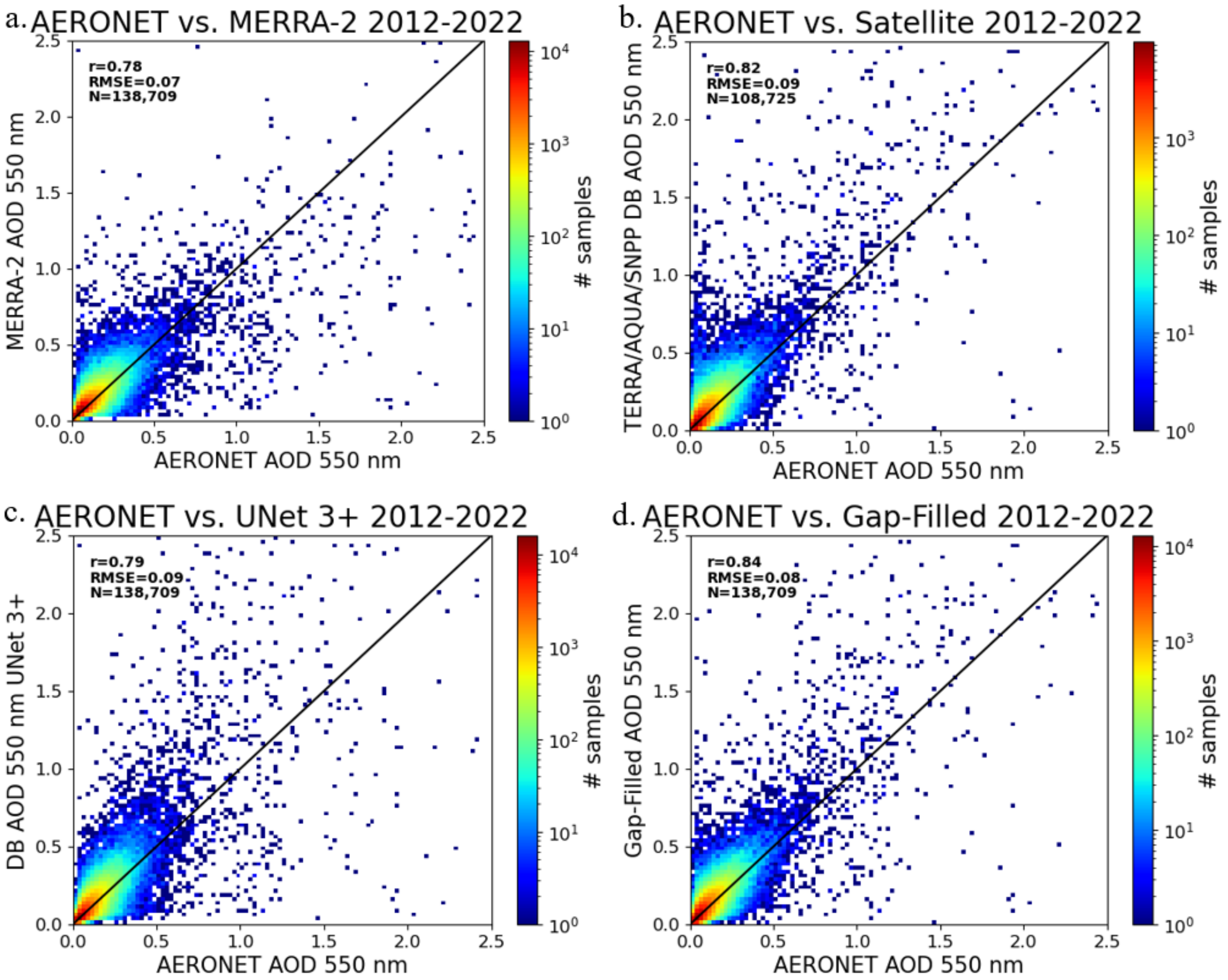

Figure 14: Scatter plots of (a) MERRA-2 AOD, (b) DB AOD, (c) UNet 3+ AOD, and (d) gap-filled AOD against AERONET AOD at 550 nm. The color scale represents the density of samples, and the black line represents the 1:1 line.

4.1.2 Spatial Evaluation

A spatial evaluation between UNet 3+ and DB AODagainst AERONET AOD is shown in Figures 15-19, the figures show results from spatial RMSE, r, and R^2 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 15), UNet 3+ performs slightly better in RMSE throughout eastern CONUS and the Midwest than DB, e.g., RMSE~0.12 in Illinois and Texas for DB compared to RMSE~0.07 for UNet 3+. The main improvement here is arguably the 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+ appears to struggle the most in the Pacific Northwest. RMSE's in this region are increased to around 0.15-0.25, compared to the <0.07 values present throughout the rest of CONUS. However, DB AOD retrievals also show similar difficulty in properly characterizing AOD loadings in this region. 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. This challenge could be resolved with the upcoming NASA PACE mission equipped with the Ocean Color Instrument (OCI) with a hyperspectral band (ultraviolet - shortwave infrared). Because of the MODIS and VIIRS spectral bands, aerosol retrievals need to assume the aerosol absorption nature and the aerosol height. Due to the spectral band capabilities of OCI, it will be possible to retrieve aerosol loading without assuming either aerosol absorption or height components (Personal communication with Dr. Andrew Sayer, NASA PACE Mission Lear Scientist).

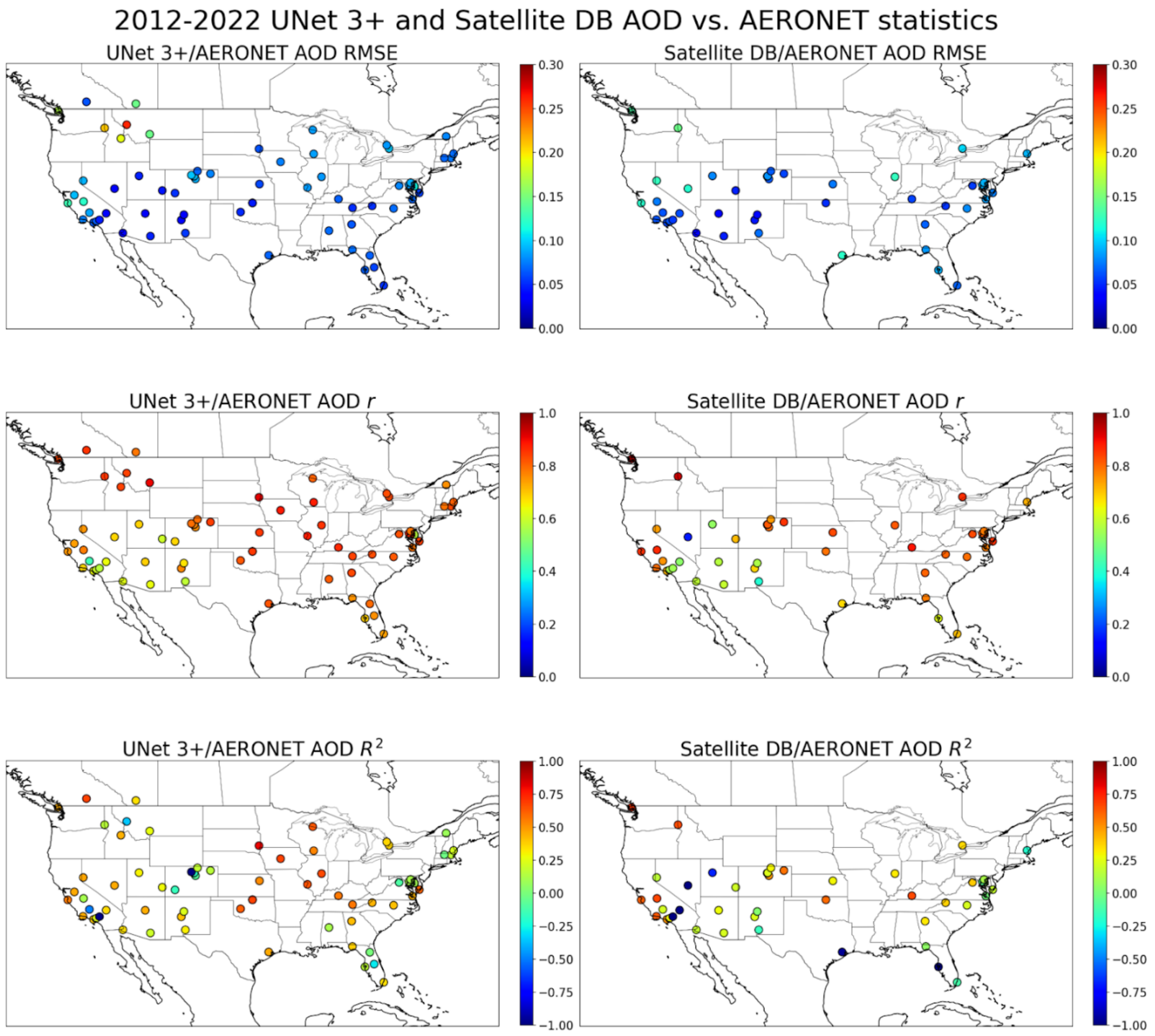

Figure 15: Spatial statistical evaluation of UNet 3+ and DB AOD against AERONET stations throughout CONUS with at least 1,000 collocated observations.

Figure 17, Figure 18, Figure 19, and Figure 19 show the same analysis as the previous one but separated by season. These seasonal maps instead select stations that have at least 200 collocated points. For winter (Figure 16), we can see UNet 3+ achieves very low RMSE (<0.01) everywhere except for a few select stations in California where it reaches ~0.11. UNet 3+ performs roughly the same as DB in terms of RMSE. Notable discrepancies include the aforementioned stations in California, where UNet 3+ (~0.11) performs worse than DB (~0.7), as well as Nevada, Texas, and Illinois, where DB (~0.12) performs worse than UNet 3+ (0.7).

The coverage gains with gap-filling are very pronounced in winter. For DB, the northern half of CONUS is missing, but recovered in the UNet 3+ maps. Since AOD values are chronically low throughout the winter and lack much variance for UNet 3+ to capture, r and R^2 values are also low.

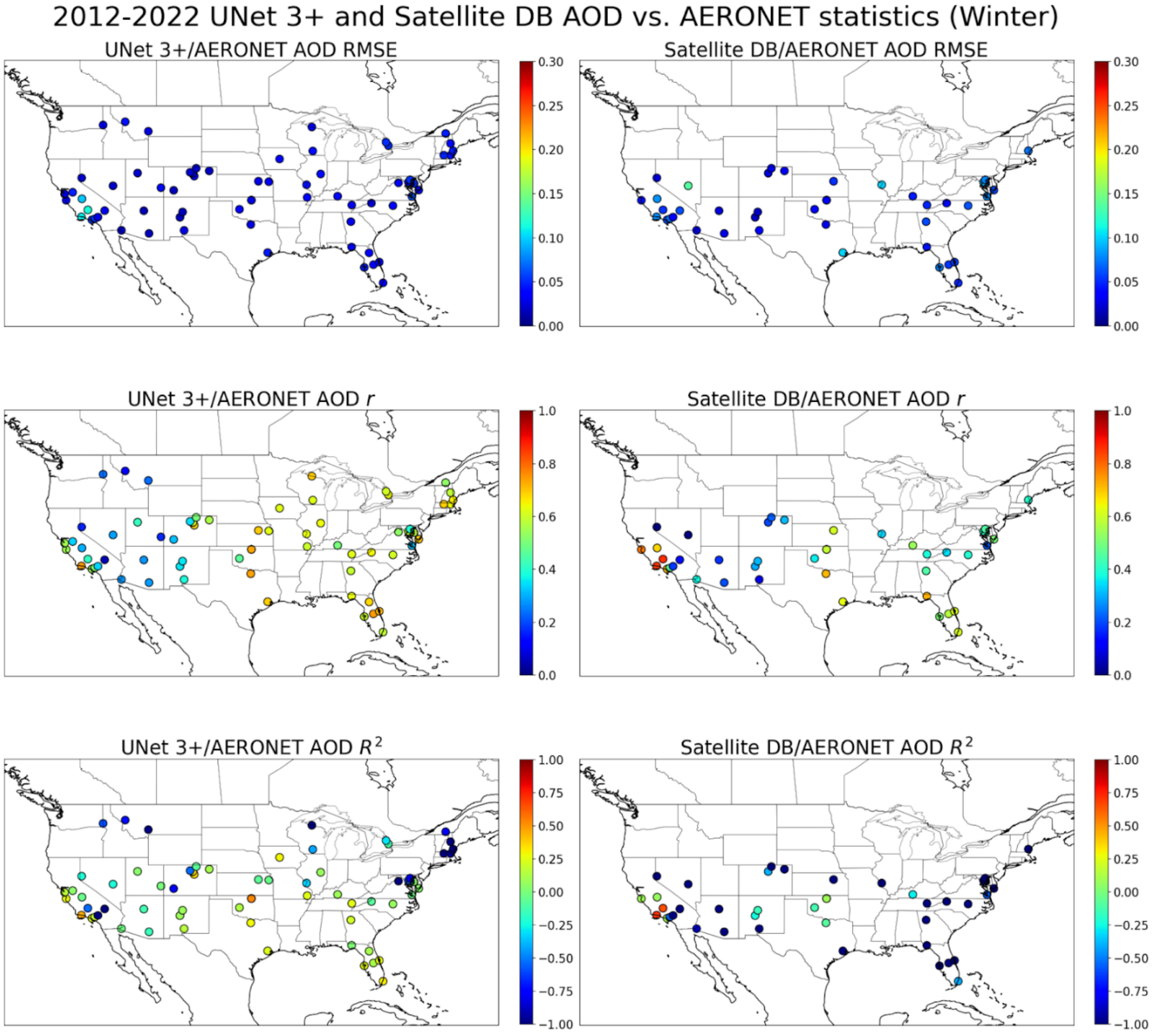

Figure 16: Spatial statistical evaluation over winter (DJF) of UNet 3+ and DB AOD against AERONET stations throughout CONUS with at least 200 collocated observations.

