Investigation of artificial neural network performance in the aerosol properties retrieval

Nishi Srivastavaa,*, D. Vigneshb and Nisheeth Saxenab

aDepartment of Physics, Birla Institute of Technology, Mesra, Ranchi
bDepartment of Computer Science and Engineering, Birla Institute of Technology, Mesra, Ranchi
*Corresponding author. E-mail: nishi.bhu@gmail.com

ABSTRACT

Aerosols are an integral part of Earth’s climate system and their effect on climate makes this field a relevant research problem. The artificial neural network (ANN) technique is an upcoming technique in different research fields. In the current work, we have evaluated the performance of an ANN with its parameters in simulating the aerosol’s properties. ANN evaluation is performed over three sites (Kanpur, Jaipur, and Gandhi College) in the Indian region. We evaluated the performance of ANN for model’s hyperparameter (number of hidden layers) and optimizer’s hyperparameters (learning rate and number of iterations). The optical properties of aerosols from AERONET (AErosol RObotic NETwork) are used as input to ANN to estimate the aerosol optical depth (AOD) and Angstrom exponent. Results emphasized the need for optimal learning rate values and the number of iterations to get accurate results with low computational cost and to avoid overfitting. We observed a 23–25% increase in computational time with an increase in iteration. Thus, a meticulous selection of these parameters should be made for accurate estimations. The result indicates that the developed ANN can be utilized to derive AOD, which is not assessed at AERONET stations.

Key words: AERONET, aerosols, AOD, artificial neural network

HIGHLIGHTS

• In designing an ANN, we must choose the optimal number of iterations based on computational cost and quality of results.
• Our finding indicates that ANN with more hidden layers can perform reasonably well at a low number of iterations.
• The specific site may need a different set of hyperparameters for the best performance of the ANN.
• The developed ANN can be utilized to derive AOD, which is not assessed at AERONET stations

1. INTRODUCTION

Aerosols contribute a tiny fraction to the atmosphere but substantially impact the whole Earth’s climate system. Aerosols emanate from natural or anthropogenic sources and have a wide range of interactions with other components of the Earth system. The impact of aerosols on the climate system significantly changes with a change in their size and composition of aerosol (Satheesh & Srinivasan 2006). Thus, accurate measurements of aerosols’ properties are essential for the exact estimate of their impact and their interaction with other components of the climate system. The properties of aerosols have high spatial variation owing to various factors. The leading causes are chemical composition, size distribution, shape, wind speed and direction, terrain properties, relative humidity, and numerous others. The measurement of aerosols involves high levels of uncertainty and, subsequently, its impact on climate also involves a high level of uncertainty (IPCC Report 2007, 2013). The uncertainties associated with aerosol measurements and their effects on the climate make this a promising field of research. High spatial and temporal variability in aerosol distribution makes it more challenging to quantify their impacts and the associated uncertainties (Srivastava et al. 2016). The researchers have implemented various approaches to examine the properties and role of aerosol in the climate system to reduce the uncertainties (Wilcox et al. 2006; Nakajima et al. 2007; Bellouin et al. 2008; Zhang et al. 2008; Yin et al. 2015). Ground-based observations, satellite measurements, and numerical/chemical transport model simulations are frequently used techniques to study aerosol properties (Chin et al. 2009; Lu et al. 2011; Yang et al. 2017; Li et al. 2019).

This is an Open Access article distributed under the terms of the Creative Commons Attribution Licence (CC BY 4.0), which permits copying, adaptation and redistribution, provided the original work is properly cited (http://creativecommons.org/licenses/by/4.0/).
There are several properties associated with aerosols that are important to study to understand their impact on climate. But in the present work, we focus primarily on their optical properties, as their variations contribute to climatic conditions. Researchers have adopted various techniques to study and characterize the properties of the aerosol. There have been several attempts to understand the chaotic behavior of climate with the available data collected from multiple satellites and ground-based networks. In this work, we have used a new developing approach popular in various research fields, ‘artificial neural network (ANN)’ techniques, to estimate the aerosol properties. This research aims to study the performance of the ANN technique in simulating the aerosol optical properties over the Indian domain.

