

Analytical Prediction of Scattering Properties of Spheroidal Dust Particles With Machine Learning

Xi Chen, Jun Wang, Joe Gomes, Oleg Dubovik, Ping Yang, Masanori Saito

▶ **To cite this version:**

Xi Chen, Jun Wang, Joe Gomes, Oleg Dubovik, Ping Yang, et al.. Analytical Prediction of Scattering Properties of Spheroidal Dust Particles With Machine Learning. Geophysical Research Letters, 2022, 49, 10.1029/2021GL097548 . insu-03686316

HAL Id: insu-03686316

https://insu.hal.science/insu-03686316

Submitted on 2 Jun 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - ShareAlike 4.0 International License

Check for updates

Geophysical Research Letters

RESEARCH LETTER
10.1029/2021GL097548

Key Points:

- A neural network (NN) model is trained to characterize the optical properties of spheroidal dust particles
- Analytical Jacobians of the optical properties with respect to microphysical parameters are derived from NN model directly
- Adding analytical Jacobians from a linearized T-matrix model in the training improves the NN-derived Jacobians

Supporting Information:

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

Correspondence to:

J. Wang and J. Gomes,
jun-wang-1@uiowa.edu;
joe-gomes@uiowa.edu

Citation:

Chen, X., Wang, J., Gomes, J., Dubovik, O., Yang, P., & Saito, M. (2022). Analytical prediction of scattering properties of spheroidal dust particles with machine learning. *Geophysical Research Letters*, *49*, e2021GL097548. https://doi.org/10.1029/2021GL097548

Received 9 DEC 2021
Accepted 11 APR 2022

Author Contributions:

Conceptualization: Xi Chen, Jun Wang, Joe Gomes
Data curation: Xi Chen, Jun Wang, Oleg Dubovik, Ping Yang, Masanori Saito
Formal analysis: Xi Chen, Jun Wang, Joe Gomes
Funding acquisition: Jun Wang
Investigation: Xi Chen, Jun Wang, Joe Gomes
Methodology: Xi Chen, Jun Wang, Joe Gomes
Project Administration: Jun Wang
Resources: Xi Chen, Jun Wang

© 2022. The Authors.
This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

Analytical Prediction of Scattering Properties of Spheroidal Dust Particles With Machine Learning

Xi Chen[1,2], Jun Wang[1,2,3], Joe Gomes[1], Oleg Dubovik[4], Ping Yang[5,6,7], and Masanori Saito[5]

[1]Department of Chemical and Biochemical Engineering, Iowa Technology Institute, The University of Iowa, Iowa City, IA, USA, [2]Center for Global and Regional Environmental Research, The University of Iowa, Iowa City, IA, USA, [3]Department of Physics and Astronomy, The University of Iowa, Iowa City, IA, USA, [4]Laboratoire d'Optique Atmosphérique, CNRS–Université de Lille, Lille, France, [5]Department of Atmospheric Sciences, Texas A&M University, College Station, TX, USA, [6]Department of Oceanography, Texas A&M University, College Station, TX, USA, [7]Department of Physics and Astronomy, Texas A&M University, College Station, TX, USA

Abstract A neural network (NN) model is trained with a database widely used in the aerosol remote sensing community to rapidly predict the single-scattering optical properties of spheroidal dust particles. Analytical solutions for their Jacobians with respect to microphysical properties are derived based on the functional form of the NN. The Jacobian predictions are improved by adding Jacobians from a linearized T-matrix model into the training. Out-of-database testing implies that NN-based predictions perform better than the business-as-usual method that interpolates optical properties from the database. Independent validation further demonstrates the efficacy of the NN-based predictions by reducing computational costs while maintaining accuracy. This work represents the first use of machine learning-based function approximation to computationally expedite the application of the existing spheroidal dust properties database; the resultant NN model can be implemented in atmospheric models and satellite retrieval algorithms with high accuracy, computational efficiency, and the rigor of analytical solutions.