In the spring (Figure 17), RMSE is still generally low for both UNet 3+ and DB. UNet 3+ RMSEs are all below 0.1, while just a few DB RMSEs peak past 0.1. However, a clear distinction between east and west can be seen in the r and R^2 maps, where the west performs noticeably worse than the east for both UNet 3+ and satellite. Namely, for UNet 3+, r values do not drop below 0.7 in the eastern U.S. but do not exceed 0.65 in the western U.S. except for one station. This pattern is mirrored in DB, where several stations in the western U.S. are near 0. This drop in performance is in line with previous investigations of AOD retrievals in the semi-arid western U.S. (Loría-Salazar et al., 2016). It hints at an intrinsic limit in AOD gap-filling performance with the given data inputs. Another explanation as to the lower correlations in the western U.S. is described in Sayer et al., (2019), which notes the lower correlation between Terra, Aqua, and VIIRS collection 6.1 DB AOD and AERONET over this region due to the relatively lower dynamic range of AOD (e.g. over clean mountainous sites) and positive outliers such as dry lake beds and cities in rugged terrain which are difficult to characterize (Sayer et al., 2019).

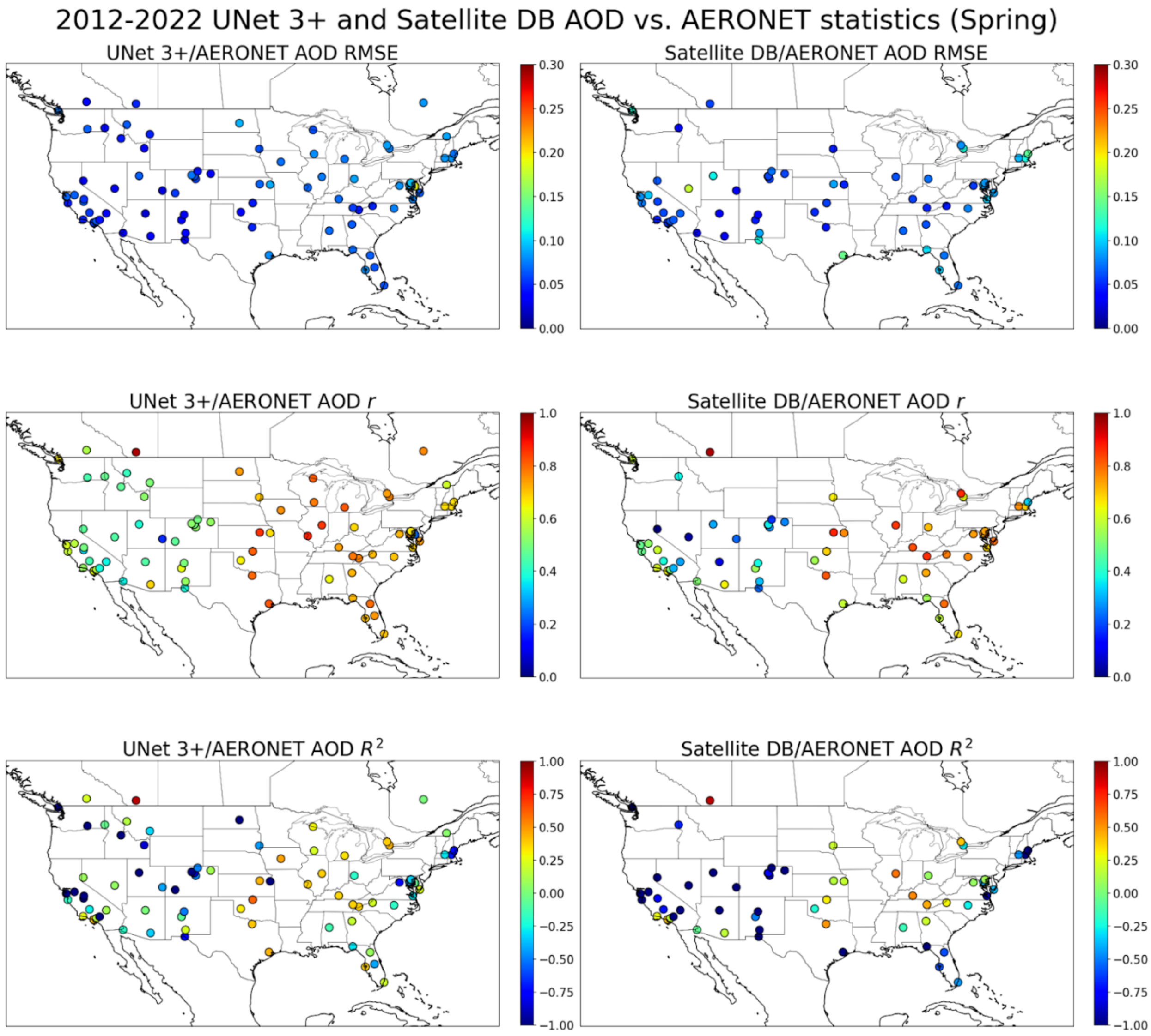

Figure 17: Spatial statistical evaluation over spring (MAM) of UNet 3+ and DB AOD against AERONET stations throughout CONUS with at least 200 collocated observations.

During the summer (Figure 18), the Pacific Northwest has lower evaluation metrics (e.g., RMSE>0.2 and R^2~0 for a few stations). This is likely because summer is peak wildfire season and characterizing the exact values of elevated AOD levels (AOD>0.3) in smoke clouds is still difficult for UNet 3+. DB also struggles with this region in the summer, with all AERONET stations in this area having an RMSE>0.1. Compared to winter and spring, RMSE's become higher in more areas due to the increased dynamic range of AOD during this season. This is apparent in California, Colorado, the Midwest, and parts of the U.S. east coast. UNet 3+

estimates are also now more correlated with AERONET retrievals, with all stations having a correlation of at least r~0.55. However, the R^2 map for UNet 3+ reveals a select number of stations in California, Montana, Colorado, Alabama, Florida, and Virginia with negative R^2, which implies biased estimates over those areas. Most of these stations also suffer from negative R^2 values with DB, suggesting that these biases are learned from training against DB retrievals.

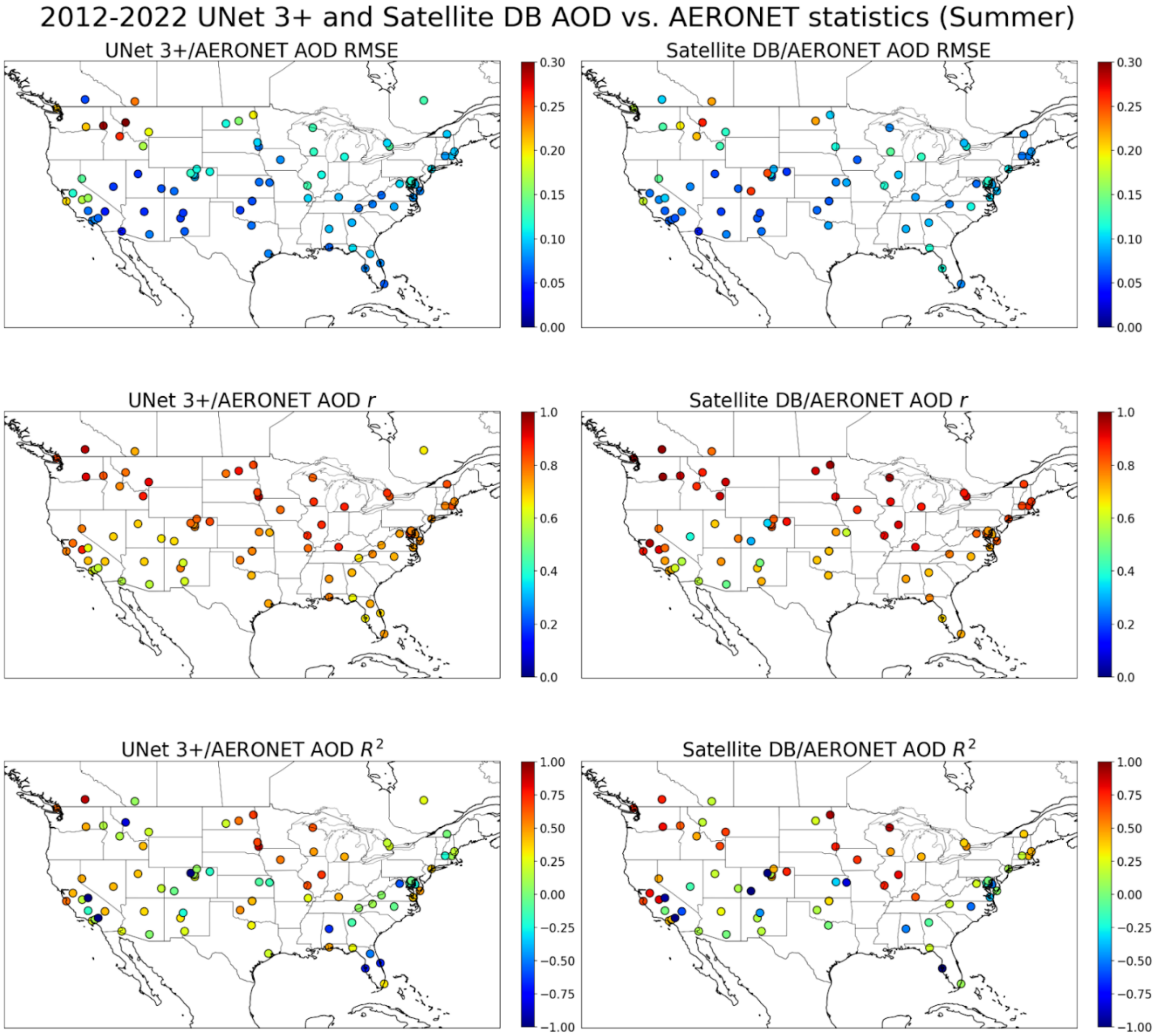

Figure 18: Spatial statistical evaluation over summer (JJA) of UNet 3+ and DB AOD against AERONET stations throughout CONUS with at least 200 collocated observations.

Model performance in the fall (Figure 19) seems to be better than in the summer, particularly for the eastern U.S. which sees RMSEs drop and rs rise. The Pacific Northwest, California, and Florida remain challenging locations for the model.