The wide use of the ANN can be seen in several research fields, including computer science, medical biology, economics, chemistry, physics, agriculture, and meteorology (Yang et al. 2000; Ahmed et al. 2012; Krishna 2013). Researchers are utilizing ANN in various atmospheric studies and applications (Perez & Reyes 2006; Radosavljevic et al. 2007; Luis et al. 2015; García et al. 2016; Bethania et al. 2017).

Researchers use various neural networks to simulate research problems. The primarily used neural networks are the single-layer perceptron neural network, multilayer perceptron (MLP) neural network, convolution neural network, and recurrent neural network. In the present study, we have opted for an MLP neural network, an extension of a single-layer perceptron neural network. A detailed description of the ANN is provided under the ANN calculation section.

ANN applications are also seen in the various specialized field of atmospheric science, including aerosol studied and air pollution. Radosavljevic et al. (2007) have predicted the aerosol optical depth (AOD) values with the help of neural networks. They have used the global data from MODIS (Moderate Resolution Imaging Spectroradiometer) satellite to train their network. Their findings showed that neural networks work more efficiently in comparison to the operational retrieval algorithm. Ali et al. (2013) have used neural networks in the field of data assimilations. They used a chemical transport model to produce the three-dimensional aerosol distributions over Europe. They proposed a methodology based on neural networks to compute the missing AOD data from satellites. Luis et al. (2015) have used neural networks to estimate missing AOD values at a station with the help of AOD information at other stations, ANN, and air mass trajectories. During cloudy days, it is not easy to monitor the AOD. Thus, the proposed technique is helpful for cloudy days for estimating AOD and the other optical properties. In another work, García et al. (2016) have used ANN combined with the observational data to reconstruct a time series for 73 years for AOD over a location in Spain for summer months because of high aerosol loading during this season. The AOD estimated by them was in good agreement with the observations like solar spectrometer Mark-I and AERONET (AERosol RObotic NETWORK). In another study by Bethania et al. (2017) AOD finding was improved from MODIS data with machine learning over the location that is poorly characterized and far away from the AERONET stations. They used ANN along with the support vector machines to obtain corrected AOD values taken from MODIS. In 2018, Nicolae et al. (2018) developed an ANN algorithm based on lidar data to determine the most prominent aerosol type from the multispectral lidar data. Their algorithm was able to recognize the aerosol types in the majority of cases. In recent work, Yali et al. (2020) have used ANN to predict aerosol particle distribution. They used AOD and parameters derived from the Mie theory as inputs to the ANN. Yín et al. (2020) used a convolutional neural network model to identify and classify atmospheric particles in the scanning electron microscopy images.

ANNs are also used in the assessment of air pollution at various locations in the world. Perez & Reyes (2006) developed a model with the help of ANN to forecast the maximum PM$_{10}$ concentration ahead of 1 day in the city of Santiago, Chile. They used meteorological variables and pollutants concentration as input to the model to forecast the PM$_{10}$ concentration. Gupta & Christopher (2009) also used ANN and satellite-derived products and surface-measured aerosol products to assess air quality. In a recent study, Léskiewicz et al. (2018) have studied the contribution of bio-aerosols in the degradation of air quality. In this work, they performed a real-time analysis of fingerprints of various types of bio-aerosols with the help of ANN, and they achieved a very high degree of accuracy in the real-time aerosol classification.

This research paper has emphasized the evaluation of ANN in simulating the optical properties of aerosols through the MLP approach in ANN. In this work, we have focused on the performance of ANN in estimating aerosol properties with variation in their hyperparameters. Optical properties of aerosols from AEROENT are used as input and to train the neural network to estimate AOD and Angstrom exponent (AE). AERONET network provides information regarding the aerosol properties on the local sites over the various locations of the globe.