Plain Language Summary Dust particles affect both solar and terrestrial radiative transfer, but whether they cool or warm the climate is currently an open question in the literature. Accurate estimation of dust scattering and absorption properties, while critical for climate studies, is hindered by the fact that dust particles have irregular shapes and large size ranges; hence, no single method can be applied for all particle sizes and shapes. Often, a comprehensive look-up table of these properties is created by combining multiple methods. The application of such databases, however, is cumbersome and inaccurate due to the need for multi-variable interpolation. Furthermore, the look-up-table approach lacks the mathematical rigor needed to determine the sensitivity (Jacobians) of the single-scattering properties to the dust size, shape, and refractive index that are needed in remote sensing algorithms. The aforementioned challenges are tackled here by developing a novel approach within the neural network (NN) framework. This NN-based approach is fast, accurate, and able to predict Jacobians with analytical formulas. The NN model can be readily applied to the dust retrieval algorithm and radiative forcing modeling. The concept of deriving Jacobians from the NN model in this study can also be generalized for application to other problems involving gradient calculations.

1. Introduction

Dust aerosol has a widespread impact on air quality and visibility (Huang et al., 2008; Moulin et al., 1998) and modulates the Earth's radiation budget (Pachauri et al., 2014) via scattering and absorbing both shortwave and thermal infrared radiation (Xu et al., 2017), leading to large uncertainties in climate projection. The radiative forcing and spectral fingerprints of dust particles depend on their bulk optical properties, including the extinction coefficient, single-scattering albedo, and phase matrix. While these properties can be computed exactly based on the Lorenz-Mie theory for spherical particles (Mishchenko et al., 2002), most mineral dust particles have highly irregular shapes with a wide range of particle sizes, posing significant challenges to both remote sensing and climate modeling (Wang et al., 2003, 2004).

Various techniques have been developed for estimating non-spherical particle properties, such as the discrete dipole approximation (DDA; Draine & Flatau, 1994; Yurkin et al., 2007), the extended boundary condition method (EBCM) as an implementation of the T-matrix (Mishchenko et al., 1997; Mishchenko & Travis, 1994),

Software: Xi Chen, Jun Wang, Joe Gomes
Supervision: Jun Wang, Joe Gomes
Validation: Xi Chen, Jun Wang, Joe Gomes
Visualization: Xi Chen, Jun Wang, Joe Gomes
Writing – original draft: Xi Chen, Jun Wang
Writing – review & editing: Xi Chen, Jun Wang, Joe Gomes, Oleg Dubovik, Ping Yang, Masanori Saito

the invariant imbedding T-matrix method (IITM; Bi & Yang, 2014; Johnson, 1988; Yang et al., 2000, 2019), the finite-difference time domain (Kane, 1966; Sun et al., 1999; Yang et al., 2000; Yang & Liou, 1996a), and the physical-geometric optics method (PGOM; Yang & Liou, 1996b). Collectively, each of these models has inherent limitations on certain particle sizes or shapes. For example, in the case of spheroidal particles with an aspect ratio of 2, EBCM is applicable to size parameters ($x_s = \frac{2\pi r}{\lambda}$, r is particle radius and λ is incident light wavelength) up to approximately 110, while IITM is applicable to size parameters up to 500. Hence, past works have combined different techniques to generate a large database, in the form of look-up tables (LUTs), by assuming the particles to be axially symmetric spheroids (Dubovik et al., 2006) or three-axial ellipsoids (Meng et al., 2010). Saito et al. (2021) used a combination of IITM and PGOM to develop a dust optical property database assuming dust particles to be an ensemble of irregular hexahedral particles. In these LUTs, the dust optical properties are archived for a set of discrete microphysical parameter points (such as particle size, shape, and index of refraction). Consequently, interpolation is needed when desired input parameters do not fall at those entry points, especially in cases where the bulk optical properties are desired from integration over a given particle size distribution. Such interpolation-based application of LUT can lead to non-trivial errors due to non-linearity.