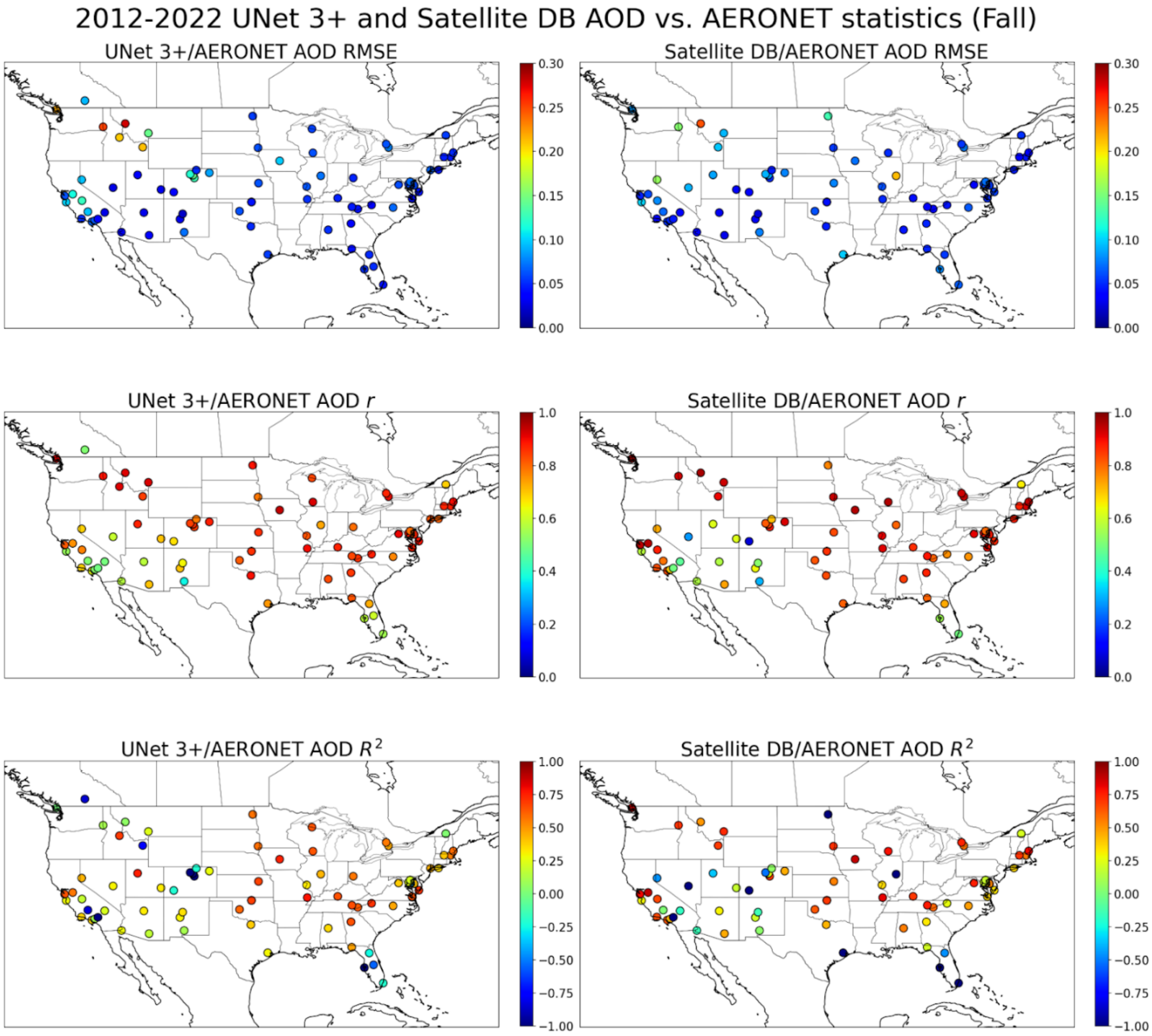

Figure 19: Spatial statistical evaluation over fall (SON) of UNet 3+ and DB AOD against AERONET stations throughout CONUS with at least 200 collocated observations.Figure 20 shows residuals as a function of AERONET AOD values ($\text{AOD}_{\text{UNet 3+}}$ - $\text{AOD}_{\text{AERONET}}$), with the blue line showing the mean residual and red area denoting ±1 standard deviation, i.e. 68.2. For low-medium (0.1<AOD<0.3) AOD events, UNet 3+ has a low bias (<0.2 at 550 nm). However,

for higher AOD loadings (AOD >0.3) both the uncertainty and bias grows significantly (> 1.5 at 550 nm).x

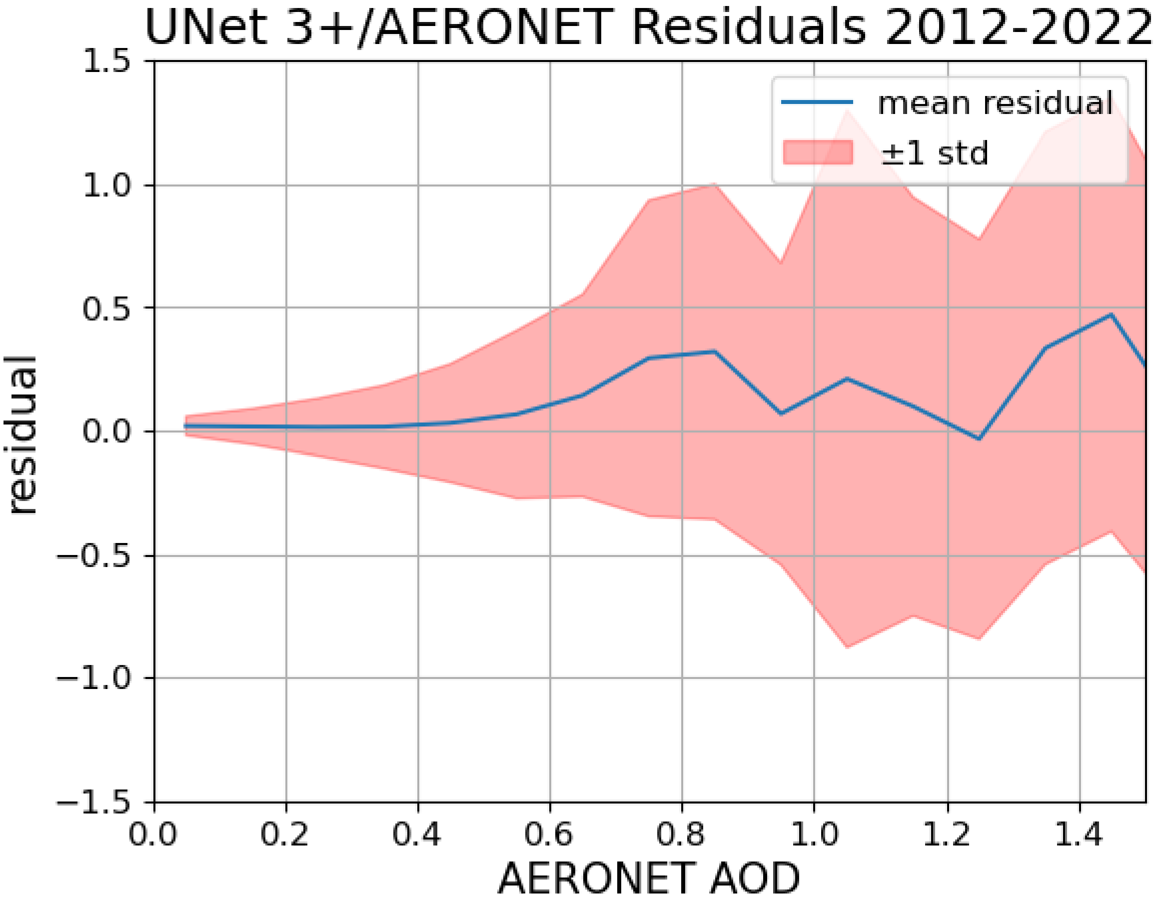

Figure 20: Residuals of UNet 3+ AOD against AERONET AOD. Blue denotes the average residual and red denotes one standard deviation.

Figure 21 shows the composite of the daily differences between UNet 3+ predictions and DB AOD retrievals. Over CONUS, UNet 3+ and DB AOD are well aligned apart from the Pacific Northwest, where UNet 3+ has overestimated AOD levels by around 0.1 due to wildfire smoke, and east of Hudson Bay, where UNet 3+ has underestimated baseline AOD loadings (~-0.1).

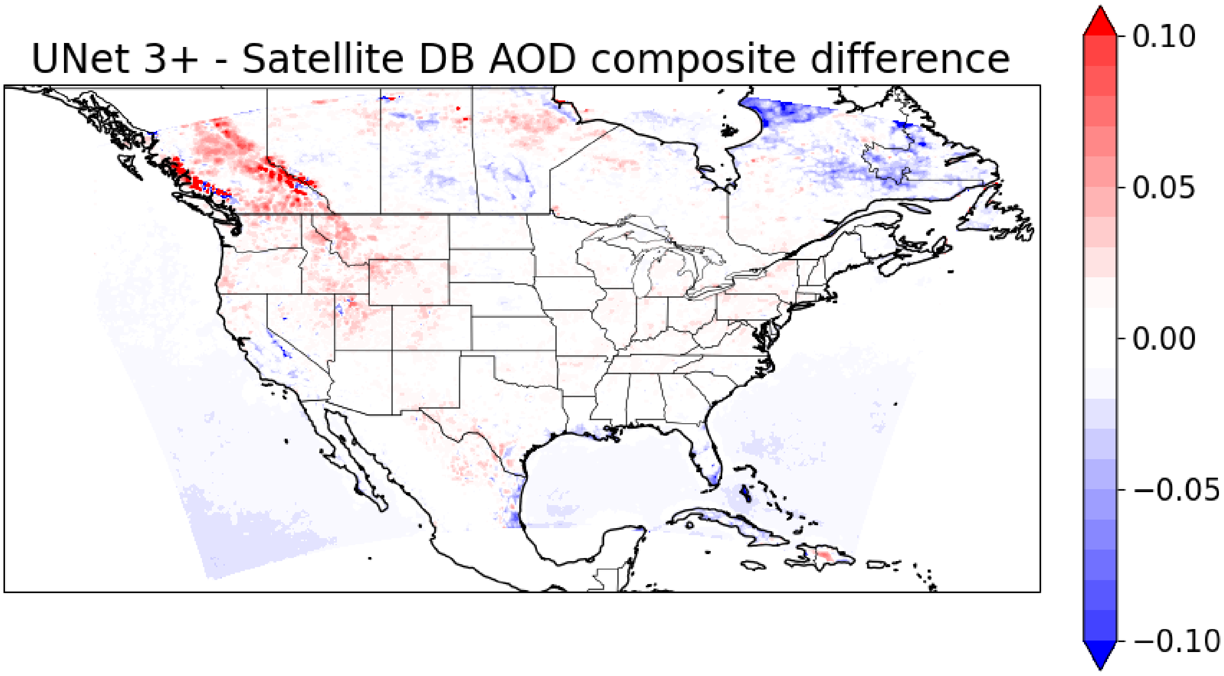

Figure 21: Composite difference between UNet 3+ AOD and satellite DB AOD.

The winter composite (Figure 22) presents challenges for interpretation due to the deficient quantity of retrievals in the northern U.S. and Canada. Still, UNet 3+ accuracy in this region likely suffers over the winter. The UNet 3+ training over this region in the winter is mainly guided by MERRA-2 AOD and not DB AOD retrievals (Figure 11.a).

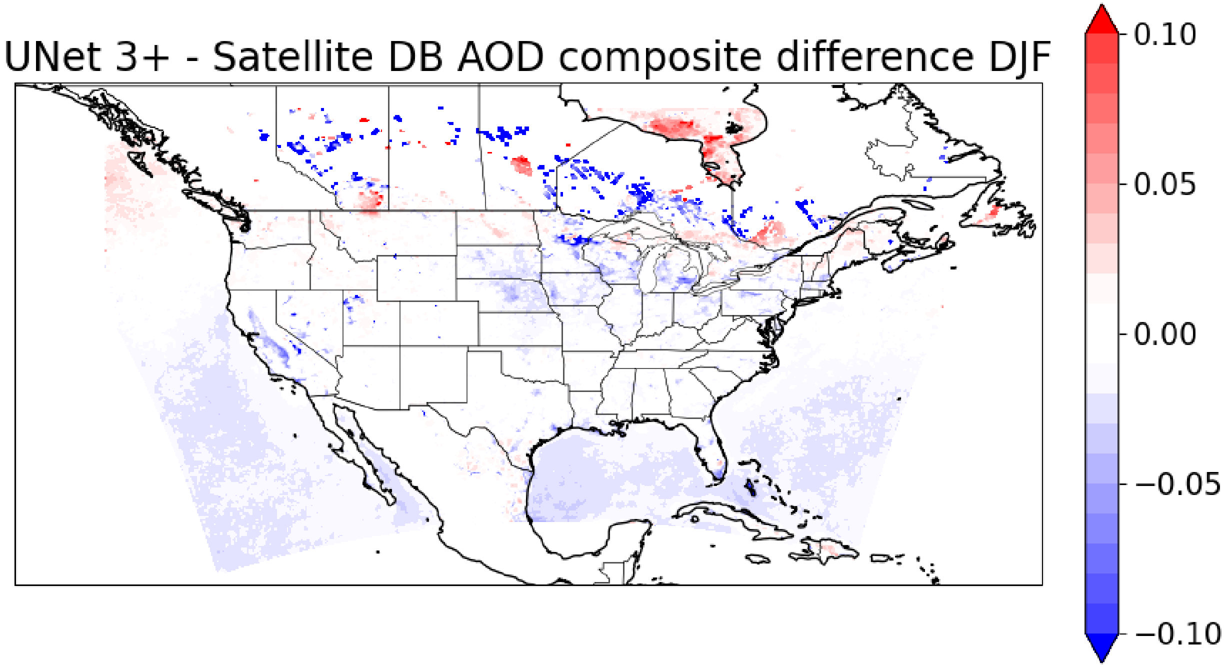

Figure 22: Composite difference between UNet 3+ AOD and satellite DB AOD for winter (DJF).