In the present work, the developed ANN is designed to achieve two significant optical parameters of aerosols (i.e., AOD and AE) as output. Both of these parameters are crucial in determining the climatic effect of aerosols. AOD describes the extent to which incoming solar radiation gets attenuated by aerosols in the atmosphere. On the other hand, AE gives a
qualitative measure of the particle size of various aerosols. AE is also used to derive other essential parameters associated with aerosols (Schuster et al. 2006).

The present work is structured in the following way: in Section 2, we have given a detailed description of the data used in the work. In Section 3, we have described ANN briefly and the route path to perform the present work. The following section gives detail about the findings of the work, and lastly, in the conclusion section, we have discussed the significant outcome of our work.

2. DATA AND SITE DESCRIPTION

We have evaluated the ANN’s performance to calculate the AOD and AE over three sites with different topography in the present work. We have taken Kanpur (26.513N, 80.232E), Jaipur (26.906N, 75.806E), and Gandhi College (25.871N, 84.128E) sites to develop and study the neural network. The Kanpur site falls in the Indo-Gangetic Plain (IGP), which is highly polluted due to various geographical, meteorological, and other reasons. Jaipur falls under the semi-arid terrain as in the vicinity of the Thar Desert, which experiences a semi-arid climate with moderate rainfall. Gandhi College is in the eastern part of IGP with a moderate climate. Owing to its rural location, most of the land is under cultivation processes and produces natural aerosols. This makes the retrieval of AOD from satellites more intricate than for an urban site like Kanpur, where land is almost consistent.

In the development of the neural network, we have used information from AERONET data as input to the network and training purposes of the ANN. To compare estimated AOD with satellite data, we have used MODIS satellite AOD data at 550 nm. Ten years of data was used to perform this exercise for Kanpur (2008–2018). On the other two sites, we have used data from 2009 to 2017 due to data unavailability for 2008 and 2018 (over Gandhi College few months data are available in 2018). Data breaks are there for Jaipur and Gandhi College stations.

AERONET is a worldwide network of radiometers to retrieve aerosol optical properties (Holben et al. 1998). This project was established worldwide by NASA and PHOTONS in collaboration with the national agency and local authority. Aerosol properties are measured with the help of the sun photometers installed at each site. The direct sun measurements are performed at eight spectral bands (340, 380, 440/441, 500, 670, 870, 940, and 1,020 nm) and sky radiance measurements at four spectral channels (440, 675, 870, and 1,020 nm). Various aerosol properties involve AOD, AE, single scattering albedo, aerosol size distribution, and various others are derived with the help of these information. Under the cloud-free conditions, the expected uncertainty in computed AOD is approximately $\pm 0.010$ to $\pm 0.021$ (Holben et al. 1998; Eck et al. 1999). Each station provides three levels of data: level 1.0 (unscreened), level 1.5 (cloud-screened and quality controlled), and level 2.0 (quality-assured). We have used level 2.0 data for our studies, a quality-assured product after cloud screening and necessary post-calibration. Details about the AERONET can be found in the introductory paper by Holben et al. (1998, 2001).

To compare the calculated AOD from ANN, we used the MODIS satellite AOD data. For this comparison, we used level-3 data from MODIS. The level-3 product is a daily aggregation of the level-2 data on a regular grid. Deep Blue and Dark Target algorithms are two well-known retrieval algorithms for aerosol retrieval from MODIS. The Deep Blue algorithm works well over bright surfaces also. In the present work, we will use the Deep Blue dataset, which uses enhanced Deep Blue algorithms (Hsu et al. 2013; Levy et al. 2013; Wei et al. 2019). MODIS sensor is an important sensor aboard the Terra satellite. MODIS plays a vital role for policymakers, as its datasets are used widely in the development of an interactive Earth system model to predict global change properly. The extensive statistical quantitative evaluation about the quality and uncertainty of the retrievals have found expected error $= [\pm (0.05 \pm 20\%)]$ (Hsu et al. 2013; Levy et al. 2013; Wei et al. 2019). As satellite data provide the gridded data, the grid on which the study sites were situated was extracted from the dataset for proper comparison.