Furthermore, the optimal estimation algorithm for aerosol retrieval requires the input of not only the spectral intensity but also its derivatives with respect to the state vector (aerosol microphysical properties) of the retrieval, that is, Jacobian matrices (Rodgers, 2000). To meet this need, Wang et al. (2014) developed a Unified Linearized Vector Radiative Transfer Model that integrates a linearized Mie model and a linearized EBCM T-matrix code (R. Spurr et al., 2012) into the Vector LInearized Discrete Ordinate Radiative Transfer (R. J. D. Spurr, 2006). However, challenges remain for computing Jacobians for large non-spherical particles from linearized EBCM T-matrix code (hereafter shortened to "T-matrix model"). As a replacement, the finite difference (FD) method could approximate these Jacobians using a LUT, but the resolution of input data points in the LUT is often too coarse for the FD method. Moreover, both FD and interpolation will add a burden for computing in the retrieval, especially from the hyperspectral measurements.

We present a method based on machine learning (ML) to derive both optical properties and their respective Jacobians for spheroidal particles covering a large size range accurately and efficiently, thereby meeting the emergent requirements in both remote sensing and atmospheric modeling of dust particles. In recent years, ML methods have been increasingly applied in the field of atmospheric remote sensing. For example, a neural network (NN) and a convolutional NN are used in the classification of satellite images (Hughes & Hayes, 2014; Mohajerani et al., 2018; Saponaro et al., 2013). Furthermore, an NN can provide a fast and accurate replacement for radiative transfer model (RTM) calculation (Takenaka et al., 2011), especially for hyperspectral and multi-angle simulations, which accelerates the optimal estimation retrieval of atmospheric properties for aerosol and clouds (Chen et al., 2018; Nanda et al., 2019; Segal-Rozenhaimer et al., 2018).

This study is the first application of ML methods for aerosol particle scattering and their derivatives approximation at the same time, which differs from the past study that only predicts aerosol scattering properties with ML (Yu et al., 2022). Here, the rigor and constraints of aerosol optical properties (e.g., the normalization property of phase function) are retained to a degree at least comparable to the LUT or FD method. This approach is conducted by training an NN model with a database for spheroid mixed dust particles developed by Dubovik et al. (2002,2006) (hereafter, the Dubovik data) to derive dust single-scattering optical properties and their Jacobians to microphysical parameters, simultaneously. Moreover, the new method adds a limited number of analytically computed Jacobians from a linearized T-matrix model to the training process to improve NN Jacobians prediction. Finally, the performance of this NN model in predicting optical properties and their Jacobians is evaluated. The NN model and its training method are presented in Section 2. Data used in this study is described in Section 3, while the validation of the NN model predictions is presented in Section 4. Section 5 provides a summary and discussion.

2. Methods

NN is a machine learning model in which artificial neural networks adapt and learn in a data-driven fashion. As a universal function approximator, the feed-forward NN is ideally suited for modeling nonlinear processes. Specifically, it is used here to establish the relationships between the microphysical parameters of dust particles and corresponding single-scattering properties, as well as to derive their analytical derivatives. The description of

the dataset is given in Section 2.1. The formulation of feed-forward multi-layer NN and derivation of Jacobians are introduced in Section 2.2. The model optimization procedure is given in Section 2.3. Section 2.4 describes the tuning of NN model hyperparameters. The process to evaluate NN predictions is presented in Section 2.5.

2.1. Data and Transformation

Dubovik's LUT is used as the dataset in this study. The input features $\boldsymbol{x}$ are defined as the microphysical parameters of spheroid dust particles, including the particle effective radius (r), the real and imaginary parts of the refractive index (n and k), and scattering angle (θ). Accordingly, the output targets $\boldsymbol{y}$ are the particle single-scattering properties, including volume extinction coefficient (c_{ext}, μm^{-1}), absorption coefficient (c_{abs}), and six nonzero elements of the scattering phase matrix $\mathbf{F}$:

$$\mathbf{F} = \begin{bmatrix} F_{11} & F_{12} & 0 & 0 \\ F_{12} & F_{22} & 0 & 0 \\ 0 & 0 & F_{33} & F_{34} \\ 0 & 0 & -F_{34} & F_{44} \end{bmatrix}$$

The details about the range of features and datapoints definition in Dubovik's LUT are provided by Dubovik et al. (2002) and are briefly summarized in Table S1 and Figure S3 in the supplement. In total, there are more than 2 million datapoints for four features and eight targets in our research. 10 percent of total LUT data is randomly selected as our test data. Then, we select 20% from the remaining 90% as validation data and the balance as training data.