In the spring (Figure 23), there appears to be a systematic underestimation of AOD of at least 0.1 south of Hudson Bay and in Quebec and Newfoundland. A thin strip of the western coast of British Columbia also experiences systemic underestimation of at least 0.1. Finally, UNet 3+ tends to slightly overestimate AOD by around 0.03 throughout various parts of CONUS (Rocky Mountains, Appalachia and the Midwest, and Southern Texas/Northwestern Mexico).

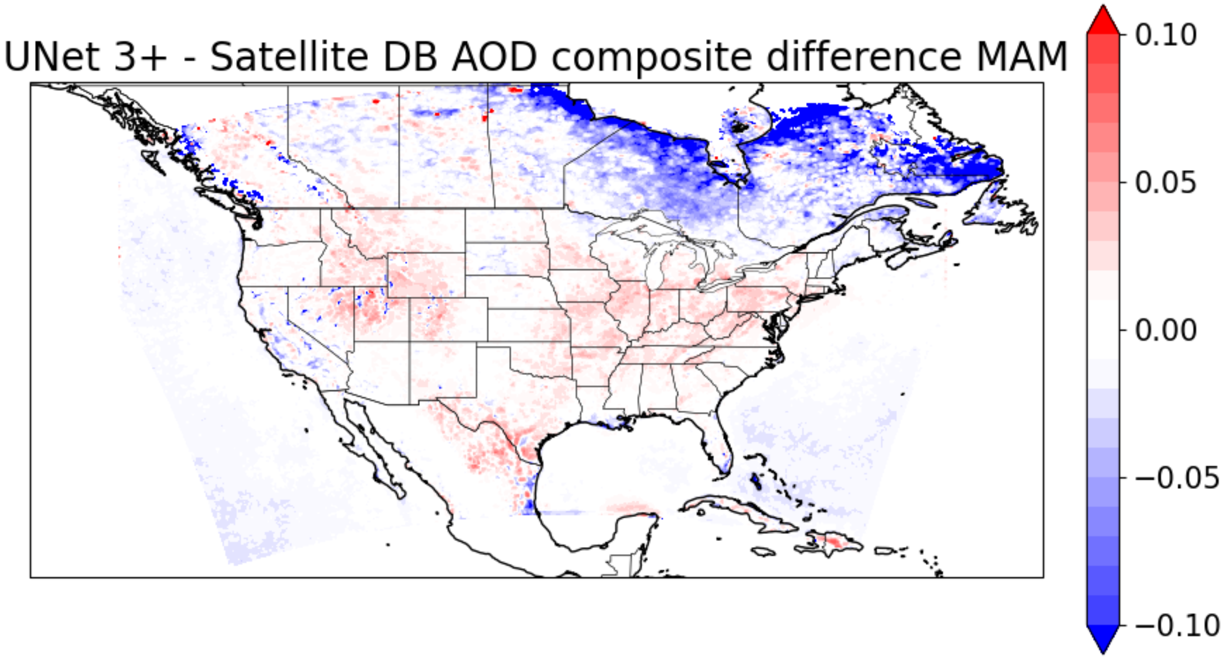

Figure 22: Composite difference between UNet 3+ AOD and satellite DB AOD for spring (MAM).

In the summer (Figure 3), UNet 3+ seems to overestimate, most notably in the Pacific Northwest (>0.1) and south of Hudson Bay (~0.05), most likely due to increased smoke presence across these regions due to wildfire season. Further analysis with AERONET needs to be performed to determine the degree of overestimation. The clusters of overestimations from the spring map appear to disperse, although they are still present. AOD in many regions bordering the Gulf of Mexico and the Caribbean is underestimated by around 0.1 in the summer likely due to the persistent cloud cover during this season in this region.

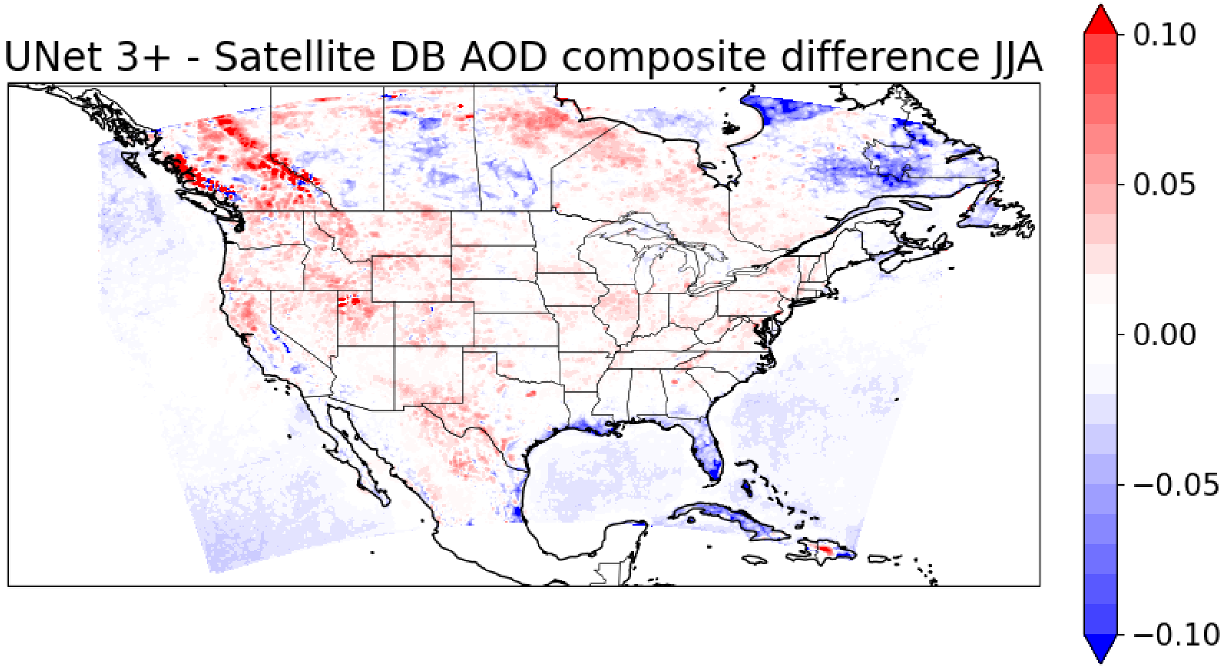

Figure 23: Composite difference between UNet 3+ AOD and satellite DB AOD for summer (JJA).

In the fall (Figure 24), UNet 3+ predictions align with DB retrievals outside of the Pacific Northwest, due to wildfire smoke, which can still occur in September and October. UNet 3+ again seems to overestimate AOD values in this region by up to 0.1. However, all other regions are within ±0.03 of DB on average.

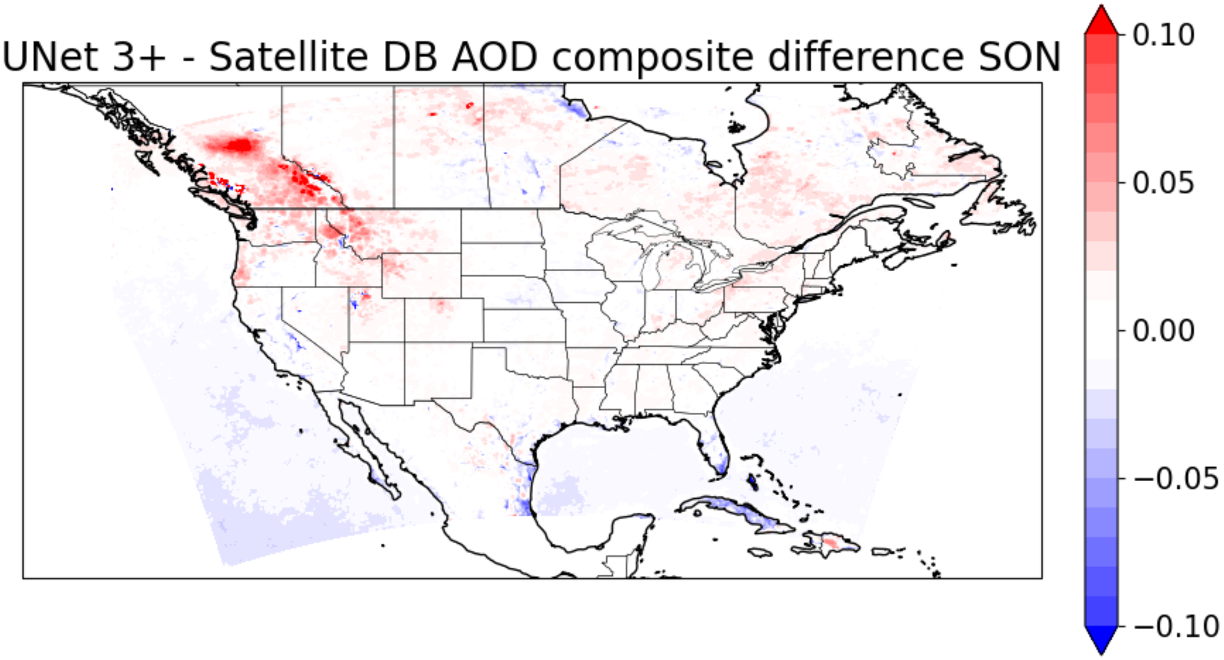

Figure 24: Composite difference between UNet 3+ AOD and satellite DB AOD for fall (SON).

4.1.3 Comparison with Existing Methods

Method	**RMSE**	**R**	**Region**	**Source**
Bayesian (still has gaps)	0.08-0.13	0.89-0.93	India	Singh et al., 2017
Bayesian Maximum Entropy (still has gaps)	0.29	0.75	China	Tang et al., 2017
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., 2021
Tensor Completion	0.21	0.91	China	Bai et al., 2022
4D-STET	Did not report	0.73	Global (land)	Wei et al., (2023)

UNet 3+	0.08	0.84	CONUS	Lee et al., in prep

Table 6: A collection of AOD gap-filling models and their performance.

Table 6 shows the statistical performance of previous gap-fill models. From Table 6, the gap-filled product from UNet 3+ performs either better or on-par with a selection of modern gap-filling methods when evaluated against AERONET for RMSE and r. UNet 3+ gap-filling scores the lowest RMSE, and its r is only surpassed by Bayesian merging (Singh et al., 2017), tensor completion (Bai et al., 2022), and the multi-stage spatiotemporal fitting developed by Zhang et al. (2021).

The Bayesian merging approach yields a daily AOD product with excellent performance metrics. Unfortunately, since the technique still relies on composing AOD retrievals from various satellites (MODIS and MISR in the original study), the resulting maps still have gaps where none of the input satellites could perform a retrieval. Also worth noting is that the authors accounted for “the error distribution of AOD from AERONET data” (Singh et al., 2017), meaning the model is directly optimized for estimating AERONET AOD densities. Despite the fact the UNet 3+ model we trained is AERONET-agnostic (i.e., it was not trained on any AERONET data), it still performed relatively well according to the statistical evaluation.

The tensor completion approach developed by Bai et al., (2022) introduces a workflow that is more similar to our methodology. AOD data from multiple satellite instruments (Terra MODIS, Aqua MODIS, Terra MISR, Suomi-NPP VIIRS, Envisat AATSR, and PARASOL POLDER) and MERRA-2 is first fed into a set of pre-processing models that regrids everything to a 1 km resolution using meteorology and land-use auxiliary variables. To perform the gap-filling, the regridded AOD tensors are passed into the higher-order singular value decomposition (HOSVD), with the MAIAC AOD product from Terra as a baseline. The authors used this method to create a long-term (2000-2020) daily 1 km resolution AOD dataset over China.