3. ANN CALCULATIONS

This section gives a brief introduction to the ANN and outlines the methodology to perform this work. The basic architecture of the ANN involves different layers acting as the building blocks of the network, and these essential layers are the input layer, hidden layer, and output layer. A schematic representation of different layers within a neural network is depicted in Figure 1.

Each of these layers has its significance and role in forming a neural network with different functionality. Being the first layer, the input layer acts as an interface for the network. All statistical data corresponding to the required input parameters is fed to the network, followed by the hidden layer that receives signals (data) from the input layer and processes them. The
number of hidden layers and neurons present in it varies within a network and depends on the complexity of the problem under consideration. The output layer is responsible for computing feedback in data into appropriate results (Dharwal and Loveneet 2016). In the case of ANN, all neurons contribute with a specific probability decided by weights to obtain calculated results in the network. Hence, a weighted sum of input data is passed on to the next layer (hidden layer) after passing it through an activation function to resolve nonlinear problems, and further use it to limit the range of amplitude of output to a certain finite value (Karlik and Olcag 2011). Several activation functions are used in ANN, such as sigmoid, Tan hyperbolic, rectified linear unit, and many more. The right choice of activation function is required to procure an appropriate solution.

We have opted for an MLP neural network, an extension of a single-layer perceptron neural network. MLP consists of a system of interconnected neurons establishing nonlinear mapping between neurons of different layers. MLP is a well-accepted model for neural networks that estimate the output function very accurately, provided the input dataset is sufficiently large. MLP can work with two processes, namely supervised technique and unsupervised technique. In the present work, we used a supervised learning approach and a backpropagation algorithm (BPA) to monitor the performance of the network, which is also referred to as the generalized delta rule and is widely used (Heermann & Nahid 1992; Sibi et al. 2013). The training of a neural network begins with random initialization of weights generating results (Napolitano et al. 2011), which may be erroneous. The minimization of error in BPA follows the gradient descent approach. The network converges with actual results determining global minima using adaptive weights, learning rates, and momentum factors during subsequent iterations. Learning rate and the momentum factor are crucial, as they play a vital role in manipulating the weights in different layers by a specific factor to prevent the network from blowing up. These parameters also prevent the network from being trapped in local minima while minimizing error in the dataset computed by the network using BPA (Cross et al. 1995). After each iteration, the actual data are compared with calculated data to activate BPA to suffice errors prevailing in results achieving greater accuracy (Sordo 2002).

The proper functioning of the neural network is based on three essential categories of datasets. They are the training set, validation set, and test set (Perez and Reyes 2002). Each dataset has its significance and plays a vital role in enhancing the performance of the network. A training set is essential for the proper learning of the network. There exist various proportions in which each type of dataset is used in ANN. For instance, the proportion of the training set required to train the network depends on the complexity of the constructed network. The size of the training set changes based on the complexity of the problem under consideration. Larger training sets become essential for suitable network training for climate simulations that rely on the network’s adaptability to perturbations in climatic systems. Hence, in the present work, 70% of the dataset has been used as a training set to contribute to the network’s learning process.

A validation set is used to alter the hyperparameters employed in the neural network. A validation dataset is a sample of data held back from training your model used to estimate model skill while tuning the model’s hyperparameters. A few known hyperparameters are the learning rate, number of hidden layers within the network, and neurons with each layer in ANN. An optimal amount of data is required to tune hyperparameters that reflect the network’s functioning and performance. We have used nearly 20% of the total dataset as the validation dataset in the development of this ANN.

The test set of data is used for the final assessment of the network. The test set is used for fine-tuning of the network to yield satisfactory results. It acts as a supporting tool to significantly correlate the expected results and those computed by the

![Figure 1](https://iwaponline.com/jwcc/article-pdf/12/6/2814/934640/jwc0122814.pdf) | Framework of the ANN.

Downloaded from http://iwaponline.com/jwcc/article-pdf/12/6/2814/934640/jwc0122814.pdf by guest
network. Hence, it improvises over the fitting of the training set to minimize standard deviation in the acquired results. Therefore, a minimum proportion of the dataset is used for fine-tuning of the network. In our ANN, we have used nearly 10% of the total dataset for the complete training of the network.