Before training, the input and output 2D matrices, $\mathbf{X}$ and $\mathbf{Y}$ consisting of feature and target vectors ($\boldsymbol{x}$, $\boldsymbol{y}$) for multiple data points, and the corresponding true Jacobians from the linearized T-matrix model at different training data points must be normalized or transformed due to the different magnitudes of each feature or target. Details on data distribution and transformation are provided in the supplement (Figure S3).

2.2. Feed-Forward Multi-Layer NN and Jacobian Derivation

In the NN framework, for each datapoint, g dust microphysical parameters are regarded as input features ($\boldsymbol{x} = [x_1, x_2, \ldots, x_g]$) and l single-scattering properties ($\boldsymbol{y} = [y_1, y_2, \ldots, y_l]$) are output targets. As illustrated in Figure S1, in forward propagation, the input layer signal of $\boldsymbol{x}$ is transformed to the output layer $\boldsymbol{y}$ via multiple fully connected neurons in multiple hidden layers. The output of the ith neuron in the jth layer (a_{ij}) is related to the neurons in the previous (j-1)th layer as follows:

$$a_{ij} = \sigma\left(\sum_{k=1}^{N_{j-1}} w_{ki}^{j-1} a_{k(j-1)} + b_i^j\right) \tag{1}$$

where $a_{k(j-1)}$ represents the kth neuron in the (j-1)$^{\text{th}}$ layer, w_{ki}^{j-1} is the weight (called model parameters with b_i^j) determining the connection between the neuron $a_{k(j-1)}$ and neuron a_{ij}, and b_i^j is the bias added for the neuron a_{ij}. The elementwise activation function $\sigma$is often defined as a simple but differential function to introduce non-linearity into the output of a neuron.

Following Equation 1, due to the differential nature of $\sigma$and linear expressions of neurons via w and b, the derivative of a_{ij} to $a_{k(j-1)}$ is:

$$\frac{\partial a_{ij}}{\partial a_{k(j-1)}} = w_{ki}^{j-1}\sigma'\left(\sum_{k=1}^{N_{j-1}} w_{ki}^{j-1} a_{k(j-1)} + b_i^j\right) = w_{ki}^{j-1}\sigma'(Z_{ij}) \tag{2}$$

where σ' represents the derivative of the activation function and $Z_{ij} = \sum_{k=1}^{N_{j-1}} w_{ki}^{j-1} a_{k(j-1)} + b_i^j$. Hence, via the chain rule, it is feasible to derive the gradient of each output target (y_p) to each input feature (x_q) analytically from the model parameters (w, Z, and σ') in all layers of the NN model as follows:

$$\frac{\partial y_p}{\partial x_q} = f\left(\frac{\partial a_{ij}}{\partial x_q}, w_{ki}^j, \sigma'(Z_{ij})\right) = g\left(w_{ki}^j, \sigma'(Z_{ij})\right) \quad (3)$$

The explicit expressions of function f and g are shown in the supplement. In other words, from a trained NN model, the Jacobians of single-scattering optical properties to microphysical parameters could be derived analytically by forward propagation without more training. This type of Jacobian derivation is more efficient than that of the FD method.

2.3. Model Optimization

Given a dataset of known input and output matrices $\mathbf{X}$ and $\mathbf{Y}$, training is performed to optimize the model parameters ($\mathbf{w}$ and $\mathbf{b}$) in the back propagation to minimize the cost function C_y defined as the error between NN predicted $\hat{\boldsymbol{y}}$ from forward propagation and true values of $\boldsymbol{y}$. Considering the definition of the neuron in Equation 1, C_y is differentiable with respect to the model parameters, w and b; thus, stochastic gradient descent is used as a relatively efficient optimization method.