It is worth noting that the vast majority of AOD gap-filling efforts have been focused on Asia, specifically China, as indicated by Table 6. 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, and other characteristics. For instance, Bai et al., (2022) found that Aqua and Terra MODIS MAIAC retrievals over China had an RMSE of 0.13-0.14 and r of 0.95, and SNPP VIIRS MAIAC retrievals had an RMSE of 0.22 and r of 0.80. This contrasts with the RMSE of 0.09 and r of 0.82 for the satellite AOD composite used for our study. The approach we are presenting in this study holds value because it is the first AOD gap-filling model that can achieve performance metrics comparable to those developed in historically much more studied regions.

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 model predictions are plausible when considering seasonality, location, and satellite imagery. This section presents some case studies that provide insight into the model’s strengths and weaknesses. As mentioned in the introduction, two prominent high AOD features that an AOD model should ideally be able to characterize correctly are plumes of wildfire smoke and elevated aerosol loadings due to temperature inversions.

4.2.1 Temperature Inversions

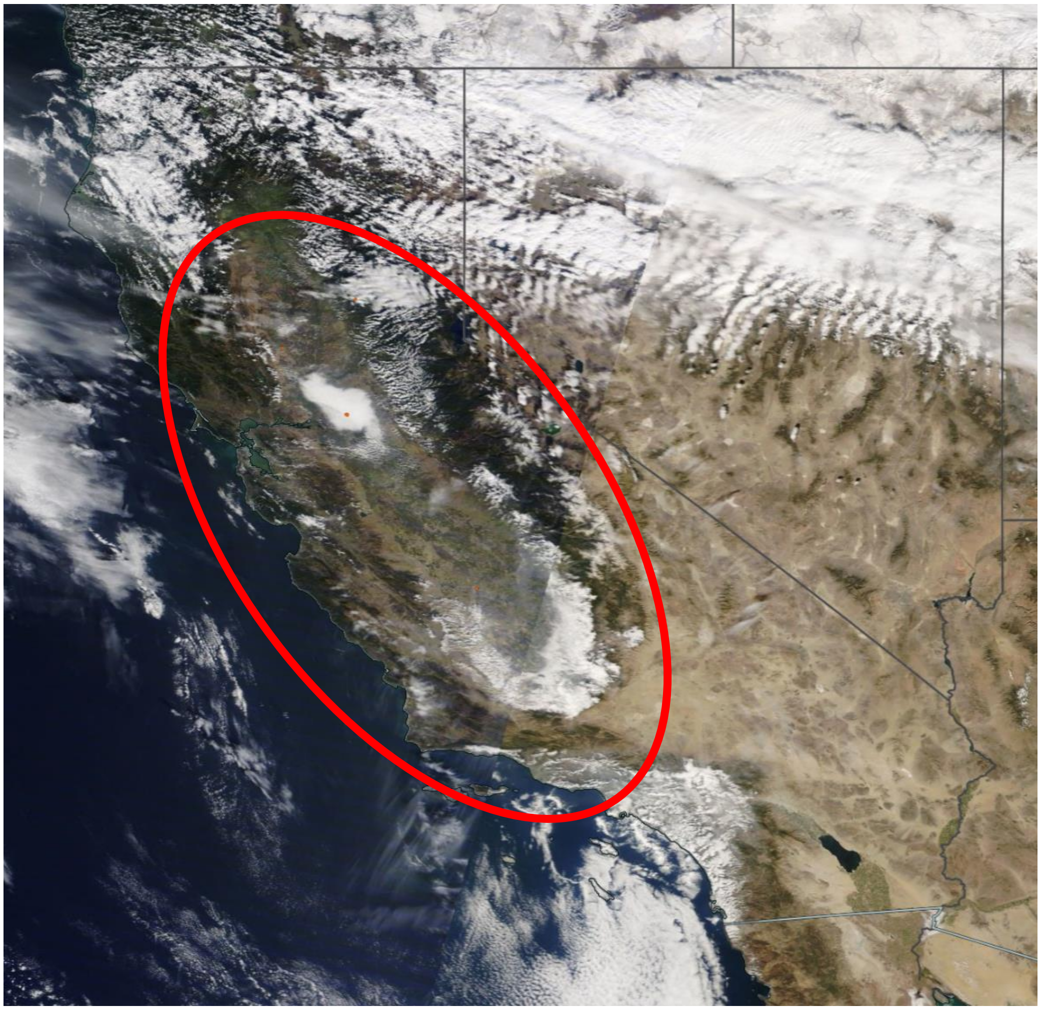

Figure 25: Haze trapped over Sacramento Valley (circled in red) during a temperature inversion on November 15, 2014. Source: NASA Terra visible image.

One region in the western U.S. that commonly experiences increased surface air pollution due to temperature inversions is the Central Valley in California. These events can be visible to the naked eye, provided no clouds cover the area. One such event occurred on November 15, 2014. Terra visible imagery shows most of the Central Valley is covered in a thin haze (Figure 25). This haze is notably strong in the southern tip of the valley and raises AOD values to around 0.2 based on available DB retrievals (Figure 26.b). However, due to cloud cover, DB retrievals over this area are incomplete (Figures 25 and 26.b). MERRA-2 only predicts increased AOD values of around 0.2 in just the northern part of the valley (Figure 26.d). On the other hand, UNet

3+ proposes increased AOD values of around 0.2 throughout the entire valley (Figure 26.a). The latter is a more accurate representation of aerosol loadings based on visible imagery. The gap-filled AOD refines the exact placement of the haze while providing reasonable estimates for the cloud-obstructed areas (Figure 26.c).

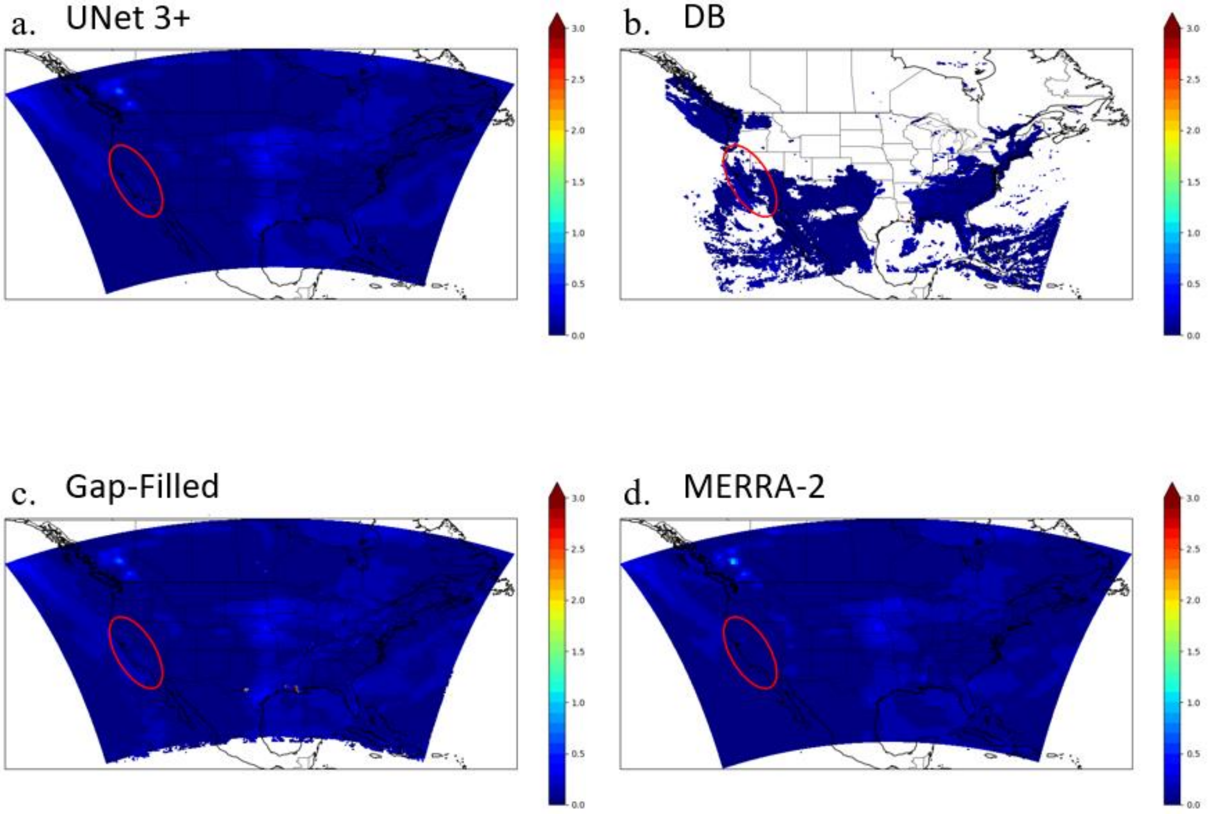

Figure 26: (a) UNet 3+, (b) DB, (c) gap-filled, and (d) MERRA-2 AOD for November 15, 2014.

The ability of UNet 3+ to predict increased AOD due to temperature inversions in California is made possible by the availability of DB retrievals in the area during such events. However, for regions like the Pacific Northwest that are more commonly covered by clouds (i.e., Idaho, Washington, Montana, Wyoming, etc.), such predictions are impossible without adding emissions data to UNet 3+.

4.2.2 Wildfire Smoke

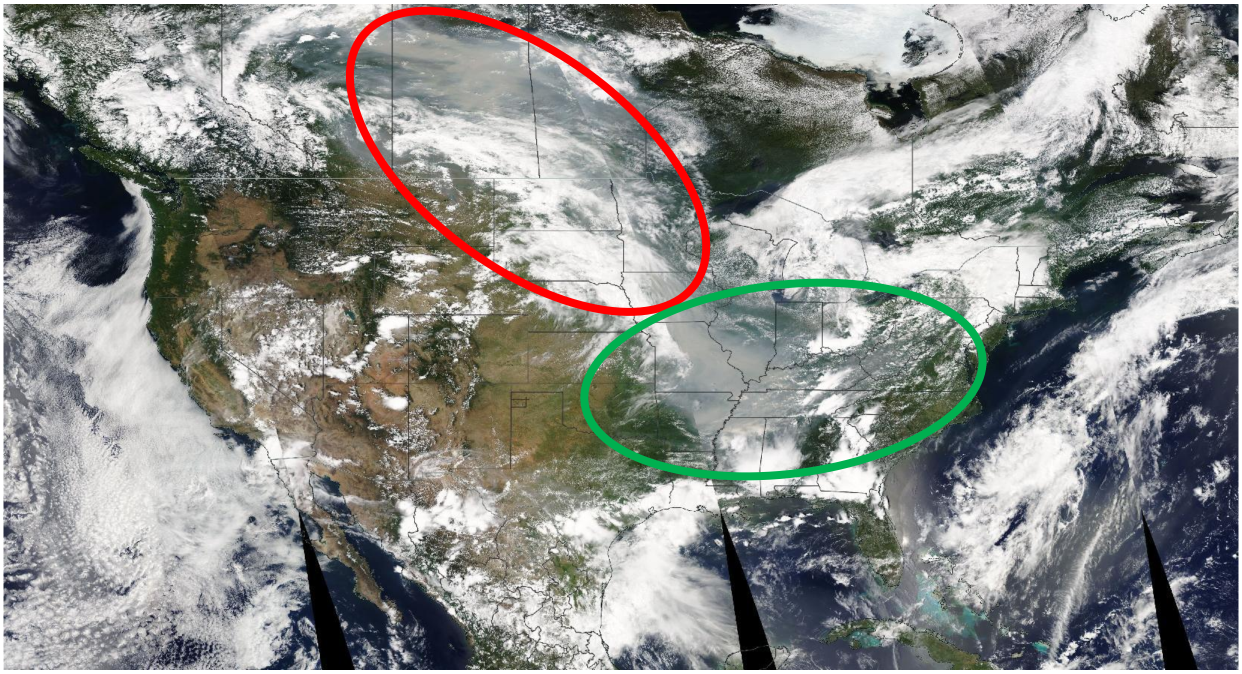

Figure 27: Wildfires over Canada (circled in red) and the Midwest (circled in green) generated heavy smoke plumes affecting North America on June 30, 2015. Source: NASA Terra visible image..

On June 30, 2015, multiple wildfires were experienced throughout Canada, primarily in Saskatchewan and Alberta (circled in red), as well as throughout the Midwest (circled in green) (Figure 27). Parts of these smoke plumes were occluded by cloud cover, leading to gaps in DB AOD retrievals (Figure 28.b). Figure 28.D shows that MERRA-2 generally retrieves high AOD (>3.0) 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 28.a). For the Canadian fires, the general 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 values ~3 for a larger area than MERRA-2, which is more in line with the DB retrievals of >3 for the entire plume. UNet 3+ suggests that the southern part of the plume (cloud-occluded) is also very dense, which we were not able to evaluate due to missing retrievals but does make sense contextualized by the available satellite retrievals in the gap-filled AOD. Similarly, for the smoke plume over the Midwest, UNet 3+ suggests more reasonably elevated AOD values of around 2.5-3 over Iowa and Missouri compared to MERRA-2, which suggests values of around 1.5 for most areas and 3 for just the source region. The same observation extends to Tennessee and Kentucky as well, where MERRA-2 proposes values of AOD from ~1.5 to ~2 but DB AOD retrievals are ~2.5.