The results obtained from the neural network significantly depend on the network features, activation function, and hyperparameters, and their selection governs the performance of any neural network. Hyperparameters are an integrated part of an ANN, as they control the learning process in an ANN. Thus, in this work, we have evaluated the performance of ANN in estimating the aerosol parameters with different hyperparameters. Model’s hyperparameter (number of hidden layers) and optimizer’s hyperparameters (learning rate and number of iterations) are varied to evaluate the performance of ANN. A brief outline and parameters used in the work are as follows:

- **Hidden layer**: Simulations are performed with two and three hidden layers to observe the difference in the performance of ANN.
- **Number of iterations**: Next, we have changed the number of iterations (i.e., 150, 250, 500, and 750) in the evaluation process.
- **Learning rate**: ANN performance is also evaluated with different learning rates. The learning rate varied from 0.5 to 2.5 at an interval of 0.5.
- **Activation function**: In the present work, we have used the sigmoid activation function in each layer.
- **The neurons in the hidden layer**: In each hidden layer, we have used five neurons; thus, the neurons varied from 10 to 15.

### 4. RESULTS AND DISCUSSION

This section discusses the results of the evaluation of the ANN with its hyperparameters in estimating aerosol optical properties. First, we discuss the results obtained by altering the optimizer’s and model’s hyperparameters. Next, we address the effect of change in inputs and, finally, the simulation of AOD at different wavelengths and AE.

#### 4.1. Quantification of moderation of optimizer’s and model’s hyperparameters on ANN performance

This section evaluated the effects of alteration in the optimizer’s hyperparameters (i.e., learning rate and the number of iterations) along with the change in the model’s hyperparameters (number of hidden layers) on the performance of ANN. First, we have changed the number of iterations and simulated the AOD with two and three layers in ANN.

We estimated AOD at 440 nm from ANN with the different number of iterations (150, 250, 500, and 750), while the learning rate was kept (1.5) for all simulations. We have assessed the performance with two and three hidden layered ANN at three study locations (Figure 2(A) and 2(B)). An increase in the number of iterations increases the computational time and cost. With a change in the number of iterations, ultimately, computational cost changes. Thus, it is essential to observe the actual change that occurred in the computational time.

Statistical comparison of calculated AOD at 440 nm for observed and estimated AOD is shown with the help of Taylor diagram (Figure 2(A) and 2(B)). Figure 2(A) gives two hidden layered ANN results, while Figure 2(B) gives results for three hidden layered ANN. In both figures, (a) represents results for Kanpur, (b) is for Jaipur, and (c) for Gandhi College. Details about the legend used in the figure are provided in Table 1. For cost-effective ANN, we should compare the improvement in results with the change in computational cost.

Two hidden layered ANN perform best for 250 iterations, while three hidden layered ANN perform best at 150 iterations. ANN showed the highest correlation over the Kanpur site (∼0.9) with observed data with these iteration numbers. For three-layered ANN for the other two sites, correlation was the highest for this iteration number (150), but it was reduced (∼0.6) compared to Kanpur. Similarly, for two-layered ANN, correlation decreased ∼0.4 over Jaipur and ∼0.6 over Gandhi College. It is clear from the plot that with other numbers of iterations, the performance of both ANN (two-/three-layered) decreased drastically. Over Kanpur, the impact of change in the number of iterations was less. Over Gandhi College, ANN performance was affected significantly by these changes, but drastic changes were noticed in the performance of ANN over Jaipur. These results showed overfitting with the increase in the number of iterations. Overfitting is a crucial issue in supervised machine learning, primarily due to the presence of noise and complexity of classifiers (Xue 2019).

Along with overfitting, computational time also increased significantly with an increase in the number of iterations. An increase in computational time leads to a significant increase in the computation cost. On increasing the number of iterations...
Figure 2 | Performance of two-layered neural network (A) and three-layered neural network (B) in AOD (440 nm) estimation with varying number of iterations: (a) Kanpur, (b) Jaipur, and (c) Gandhi College (legends explained in Table 1). (Continued).
Figure 2 | (Continued).
from 150 to 750, about a 23% increase was observed in the computational time with three hidden layers, while a 25% increase was observed in two hidden layered ANN.