Jacobians can be predicted analytically from NN model parameters using Equation 3, though the predictions may be inaccurate since Jacobian values are not typically optimized during training. We add true values of Jacobians in the training procedure to improve the NN predictions and the derived Jacobians simultaneously. Here, Jacobian predictions are not directly output by the NN but are obtained by evaluating Equation 3. Jacobians from analytical linearized T-matrix calculations (Spurr et al., 2012) are added as true values constituting additional constraints into the cost function:

$$C = C_y + s_1 C_{Jac_1} + s_2 C_{Jac_2} \quad (4)$$

As a result, the errors of targets and their Jacobians are minimized simultaneously during training to attain good accuracy for both of them. Here, C_{Jac_1} (or C_{Jac_2}) indicates the cost function of Jacobian $\frac{\partial y}{\partial x_1}$ (or $\frac{\partial y}{\partial x_2}$) defined by the NN prediction and true values (x_1 and x_2 represent the real and imaginary parts of the refractive index, n and $\ln k$). In the training, C_y is defined as the mean square error (MSE) of targets and $C_{Jac,i}$ is the mean absolute error (MAE) of Jacobians with respect to x_i in this study. Details about the expression of cost function C can be found in the supplement (descriptions in Table S2). s_1 and s_2 are the scaling factors of $C_{Jac,i}$ to adjust the weights of each loss function component given their different magnitudes. s_1 and s_2 are fixed in each NN model training and must be selected carefully; hence, they are regarded as hyperparameters and tuned together with others mentioned in Section 2.4. The mean relative errors of the Jacobian predictions are used as $C_{Jac,i}$ in model selection during the validation step (details in Table S2). Since multiple targets are needed for prediction, we considered a multi-task NN model during model optimization. After comparison, we find that the two-target model has better performance than the single-target model for prediction of some phase matrix elements. In conclusion, the structure of a single-target or two-target NN model for each target after hyperparameters tuning is summarized in Table S2.

2.4. Hyperparameter Optimization

In addition to the model parameters optimized by the NN training procedure, there are hyperparameters that define the model structure and control the learning process, such as the number of hidden layers (*nl*), the number of neurons in each layer (*ndim*), the activation function (σ), and the learning rate of the optimizer (*lr*). To perform hyperparameter optimization, the database is split into training data, validation data, and test data. The training data is used to optimize the model parameters. The validation data is used to perform hyperparameter selection by evaluating the trained models on a held-out set of data. The set of hyperparameters giving the lowest validation C_y is selected as optimal. The test data is used to perform a final model evaluation with the optimal hyperparameters to estimate the generalization error in the prospective use case. Here, the *tanh* activation function ($\sigma(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$) is selected empirically and is fixed while different configurations of *nl*, *ndim*, and *lr* are tested. Additional information regarding the hyperparameter optimization process can be found in the supplement.

The NN training is performed in the following steps. First, perform hyperparameter optimization and train from scratch a neural network, NN_0, using optimal *nl*, *ndim*, and *lr*. Second, define a new neural network model, NN_1,

with the same network architecture as NN_0 (the same *nl* and *ndim*) and initialize its model parameters from the pre-trained model NN_0. Third, perform hyperparameter optimization using a new cost function *C* defined as in Equation 4 by adding the T-matrix Jacobians as true Jacobians to get the optimal hyperparameters *lr*, s_1, and s_2. Finally, train the NN_1 model using these optimal hyperparameters with cost function *C*. As a result, the model NN_1 is able to predict dust single-scattering properties and their Jacobians to microphysical parameters simultaneously.