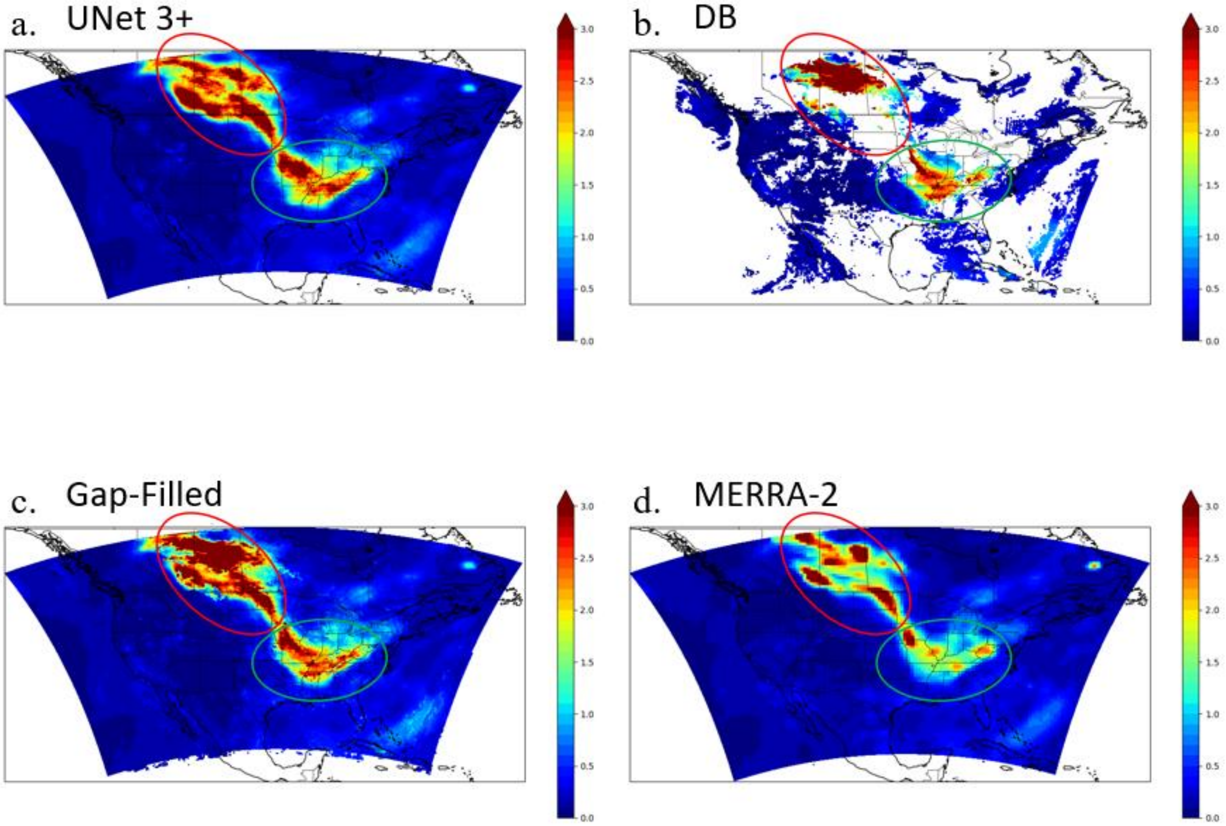

Figure 28: Wildfires over Canada (circled in red) and the U.S. (circled in green) on June 30, 2015. Source: NASA Terra.

4.3 Future Work

As the work presented in this thesis has primarily been exploratory and the numerous possibilities to be explored are gated behind cluster core hours, many of the design and engineering choices made from start to finish were based on qualitative assessment and not rigorous experimentation. As a result, future work on this project will be focused primarily on three objectives:

1. Developing a rigorous testing environment to quantify the performance of various model designs and configurations.
2. Expanding the collection of models to test the viability of different model inputs, hyperparameters, architectures, and training paradigms.
3. Exploring methods to explain the scientific mechanisms underlying trained models.

The first priority for future work is to develop an environment in which new models can be trained, validated, tested against a suite of performance metrics, and their results archived. Such an environment is required to systematically discover optimal combinations of inputs, architectures, and hyperparameters.

Once such an environment is developed, the next generation of gap-filling models can be developed. The focus of these new models will be many-fold: perform density estimation in order to provide uncertainty estimates and better capture the multi-modalities associated with transport and different aerosol types; test emissions inventories as a data source to help with urban and wintertime predictions; test regional/seasonal models; improve output resolution; develop forecasting models; and apply explainability techniques to determine key variables and dissect the mechanisms behind AOD gap-filling.

A significant focus for future work is how best to incorporate new data from PACE OCI. The at-launch application of DB to PACE OCI will be similar to that of MODIS and VIIRS, except at finer spatial resolution; however, due to increased spectral coverage, it is anticipated that improved aerosol type products (e.g., fine-resolution UV absorbing aerosol index, fine mode fraction) and an aerosol height product may become available. Additionally, PACE's two multiangle polarimeters will be able to provide more detailed aerosol microphysical categorization (e.g., size, sphericity, absorption). These upgrades could improve heritage aerosol property retrievals and offer valuable inputs to the gap-filling model. Given these improvements, assimilating PACE data into the existing gap-filling model trained on current DB data would be ideal for improving model performance. However, the methods for assimilating PACE AOD present some challenges due to new instrument calibration in the retrieval algorithm. I plan on testing statistical methods, such as taking the weighted average of all DB AOD sources or a weighted average of model outputs run on both current and next-generation data, as well as machine learning methods like transfer learning through fine-tuning the model in an online setting as PACE generates data.

Another future area of focus is bias correction. Current results show that the model can be skewed in specific regions and seasons. For example, the current model appears to overestimate AOD in the Pacific Northwest during wildfire season and underestimate AOD across southern Canada throughout the winter because it follows the trends of DB AOD. This is an inevitability due to the fact that we are training against DB AOD, which is an imperfect estimation of AOD. I plan to test methods such as subtracting the seasonal composite error between UNet 3+ and satellite products, training separate regional models, and detrending AOD data/predicting anomalies.

5. Summary

This thesis aimed to test the viability of UNet 3+, a deep learning (DL) architecture, as an architecture for aerosol optical depth (AOD) gap-filling. We tested the hypothesis that UNet 3+ trained on Deep Blue (DB) AOD and supplemental datasets, will improve spatial-temporal coverage of AOD missing data due to cloud cover and bright surfaces (e.g., semi-arid or arid deserts, snow, and ice).

The spatiotemporal gap-filling of AOD was performed daily over the CONtinental United States (CONUS) at a 12 x 12 km^2 resolution from 2012-2022. This was achieved by training the model on DB AOD and fire radiative power (FRP) from polar-orbiting satellites (i.e., Terra, Aqua, Suomi-NPP), AOD from the Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2) reanalysis, meteorological and land-use variables from the North American Mesoscale Forecast System (NAM), and smoke polygons from the Hazard Mapping System (HMS). With the above in mind, we found that:

- 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 a few data preprocessing steps, namely log transformation of AOD, data normalization, sample weighting, and training target augmentation.
- The resulting model infers aerosol loading and transport through modeled meteorological variables and DB AOD data when available. Where DB AOD is too sparse to provide enough context, the model provides upscaled MERRA-2 AOD values.
- The covariates chosen as inputs for the model enable the recognition of the first day of smoke in small fires (HMS and FRP), transport of aerosols (NAM), increased aerosol

loading due to temperature inversions, especially over California, and low aerosol loadings (MERRA-2 AOD).

- The model struggles most with areas that are sparse in satellite aerosol data. Notable areas with issues are the Pacific Northwest, Canada for most of the year, and Florida over the summer. These areas also have the most significant discrepancies with DB AOD, suggesting that future investigations should include additional data for the model to assess AOD better in these locations and/or develop separate regional/seasonal models that can specialize for AOD patterns specific to these locations.
- A statistical evaluation of UNet 3+ gap-filled AOD with respect to ground-truth AERONET AOD retrievals shows a root mean square error (RMSE)~0.08 and a Pearson correlation (r)~0.84. This performance is comparable to DB, which scores an RMSE~0.09 and r~0.82 in the same evaluation.
- The statistical evaluation metrics of this model are either better or on par with those of most preexisting AOD gap-filling methods (although most of these methods were developed specifically for parts of Asia) and roughly match the statistics of MERRA-2 and DB against AERONET. This level of performance allows for generating a spatially complete daily AOD dataset over CONUS, even the semi-arid western U.S., which has historically been challenging for this task.
- The computational requirements for the UNet 3+ gap-filling model are low compared to numerical and reanalysis models. Using an HPC node with 2 Nvidia A100 GPUs, it took ~12 hours to train the model on 11 years of data to convergence and seconds to generate AOD estimations over CONUS for one 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.
- Some potential challenges exist with the suggested approaches and, more generally, in estimating AOD and $PM_{2.5}$. The first is the sampling bias that naturally arises from AOD retrievals, whether from ground- or satellite-based sensors. As no retrievals are performed in cloudy pixels, any model trained on this data is only aware of the dynamics of clear-sky AOD, and we cannot directly validate estimations over cloudy-sky areas (not even with ground truth AERONET, as AERONET stations cannot perform retrievals during cloudy conditions either). However, because this research aims to use AOD to estimate $PM_{2.5}$, and hydrological effects tend to lower aerosol concentrations (i.e., via washout from precipitation), the methods can be justified as providing an upper-bound (acute) estimate for exposure.
- A significant focus for future work is how best to incorporate new data from the novel Plankton Aerosol Clouds Ecosystem (PACE) mission. The at-launch (February 2024) application of DB to PACE Ocean Color Instrument (OCI) was to that of MODIS and VIIRS, except at finer spatial resolution; however, due to increased spectral coverage, it is anticipated that improved aerosol-type products (e.g., fine-resolution UV absorbing aerosol index, fine mode fraction) and an aerosol height product may become available.

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.

Bibliography

Ackerman, S. A., Strabala, K. I., Menzel, W. P., Frey, R. A., Moeller, C. C., & Gumley, L. E. (1998). Discriminating clear sky from clouds with MODIS. *Journal of Geophysical Research: Atmospheres, 103*(D24), 32141–32157. https://doi.org/10.1029/1998JD200032

Albawi, S., Mohammed, T. A., & Al-Zawi, S. (2017). Understanding of a convolutional neural network. In *2017 International Conference on Engineering and Technology (ICET)* (pp. 1–6). https://doi.org/10.1109/ICEngTechnol.2017.8308186

Albrecht, B. A. (1989). Aerosols, Cloud Microphysics, and Fractional Cloudiness. *Science, 245*(4923), 1227–1230. https://doi.org/10.1126/science.245.4923.1227

Alman, B. L., Pfister, G., Hao, H., Stowell, J., Hu, X., Liu, Y., & Strickland, M. J. (2016). The association of wildfire smoke with respiratory and cardiovascular emergency department visits in Colorado in 2012: a case crossover study. *Environmental Health, 15*, 64. https://doi.org/10.1186/s12940-016-0146-8

Altaratz, O., Koren, I., Remer, L. A., & Hirsch, E. (2014). Review: Cloud invigoration by aerosols—Coupling between microphysics and dynamics. *Atmospheric Research, 140–141*, 38–60. https://doi.org/10.1016/j.atmosres.2014.01.009

Anderson, E. L., Turnham, P., Griffin, J. R., & Clarke, C. C. (2020). Consideration of the Aerosol Transmission for COVID-19 and Public Health. *Risk Analysis*, *40*(5), 902–907. https://doi.org/10.1111/risa.13500

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

Angstrom, A. (1929). On the Atmospheric Transmission of Sun Radiation and on Dust in the Air. *Geografiska Annaler*, *11*, 156–166.