This result emphasizes the need for the proper selection of the number of iterations, as an inaccurate number of iterations could lead to overfitting results and increase the computational cost. These results also indicate that higher hidden layered ANN can perform better with fewer iterations. Thus, three hidden layered ANN are more cost-effective as compared to two hidden layered ANN.

Second, we evaluated the impact of learning rate moderation on the performance of the ANN. We have varied learning rates from 0.5 to 2.5 with an interval of 0.5 with a fixed number of iterations (150) and studied the change in ANN output (Figure 3). Figure 3(A) gives the results (AOD 440 nm) with two hidden layered ANN, while Figure 3(B) gives three hidden layered ANN results. In both figures, (a) is Kanpur, (b) Jaipur, and (c) Gandhi College, and the legends used in this figure are detailed in Table 2.

The performance of two-layered ANN varied from one station to another with a change in learning rate. Two-layered ANN performed better with the 1.0 learning rate over Kanpur, where it showed about 0.8 correlation, while over Gandhi College it performed better with 1.5 learning rates. Over Jaipur, two-layered ANN performed poorly with all learning rates. With the increase in the learning rate, the performance decreases drastically over all the sites. ANN with three hidden layers performs differently than two hidden layered ANN. In three-layered ANN, model performance was nearly consistent for three sites. The best results are obtained with a 1.5 learning rate for three-layered ANN. Over Kanpur, the ANN-estimated values correlated 0.9 with the observed one, RMS was 0.45, and standard deviation was 0.70. For Jaipur and Gandhi College, ANN-calculated values had 0.6 correlation with observed value, and RMS and standard deviation were about 0.75. Results indicate that three-layered ANN performance varies from one learning rate to another. ANN performance was inferior with a high learning rate (i.e., with 2.5). ANN performed best with a moderate learning rate, i.e., 1.5, while its performance was average over all stations with other rates. These results indicate that for a particular ANN, we need to check ANN performance with different learning rates, as model performance varied significantly with the learning rate. It is difficult to state that ANN will perform better for a high learning rate or a low learning rate.

From the above results, we can also clearly observe that three-layered ANN performance was superior compared to two-layered ANN. In view of these results, for further study, we have continued with a three-layered ANN, 150 iterations, and a 1.5 learning rate.

### 4.2. Sensitivity of ANN on input parameters

In this section, we have studied ANN performance with inputs. These simulations are performed with ANN with three hidden layers, 150 iterations, and a learning rate of 1.5. Figure 4 shows the statistical variation of ANN-estimated AOD at 440 nm with two and five inputs with observed AOD from AERONET over all the three sites. With increased inputs to the ANN, the model’s performance has increased concerning all statistical parameters; standard deviation is closer to reality, and root mean square errors are also reduced significantly. This change indicates that the new input parameters taken into ANN contribute considerably to the estimation of the AOD. This result also expresses that our selection of input parameters is appropriate for the work. In this result, the same ANN also performed differently at different stations.