2.5. Model Evaluation

To evaluate the NN model, we compare the NN-predicted optical properties with true values in the test dataset, which is not used during training and hyperparameter tuning. The advantage of the NN model is illustrated by comparing the NN predictions and LUT-interpolated properties for test data. However, it is difficult to validate Jacobians developed from the established NN due to the lack of analytical Jacobians in Dubovik's LUT. The only way is to interpolate the property via LUT and use the FD method to approximate partial derivatives, which is not reliable or accurate, considering the data resolution of LUT and FD approximation. Therefore, more strict validation of our results is done here by using the linearized T-matrix method that can compute the analytical Jacobians for small particles (here we use a volume equivalent radius smaller than 0.03 μm). Furthermore, the linearized T-matrix is also used as the true Jacobians added in the NN_1 training as mentioned above. Considering T-matrix computation cost, only parts of the training datapoints of LUT (<5%) are calculated as training and validation data (called TJ_1 hereafter). Meanwhile, an equivalent number of data points is calculated as the test data for Jacobians validation (called TJ_2 hereafter). The Jacobians from the T-matrix model must be transformed to those in Dubovik's LUT using the kernel technique (Dubovik et al., 2006) before application.

3. Results

In this section, we first evaluate the performance of the trained NN by comparing the predicted targets with true targets for test data. Second, the Jacobians of predicted targets to input features are derived from trained NN and are validated by the Jacobians from TJ_2 data as test data (Section 3.3).

3.1. Validation of NN Targets

Inputting the features from the test data not learned by NN and comparing the predicted targets with the corresponding "true" targets can be regarded as an independent and effective method for evaluating the NN performance. Figure 1 shows the comparison of eight predicted targets from NN models (having no training inputs of Jacobians, NN_0) with the corresponding true values in the test data. The linear correlation coefficient between the two, R, is 1.0 for all targets, indicating the excellent performance of the NN model. Considering the magnitude of each target's variation, the relative error is an important criterion to quantify the accuracy of NN predictions. For example, although the root mean square error (RMSE) of 3.94e−3 in F_{33}/F_{11} prediction is smaller than RMSE of 1.36e−2 for $\ln(F_{11})$, its mean relative error (MRE; 4.6%) is larger due to the smaller magnitude of F_{33}/F_{11} true values.

The fits of c_{ext}, c_{abs}, $\ln(F_{11})$, and $\ln(F_{22})$ are better than others with smaller than 0.5% MRE, even smaller than 0.1% for $\ln(F_{11})$ and $\ln(F_{22})$ MRE. The predicted F_{33}/F_{11} and F_{44}/F_{11} do not match their true values as well as above, but they still have less than 5% MRE (4.60% and 3.16%). F_{12}/F_{11} and F_{34}/F_{11} are the most difficult to train, due in part to many small values close to zero, and their MREs are 5.41% and 9.40%, respectively. Even so, the RMSEs of F_{12}/F_{11}, F_{34}/F_{11}, F_{33}/F_{11}, and F_{44}/F_{11} are all comparable and lower than their counterparts in recent work using a deep learning model to train a similar LUT for super-spheroids aerosols (Yu et al., 2022), illustrating that our NN performs well in predicting each single-scattering property.