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). 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*, *110*(16), 6442–6447. https://doi.org/10.1073/pnas.1211466110

Atwater, M. A. (1970). Planetary Albedo Changes Due to Aerosols. *Science*, *170*(3953), 64–66. https://doi.org/10.1126/science.170.3953.64

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

Barbaro, E., Vilà-Guerau de Arellano, J., Krol, M. C., & Holtslag, A. A. M. (2013). Impacts of Aerosol Shortwave Radiation Absorption on the Dynamics of an Idealized Convective Atmospheric Boundary Layer. *Boundary-Layer Meteorology*, *148*(1), 31–49. https://doi.org/10.1007/s10546-013-9800-7

Barron, A. R. (1994). Approximation and estimation bounds for artificial neural networks. *Machine Learning*, *14*(1), 115–133. https://doi.org/10.1007/BF00993164

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

Bocquet, M., Elbern, H., Eskes, H., Hirtl, M., Žabkar, R., Carmichael, G. R., et al. (2015). Data assimilation in atmospheric chemistry models: current status and future prospects for coupled chemistry meteorology models. *Atmospheric Chemistry and Physics*, *15*(10), 5325–5358. https://doi.org/10.5194/acp-15-5325-2015

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). https://doi.org/10.1029/2006GB002840

Boucher, O., Randall, D., Artaxo, P., Bretherton, C., Feingold, G., Forster, P., et al. (2013). Clouds and Aerosols. In *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (pp. 571–657). Cambridge University Press. Retrieved from https://pure.mpg.de/pubman/faces/ViewItemOverviewPage.jsp?itemId=item_2007900

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.0b013e3181dbece1

Chai, J., Zeng, H., Li, A., & Ngai, E. W. T. (2021). Deep learning in computer vision: A critical review of emerging techniques and application scenarios. *Machine Learning with Applications, 6*, 100134. https://doi.org/10.1016/j.mlwa.2021.100134

Chakrabarty, R. K., Moosmüller, H., Chen, L.-W. A., Lewis, K., Arnott, W. P., Mazzoleni, C., et al. (2010). Brown carbon in tar balls from smoldering biomass combustion. *Atmospheric Chemistry and Physics*, *10*(13), 6363–6370. https://doi.org/10.5194/acp-10-6363-2010

Chang, H. H., Hu, X., & Liu, Y. (2014). Calibrating MODIS aerosol optical depth for predicting daily PM2.5 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., & Pilat, M. J. (1969). Climate: The Influence of Aerosols. *Journal of Applied Meteorology*, *8*(6), 1001–1002. https://doi.org/10.1175/1520-0450(1969)008<1001:CTIOA>2.0.CO;2

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, February 8). TransUNet: Transformers Make Strong Encoders for Medical Image Segmentation. arXiv. https://doi.org/10.48550/arXiv.2102.04306

Chen, L.-W. A., Moosmüller, H., Arnott, W. P., Chow, J. C., Watson, J. G., Susott, R. A., et al. (2007). Emissions from Laboratory Combustion of Wildland Fuels: Emission Factors and Source Profiles. *Environmental Science & Technology*, *41*(12), 4317–4325. https://doi.org/10.1021/es062364i

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

Choudhary, K., DeCost, B., Chen, C., Jain, A., Tavazza, F., Cohn, R., et al. (2022). Recent advances and applications of deep learning methods in materials science. *Npj Computational Materials*, *8*(1), 1–26. https://doi.org/10.1038/s41524-022-00734-6

Chow, J. C., Watson, J. G., Chen, L.-W. A., Arnott, W. P., Moosmüller, H., & Fung, K. (2004). Equivalence of Elemental Carbon by Thermal/Optical Reflectance and Transmittance with Different Temperature Protocols. *Environmental Science & Technology*, *38*(16), 4414–4422. https://doi.org/10.1021/es034936u

Chu, D. A., Kaufman, Y. J., Ichoku, C., Remer, L. A., Tanré, D., & Holben, B. N. (2002). Validation of MODIS aerosol optical depth retrieval over land. *Geophysical Research Letters*, *29*(12), MOD2-1. https://doi.org/10.1029/2001GL013205

Coakley, J. A., Cess, R. D., & Yurevich, F. B. (1983). The Effect of Tropospheric Aerosols on the Earth's Radiation Budget: A Parameterization for Climate Models. *Journal of the Atmospheric Sciences*, *40*(1), 116–138. https://doi.org/10.1175/1520-0469(1983)040<0116:TEOTAO>2.0.CO;2

Collobert, R., Weston, J., Bottou, L., Karlen, M., Kavukcuoglu, K., & Kuksa, P. (2011). Natural Language Processing (Almost) from Scratch. *The Journal of Machine Learning Research*, *12*(null), 2493–2537.

Cybenko, G. (1989). Approximation by superpositions of a sigmoidal function. *Mathematics of Control, Signals and Systems*, *2*(4), 303–314. https://doi.org/10.1007/BF02551274

DMSP Operational Linescan System (OLS). (2022, October 12). Retrieved February 9, 2024, from https://www.ncei.noaa.gov/products/dmsp-operational-linescan-system

Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., et al. (2021, June 3). An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale. arXiv. https://doi.org/10.48550/arXiv.2010.11929

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: Atmospheres*, *105*(D8), 9791–9806. https://doi.org/10.1029/2000JD900040

Dubovik, Oleg, 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: Atmospheres*, *111*(D11). https://doi.org/10.1029/2005JD006619

Funahashi, K.-I. (1989). On the approximate realization of continuous mappings by neural networks. *Neural Networks*, *2*(3), 183–192. https://doi.org/10.1016/0893-6080(89)90003-8

Gao, B.-C., Kaufman, Y. J., Tanre, D., & Li, R.-R. (2002). Distinguishing tropospheric aerosols from thin cirrus clouds for improved aerosol retrievals using the ratio of 1.38-µm and 1.24-µm channels. *Geophysical Research Letters*, *29*(18), 36-1-36–4. https://doi.org/10.1029/2002GL015475

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

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: Atmospheres*, *94*(D2), 2197–2206. https://doi.org/10.1029/JD094iD02p02197

Gonzalez-Alonso, L., Val Martin, M., & Kahn, R. A. (2019). Biomass-burning smoke heights over the Amazon observed from space. *Atmospheric Chemistry and Physics*, *19*(3), 1685–1702. https://doi.org/10.5194/acp-19-1685-2019

Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., et al. (2014, June 10). Generative Adversarial Networks. arXiv. https://doi.org/10.48550/arXiv.1406.2661

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

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

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

Holmes, H. A., Zhai, X., Redman, J., Digby, K., Ivey, C., Balachandran, S., et al. (2014). Improved Spatiotemporal Source-Based Air Pollutant Mixture Characterization for Health Studies. In D. Steyn & R. Mathur (Eds.), *Air Pollution Modeling and its*

Application XXIII (pp. 25–30). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-04379-1_4

Hornik, K., Stinchcombe, M., & White, H. (1989). Multilayer feedforward networks are universal approximators. *Neural Networks*, *2*(5), 359–366. https://doi.org/10.1016/0893-6080(89)90020-8

Hsu, N. C. (2017). Changes to MODIS Deep Blue aerosol products between collection 6 and collection 6.1.

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

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

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. Y. (2016, December). *Assessing long-term climate data continuity of the Deep Blue aerosol products from MODIS to VIIRS*. Presented at the American Geophysical Union, San Francisco, CA.

Hu, X., Waller, L. A., Lyapustin, A., Wang, Y., & Liu, Y. (2014). 10-year spatial and temporal trends of $PM_{2.5}$ concentrations in the southeastern US estimated using high-resolution satellite data. *Atmospheric Chemistry and Physics*, *14*(12), 6301–6314. https://doi.org/10.5194/acp-14-6301-2014

Hu, Xuefei, Waller, L. A., Lyapustin, A., Wang, Y., Al-Hamdan, M. Z., Crosson, W. L., et al. (2014). Estimating ground-level PM2.5 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). Barcelona, Spain: IEEE. https://doi.org/10.1109/ICASSP40776.2020.9053405

Iacobellis, S., Norris, J., Tyree, M., & Daniel, R. (2009). CLIMATE VARIABILITY AND CALIFORNIA LOW-LEVEL TEMPERATURE INVERSIONS. Retrieved from https://www.semanticscholar.org/paper/CLIMATE-VARIABILITY-AND-CALIFORNIA-LOW-LEVEL-Iacobellis-Norris/5a1350c4e18bf70aaa47f0dfc1c66b51bc9cab6c

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

Information (NCEI), N. C. for E. (n.d.). North American Mesoscale Forecast System (NAM) [12 km]. Retrieved February 2, 2024, from https://www.ncei.noaa.gov/access/metadata/landing-page/bin/iso?id=gov.noaa.ncdc:C00630

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., Black, T., Pyle, M., Chuang, H., Rogers, E., & DiMego, G. (2005). The NCEP WRF Core. Boulder, CO, NCAR, 2.9.

Jeong, M.-J., Hsu, N. C., Kwiatkowska, E. J., Franz, B. A., Meister, G., & Salustro, C. E. (2011). Impacts of Cross-Platform Vicarious Calibration on the Deep Blue Aerosol Retrievals for Moderate Resolution Imaging Spectroradiometer Aboard Terra. *IEEE Transactions on Geoscience and Remote Sensing*, *49*(12), 4877–4888. https://doi.org/10.1109/TGRS.2011.2153205

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

Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. *Science*, *349*(6245), 255–260. https://doi.org/10.1126/science.aaa8415

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., 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: Atmospheres*, *102*(D14), 17051–17067. https://doi.org/10.1029/96JD03988

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: Atmospheres*, *103*(D24), 32215–32238. https://doi.org/10.1029/98JD01644

Kennedy, I. M. (2007). The health effects of combustion-generated aerosols. *Proceedings of the Combustion Institute*, *31*(2), 2757–2770. https://doi.org/10.1016/j.proci.2006.08.116

Kennedy, M., & Kopp, S. (2001). *Understanding Map Projections* (8th ed.). ESRI Press.

Khain, A., Cohen, N., Lynn, B., & Pokrovsky, A. (2008). Possible Aerosol Effects on Lightning Activity and Structure of Hurricanes. *Journal of the Atmospheric Sciences*, *65*(12), 3652–3677. https://doi.org/10.1175/2008JAS2678.1

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 & Human Biology*, *26*, 186–198. https://doi.org/10.1016/j.ehb.2017.03.006

Kingma, D. P., & Ba, J. (2017, January 29). Adam: A Method for Stochastic Optimization. arXiv. https://doi.org/10.48550/arXiv.1412.6980

Koren, I., Kaufman, Y. J., Rosenfeld, D., Remer, L. A., & Rudich, Y. (2005). Aerosol invigoration and restructuring of Atlantic convective clouds. *Geophysical Research Letters*, *32*(14). https://doi.org/10.1029/2005GL023187

Koren, I., Remer, L. A., Altaratz, O., Martins, J. V., & Davidi, A. (2010). Aerosol-induced changes of convective cloud anvils produce strong climate warming. *Atmospheric Chemistry and Physics*, *10*(10), 5001–5010. https://doi.org/10.5194/acp-10-5001-2010

Kramer, H. A., Mockrin, M. H., Alexandre, P. M., & Radeloff, V. C. (2019). High wildfire damage in interface communities in California. *International Journal of Wildland Fire, 28*(9), 641–650. https://doi.org/10.1071/WF18108

Lecun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE, 86*(11), 2278–2324. https://doi.org/10.1109/5.726791

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature, 521*(7553), 436–444. https://doi.org/10.1038/nature14539

Lee, Hoesung, 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, Hyewon, 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., Hsu, N. C., Bettenhausen, C., Sayer, A. M., Seftor, C. J., Jeong, M.-J., et al. (2016). Evaluating the Height of Biomass Burning Smoke Aerosols Retrieved from Synergistic Use of Multiple Satellite Sensors over Southeast Asia. *Aerosol and Air Quality Research*, *16*(11), 2831–2842. https://doi.org/10.4209/aaqr.2015.08.0506

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

Liang, S. (2018). Comprehensive Remote Sensing. Retrieved February 9, 2024, from http://www.sciencedirect.com:5070/referencework/9780128032213/comprehensive-remote-sensing

Lim, S. S., Vos, T., Flaxman, A. D., Danaei, G., Shibuya, K., Adair-Rohani, H., et al. (2012). A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. *The Lancet*, *380*(9859), 2224–2260. https://doi.org/10.1016/S0140-6736(12)61766-8

Lohmann, U., & Feichter, J. (2005). Global indirect aerosol effects: a review. *Atmospheric Chemistry and Physics*, *5*(3), 715–737. https://doi.org/10.5194/acp-5-715-2005

Loría-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

Loría-Salazar, S. Marcela, 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

Lu, Z., Pu, H., Wang, F., Hu, Z., & Wang, L. (2017). The Expressive Power of Neural Networks: A View from the Width. In *Advances in Neural Information Processing Systems* (Vol. 30). Curran Associates, Inc. Retrieved from https://proceedings.neurips.cc/paper_files/paper/2017/hash/32cbf687880eb1674a07bf717761dd3a-Abstract.html