### 4.3. ANN-estimated AOD and AE

After selecting all suitable hyperparameters that lead to the best performance, we have estimated the AOD and AE with this ANN over the study sites during the study period. These results are simulated with three hidden layered ANN with five inputs, 150 iterations, and a 1.5 learning rate. We have studied the AOD at 440, 500, 675 nm, and the AE for 440–675 nm. Figures 5–7
Figure 3 | Performance of two-layered neural network (A) and three-layered neural network (B) in AOD (440 nm) estimation with varying learning rate: (a) Kanpur, (b) Jaipur, and (c) Gandhi College (legends explained in Table 2). (Continued).
Figure 3 | (Continued).
show the time series of actual AOD with ANN-estimated AOD at Kanpur, Jaipur, and Gandhi College. Figure 8 shows the statistical information for the estimated AOD (500 nm) from ANN at all three sites. Time-series plots clearly show that the ANN with the above-explained specifications has generated the station-measured AOD with reasonable accuracy. These plots showed that ANN could capture the pattern of variation of AOD with some discrepancies. Careful analysis of time series indicates that at 440 nm, ANN captured the variation for all sites with some underestimations in the winter months. For some months, ANN overestimated the AOD values as well. The same pattern was also followed for AOD at 500 nm for all stations. At 675 nm wavelength, the estimated AOD was very close to observed AOD for Kanpur and Gandhi College, but for Jaipur station, ANN underestimated the values for most months. As all the sites have different dominating aerosols, and change in the aerosol system caused a difference in the performance of ANN. Kanpur and Gandhi College stations fall in the IGP region where fine mode aerosols dominate, while over Jaipur, coarse mode aerosols dominate. Thus, this result indicates that ANN performed better in capturing the fine mode dominated AOD than coarse mode dominated AOD.

Figure 8 shows the scatter plot between the estimated and observed AOD at 500 nm for all three locations. These plots also provide information regarding the $R^2$ values, and these values for Kanpur were 0.52, Jaipur (0.25), and Gandhi College (0.62). $R^2$ is a goodness-of-fit measure for linear regression models and gives information regarding the estimated values’ closeness to reality. From time-series plots, it is clear that ANN was able to capture the variation of AODs. Still, the estimated values departed from actual values occasionally, i.e., for a few months underestimated and a few overestimated. The correlation between the estimated and observed AOD for three wavelengths for all sites is given in Table 3. For Kanpur, ANN performed best, while over the other two stations, ANN performance was slightly inferior compared to Kanpur.

Next, we have studied AE with the help of ANN, and variation of observed and estimated AE is shown in Figure 9. ANN was able to capture the pattern of AE variation for all locations, but the values were underestimated for most of the months. ANN performance can also be judged by the correlation values (shown in Table 3), where a high correlation was seen for all three sites. Similar to correlation values, high $R^2$ values were also observed for ANN-estimated AE. $R^2$ values were for Kanpur (0.72), Jaipur (0.60), and Gandhi College (0.68).

### 4.4. Comparison of ANN-derived AOD with MODIS-measured AOD

In the last part of the work, we have compared estimated AOD with satellite data to strengthen the accuracy of the ANN. For this, we have compared the ANN capability in estimating the AOD at 550 nm for all the sites i.e. Kanpur, Jaipur and Gandhi College (Figure 10). We have estimated AOD at 550 nm as most of the AOD measurements are performed at this wavelength, and thus it is helpful to check model performance at this wavelength. AERONET station does not measure the AOD at 550 nm, and thus, we have compared the ANN-estimated AOD at 550 nm with MODIS-measured AOD. The AOD at 550 nm has been derived with AOD values at 440, 675 nm, and AE for 440–675 nm values found from the ANN calculations. For these parameters, simulation was performed with three hidden layers, 150 iterations, and 1.5 as the learning rate. Estimated AOD compares well with observed AOD from MODIS at 550 nm. Though ANN was not calculating AOD at 550 nm, indirectly, we can utilize this to estimate AOD at other wavelengths. A high correlation was observed in estimated and satellite-observed AOD at 550 nm (Table 3).
Figure 4 | Performance of three-layered ANN with two and five inputs (number of iterations: 150; learning rate: 1.5): (a) Kanpur, (b) Jaipur, and (c) Gandhi College.
Figure 5 | Comparison of ANN-estimated AOD with AERONET-observed AOD at Kanpur for wavelength: (a) 440, (b) 500, and (c) 675 nm.
Figure 6 | Comparison of ANN-estimated AOD with AERONET-observed AOD at Jaipur for wavelength: (a) 440, (b) 500, and (c) 675 nm.
Figure 7 | Comparison of ANN-estimated AOD with AERONET-observed AOD at Gandhi College for wavelength (a) 440, (b) 500, and (c) 675 nm.
Figure 8 | Comparison of ANN-estimated AOD and AERONET-observed AOD at 500 nm: (a) Kanpur, (b) Jaipur, and (c) Gandhi College.