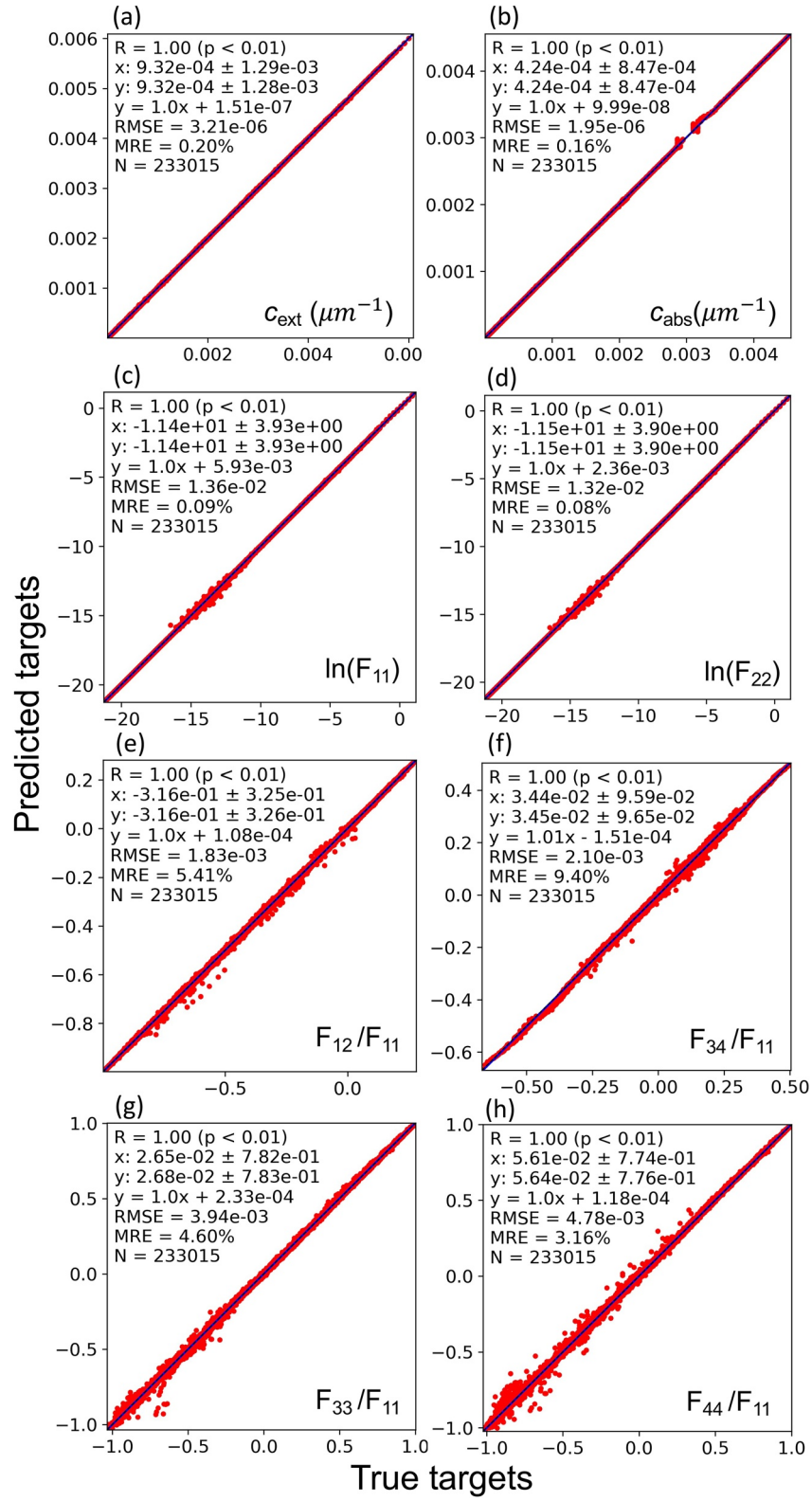


Figure 1. The scatter plots of the predicted targets and corresponding true values for test data. (a–h) are for c_{ext}, c_{abs}, ln(F_{11}), ln(F_{22}), F_{12}/F_{11}, F_{34}/F_{11}, F_{33}/F_{11}, and F_{44}/F_{11}. In each subplot, the *x*-axis represents the true value while the *y*-axis is the prediction from the neural network model. The correlation coefficient (R), the mean and standard deviation values respective for x and y, the fitting equation between x and y, the root mean square error, the mean relative error, and the total number of data used for each target are shown in each subplot. The solid black line is the one-by-one line.