Lyapustin, A., & Wang, Y. (2008, January 1). *The Time Series Technique for Aerosol Retrievals over Land from MODIS: Algorithm MAIAC*. Retrieved from https://ntrs.nasa.gov/citations/20080023335

Lyapustin, A., Martonchik, J., Wang, Y., Laszlo, I., & Korkin, S. (2011). Multiangle implementation of atmospheric correction (MAIAC): 1. Radiative transfer basis and look-up tables. *Journal of Geophysical Research: Atmospheres*, *116*(D3). https://doi.org/10.1029/2010JD014985

Lyapustin, A., Wang, Y., Laszlo, I., Kahn, R., Korkin, S., Remer, L., et al. (2011). Multiangle implementation of atmospheric correction (MAIAC): 2. Aerosol algorithm. *Journal of Geophysical Research: Atmospheres*, *116*(D3). https://doi.org/10.1029/2010JD014986

Lyapustin, A., Wang, Y., Laszlo, I., Hilker, T., G.Hall, F., Sellers, P. J., et al. (2012). Multi-angle implementation of atmospheric correction for MODIS (MAIAC): 3. Atmospheric correction. *Remote Sensing of Environment*, *127*, 385–393. https://doi.org/10.1016/j.rse.2012.09.002

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

Ma, J., Sheridan, R. P., Liaw, A., Dahl, G. E., & Svetnik, V. (2015). Deep Neural Nets as a Method for Quantitative Structure–Activity Relationships. *Journal of Chemical Information and Modeling*, *55*(2), 263–274. https://doi.org/10.1021/ci500747n

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

Martins, J. V., Tanré, D., Remer, L., Kaufman, Y., Mattoo, S., & Levy, R. (2002). MODIS Cloud screening for remote sensing of aerosols over oceans using spatial variability. *Geophysical Research Letters*, *29*(12), MOD4-1-MOD4-4. https://doi.org/10.1029/2001GL013252

Matz, C. J., Egyed, M., Xi, G., Racine, J., Pavlovic, R., Rittmaster, R., et al. (2020). Health impact analysis of PM2.5 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

Mauderly, J. L., & Chow, J. C. (2008). Health Effects of Organic Aerosols. *Inhalation Toxicology*, *20*(3), 257–288. https://doi.org/10.1080/08958370701866008

McCormick, R. A., & Ludwig, J. H. (1967). Climate Modification by Atmospheric Aerosols. *Science*, *156*(3780), 1358–1359. https://doi.org/10.1126/science.156.3780.1358

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.

Muhammad Ali, P. J., & Faraj, Rezhna H. (2014). Data Normalization and Standardization: A Technical Report. *Machine Learning Technical Reports*, 1–6.

Neale, T., & May, D. (2018). Bushfire simulators and analysis in Australia: insights into an emerging sociotechnical practice. *Environmental Hazards*, *17*(3), 200–218. https://doi.org/10.1080/17477891.2017.1410462

North American Mesoscale Forecast System. (2020, August 10). Retrieved February 24, 2024, from https://www.ncei.noaa.gov/products/weather-climate-models/north-american-mesoscale

O'Shea, K., & Nash, R. (2015, December 2). An Introduction to Convolutional Neural Networks. arXiv. https://doi.org/10.48550/arXiv.1511.08458

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., 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

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

Petty, G. W. (2006). *A First Course in Atmospheric Radiation Second Edition*. Sundog Publishing LLC.

Péwé, T. L., & Science, A. A. for the A. of. (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

Prather, K. A., Hatch, C. D., & Grassian, V. H. (2008). Analysis of Atmospheric Aerosols. *Annual Review of Analytical Chemistry*, *1*(1), 485–514. https://doi.org/10.1146/annurev.anchem.1.031207.113030

Ramanathan, V., Crutzen, P. J., Kiehl, J. T., & Rosenfeld, D. (2001). Aerosols, Climate, and the Hydrological Cycle. *Science*, *294*(5549), 2119–2124. https://doi.org/10.1126/science.1064034

Randles, C. A., Silva, A. M. da, Buchard, V., Colarco, P. R., Darmenov, 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

Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You Only Look Once: Unified, Real-Time Object Detection (pp. 779–788). Presented at the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Retrieved from https://www.cvfoundation.org/openaccess/content_cvpr_2016/html/Redmon_You_Only_Look_CVPR_2016_paper.html

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

Remer, L. A., Tanré, D., Kaufman, Y. J., Ichoku, C., Mattoo, S., Levy, R., et al. (2002). Validation of MODIS aerosol retrieval over ocean. *Geophysical Research Letters*, *29*(12), MOD3-1-MOD3-4. https://doi.org/10.1029/2001GL013204

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, May 18). U-Net: Convolutional Networks for Biomedical Image Segmentation. arXiv. https://doi.org/10.48550/arXiv.1505.04597

Rosenfeld, D., Andreae, M. O., Asmi, A., Chin, M., de Leeuw, G., Donovan, D. P., et al. (2014). Global observations of aerosol-cloud-precipitation-climate interactions. *Reviews of Geophysics*, *52*(4), 750–808. https://doi.org/10.1002/2013RG000441

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

Saharia, C., Chan, W., Chang, H., Lee, C. A., Ho, J., Salimans, T., et al. (2022, May 3). Palette: Image-to-Image Diffusion Models. arXiv. https://doi.org/10.48550/arXiv.2111.05826

Sayer, A. M., Hsu, N. C., Bettenhausen, C., Jeong, M.-J., Holben, B. N., & Zhang, J. (2012). Global and regional evaluation of over-land spectral aerosol optical depth retrievals from

SeaWiFS. *Atmospheric Measurement Techniques*, *5*(7), 1761–1778. https://doi.org/10.5194/amt-5-1761-2012

Sayer, A. M., Hsu, N. C., Bettenhausen, C., Ahmad, Z., Holben, B. N., Smirnov, A., et al. (2012). SeaWiFS Ocean Aerosol Retrieval (SOAR): Algorithm, validation, and comparison with other data sets. *Journal of Geophysical Research: Atmospheres*, *117*(D3). https://doi.org/10.1029/2011JD016599

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

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), 13,496-13,526. 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

Sayer, Andrew 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

Schraufnagel, D. E. (2020). The health effects of ultrafine particles. *Experimental & Molecular Medicine*, *52*(3), 311–317. https://doi.org/10.1038/s12276-020-0403-3

Shaw, G. E. (1983). Sun Photometry. *Bulletin of the American Meteorological Society*, *64*(1), 4–10. https://doi.org/10.1175/1520-0477(1983)064<0004:SP>2.0.CO;2

Shiraiwa, M., Ueda, K., Pozzer, A., Lammel, G., Kampf, C. J., Fushimi, A., et al. (2017). Aerosol Health Effects from Molecular to Global Scales. *Environmental Science & Technology*, *51*(23), 13545–13567. https://doi.org/10.1021/acs.est.7b04417

Shuman, J. K., Balch, J. K., Barnes, R. T., Higuera, P. E., Roos, C. I., Schwilk, D. W., et al. (2022). Reimagine fire science for the anthropocene. *PNAS Nexus*, *1*(3), pgac115. https://doi.org/10.1093/pnasnexus/pgac115

Silcox, G. D., Kelly, K. E., Crosman, E. T., Whiteman, C. D., & Allen, B. L. (2012). Wintertime PM2.5 concentrations during persistent, multi-day cold-air pools in a mountain valley. *Atmospheric Environment*, *46*, 17–24. https://doi.org/10.1016/j.atmosenv.2011.10.041

Simonyan, K., & Zisserman, A. (2015, April 10). Very Deep Convolutional Networks for Large-Scale Image Recognition. arXiv. https://doi.org/10.48550/arXiv.1409.1556

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

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

Stephens, G. L., Vane, D. G., Boain, R. J., Mace, G. G., Sassen, K., Wang, Z., et al. (2002). THE CLOUDSAT MISSION AND THE A-TRAIN: A New Dimension of Space-Based

Observations of Clouds and Precipitation. *Bulletin of the American Meteorological Society*, *83*(12), 1771–1790. https://doi.org/10.1175/BAMS-83-12-1771

Tanré, D., Herman, M., & Kaufman, Y. J. (1996). Information on aerosol size distribution contained in solar reflected spectral radiances. *Journal of Geophysical Research: Atmospheres*, *101*(D14), 19043–19060. https://doi.org/10.1029/96JD00333

Tanré, D., Kaufman, Y. J., Herman, M., & Mattoo, S. (1997a). Remote sensing of aerosol properties over oceans using the MODIS/EOS spectral radiances. *Journal of Geophysical Research: Atmospheres*, *102*(D14), 16971–16988. https://doi.org/10.1029/96JD03437

Tanré, D., Kaufman, Y. J., Herman, M., & Mattoo, S. (1997b). Remote sensing of aerosol properties over oceans using the MODIS/EOS spectral radiances. *Journal of Geophysical Research: Atmospheres*, *102*(D14), 16971–16988. https://doi.org/10.1029/96JD03437

Tanré, D., Remer, L. A., Kaufman, Y. J., Mattoo, S., Hobbs, P. V., Livingston, J. M., et al. (1999). Retrieval of aerosol optical thickness and size distribution over ocean from the MODIS airborne simulator during TARFOX. *Journal of Geophysical Research: Atmospheres*, *104*(D2), 2261–2278. https://doi.org/10.1029/1998JD200077

Trinh, T. T., Trinh, T. T., Le, T. T., Nguyen, T. D. H., & Tu, B. M. (2019). Temperature inversion and air pollution relationship, and its effects on human health in Hanoi City, Vietnam. *Environmental Geochemistry and Health*, *41*(2), 929–937. https://doi.org/10.1007/s10653-018-0190-0

Twomey, S. (1974). Pollution and the planetary albedo. *Atmospheric Environment (1967)*, *8*(12), 1251–1256. https://doi.org/10.1016/0004-6981(74)90004-3

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

Wallace, J., Corr, D., & Kanaroglou, P. (2010). Topographic and spatial impacts of temperature inversions on air quality using mobile air pollution surveys. *Science of The Total Environment*, *408*(21), 5086–5096. https://doi.org/10.1016/j.scitotenv.2010.06.020

Wang, C. (2005). A modeling study of the response of tropical deep convection to the increase of cloud condensation nuclei concentration: 1. Dynamics and microphysics. *Journal of Geophysical Research: Atmospheres*, *110*(D21). https://doi.org/10.1029/2004JD005720

Wang, Y., Khalizov, A., Levy, M., & Zhang, R. (2013). New Directions: Light absorbing aerosols and their atmospheric impacts. *Atmospheric Environment*, *81*, 713–715. https://doi.org/10.1016/j.atmosenv.2013.09.034

Wei, J., Huang, W., Li, Z., Xue, W., Peng, Y., Sun, L., & Cribb, M. (2019). Estimating 1-km-resolution PM2.5 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, X., Chang, N. B., Bai, K., & Gao, W. (2020). Satellite remote sensing of aerosol optical depth: advances, challenges, and perspectives. *Critical Reviews in Environmental Science and Technology*, *50*(16), 1640–1725. https://doi.org/10.1080/10643389.2019.1665944

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). 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

Wu, Y., Arapi, A., Huang, J., Gross, B., & Moshary, F. (2018). Intra-continental wildfire smoke transport and impact on local air quality observed by ground-based and satellite remote sensing in New York City. *Atmospheric Environment*, *187*, 266–281. https://doi.org/10.1016/j.atmosenv.2018.06.006

Zhang, Q., Meng, J., Quan, J., Gao, Y., Zhao, D., Chen, P., & He, H. (2012). Impact of aerosol composition on cloud condensation nuclei activity. *Atmospheric Chemistry and Physics*, *12*(8), 3783–3790. https://doi.org/10.5194/acp-12-3783-2012

Zhang, R., Wang, G., Guo, S., Zamora, M. L., Ying, Q., Lin, Y., et al. (2015). Formation of Urban Fine Particulate Matter. *Chemical Reviews*, *115*(10), 3803–3855. https://doi.org/10.1021/acs.chemrev.5b00067

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 & 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). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-00889-5_1

Zou, J., Han, Y., & So, S.-S. (2009). Overview of Artificial Neural Networks. In D. J. Livingstone (Ed.), *Artificial Neural Networks: Methods and Applications* (pp. 14–22). Totowa, NJ: Humana Press. https://doi.org/10.1007/978-1-60327-101-1_2