Table 3 | Correlation between ANN-estimated and AERONET/MODIS-measured AOD at various wavelengths and AE (440–675 nm)

<table>
<thead>
<tr>
<th>Station</th>
<th>AOD (440 nm)</th>
<th>AOD (500 nm)</th>
<th>AOD (675 nm)</th>
<th>AE (440–675 nm)</th>
<th>AOD (550 nm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kanpur</td>
<td>0.90</td>
<td>0.72</td>
<td>0.94</td>
<td>0.85</td>
<td>0.64</td>
</tr>
<tr>
<td>Jaipur</td>
<td>0.58</td>
<td>0.50</td>
<td>0.91</td>
<td>0.78</td>
<td>0.62</td>
</tr>
<tr>
<td>Gandhi College</td>
<td>0.62</td>
<td>0.73</td>
<td>0.79</td>
<td>0.83</td>
<td>0.55</td>
</tr>
</tbody>
</table>
Figure 9 | Comparison of ANN-estimated AE (440–675 nm) with AERONET-observed AOD: (a) Kanpur, (b) Jaipur, and (c) Gandhi College.
Figure 10 | Comparison of ANN estimated AOD with MODIS estimated AOD at 550 nm: (a) Kanpur, (b) Jaipur, and (c) Gandhi College.
5. CONCLUSIONS

In the present work, we performed an ANN-based sensitivity analysis to estimate aerosol optical properties over stations in a highly polluted Indo-Gangetic Basin (Kanpur and Gandhi College) and on a site in a semi-arid region (Jaipur). The significant findings from this work are listed below:

• The varying number of iterations showed that with an increased number of iterations, overfitting of results was observed for all the sites. Simultaneously, the computational time has also increased significantly (~25%). This result infers that increasing the number of iterations may not always lead to better performance of ANN, as in this case, where increasing the number of iterations leads to overfitting. Thus, the selection of this hyperparameter is crucial, as an increased number of iterations increases the computation time/cost. Accordingly, we must choose an optimal number of iterations based on computational cost and quality of results.

• We also evaluated the performance of the developed ANN with varying learning rates. Results indicate that the effect of change in learning rate varies with the number of hidden layers. With alteration in the hidden layers, the performance of ANN changed with the same learning rate. ANN with two hidden layers performed well at a low learning rate, while ANN with three hidden layers performed well at a moderate learning rate.

• The performance of ANN depends on the number of hidden layers in the ANN. Our finding indicates that ANN with more hidden layers can perform reasonably well at a low number of iterations. With fewer hidden layers, we have to increase the number of iterations to better estimate results.

• Input parameters to the ANN have a vital influence on the accuracy of the results. The precision of results increased with an increase in inputs to the ANN. This result also indicates an accurate selection of input parameters, as results have shown improvement in all statistical parameters depicted in the Taylor diagram.

• Simulation results indicated that the AOD and AE were well simulated with the developed ANN, though the performance of ANN varied from site to site. ANN performance was best over Kanpur, then Gandhi College, and least at Jaipur. This result indicates that with a change in the location, the aerosol system changes drastically, and thus the same ANN may not perform well for all the places with the same accuracy. Therefore, a specific site may need a different set of hyperparameters for the best performance of the ANN.

• Finally, we have compared the calculated AOD with MODIS-measured AOD at 550 nm, and the result indicated a reasonable estimation of the AOD with ANN. As AOD at 550 nm was indirectly estimated from ANN, this result suggests that the developed ANN can be utilized to derive AOD, which is not assessed at AERONET stations.

ACKNOWLEDGEMENTS

The authors thank Brent Holben, NASA GSFC, for providing AERONET data at Kanpur, Jaipur, and Gandhi College, and the site PI & staff for their effort in establishing and maintaining the site. We also thank NASA for MODIS level-3 AOD data.

DATA AVAILABILITY STATEMENT

All relevant data are included in the paper or its Supplementary Information.

REFERENCES


First received 16 November 2020; accepted in revised form 11 May 2021. Available online 16 June 2021.