3.2. Comparison With Linear Interpolation

The predicted targets from NN for test data points are compared with results interpolated linearly from the LUTs to reveal the possibility and advantages of replacing the LUTs approach with the NN model. Here, the NN model is trained from only LUT without additional Jacobians (NN_0), similar to Figure 1 and Section 3.1. To simplify, only one-dimensional linear interpolation is applied for each test data point based on the two closest data points in the rest. For this purpose, a small number of data points (<2000) are selected from the whole LUT as the test data. As a result, the true values of these test data from the original LUT are compared with those predicted using NN and linear interpolation, respectively, as shown in Figure 2. It is clear that although the predicted c_{ext}, c_{abs}, $\ln(F_{11})$, and $\ln(F_{22})$ from both methods have a strong correlation with the true values, the NN predictions present <0.2% MRE, much less than the interpolated results (1–3% MRE). The RMSE of NN predictions is also one order of magnitude smaller than those from interpolation for these four targets (Figures 2a–2d). For F_{33}/F_{11} and F_{44}/F_{11}, there are obviously many more outliers from linear interpolation compared with NN predictions, even though the correlation between interpolated and true values is strong, with 1.0 R. Both the MRE and RMSE of interpolation results are larger than the NN-predicted targets for these two targets (Figures 2g and 2h). Moreover, it is difficult to estimate F_{12}/F_{11} and F_{34}/F_{11} from linear interpolation when considering many results that are far from correct values, as demonstrated by Figures 2e and 2f. Not only are the MRE and RMSE of interpolations far larger than those for NN predictions, but the correlation coefficients of interpolation decrease to 0.98 and 0.87 for F_{12}/F_{11} and F_{34}/F_{11}, respectively. Undoubtedly, the accuracy of the interpolation depends on the grid resolution of the LUT. In other words, higher resolution results in better performance of interpolation. Based on the current resolution of this LUT (details in the supplement), linear interpolation is not suitable to approximate the single-scattering properties for unknown data points. On the other hand, greater than 1% MRE of interpolation can lead to incorrect Jacobians derived from the FD method, which requires highly accurate interpolation within small differences in input data. Even though reducing the grid interval of LUT improves the accuracy of interpolation, it is not efficient and requires computation cost and storage. The NN model, on the other hand, can provide good prediction when training from coarse-resolution data. In conclusion, the NN model shows better performance than the interpolation method, especially for data with strong nonlinearity.

Considering the range of feature values in the training data, the performance of NN to predict features outside the training data is evaluated in the supplement (Figure S5). The prediction RMSE and MRE are found to be larger than those cases that have features within the training data range, indicating that more care needs to be taken when applying this NN in the extrapolation problem. However, given that the Dubovik LUT covers the most possible range of dust spheroids parameters (Dubovik et al., 2006), our NN is robust for applications in radiative effect simulation and remote sensing of dust particles.

3.3. Validation of NN Jacobians

The NN Jacobians are validated using TJ_2 data as test data. Considering the application in remote sensing of dust microphysical properties, we mainly focus on the Jacobians with respect to refractive index (n and k) in this study. First, comparing with Jacobians predicted from the NN_0 model (without Jacobians for training), their predictions from NN_1 are more accurate by adding true Jacobians in the cost function (Section 2.3) when showing c_{ext} in Figure S4 as an example. Both the MRE and RMSE of Jacobians ($\frac{\partial c_{\text{ext}}}{\partial \ln k}$ and $\frac{\partial c_{\text{ext}}}{\partial n}$) are reduced in NN_1, while the errors of c_{ext} remain low (more details are found in the supplement). This demonstrates the feasibility and efficiency of adding the true Jacobians in the training for the highly accurate prediction of both targets and Jacobians. Then, the Jacobians of all targets predicted from NN_1 models are evaluated with TJ_2 true values in Figure 3. In general, all Jacobians predicted from NN_1 have a strong correlation with true values ($R = 1.0$), although the RMSE and MRE differ due to their different magnitudes. The Jacobians of c_{ext} and c_{abs} are predicted with the highest accuracy with MRE <1% compared with other targets' Jacobians. The Jacobians of phase matrix elements have worse prediction performance with 2–5% MRE, in which the $\frac{\partial(F_{12}/F_{11})}{\partial n}$ and $\frac{\partial(F_{34}/F_{11})}{\partial n}$ have the largest uncertainties (5.42% and 6.32% MRE), indicating that they are the most difficult to train, due in part to a lot of close-to-zero values. All targets have similar accuracy to NN_0 predictions without Jacobians training. Note that the MRE is calculated for each Jacobian with an absolute value larger than a threshold, which is defined as 1e−6 for the c_{ext}, c_{abs}, F_{11}, and F_{22} Jacobians but 1e−4 for F_{12}/F_{11}, F_{34}/F_{11}, F_{33}/F_{11}, and F_{44}/F_{11} Jacobians given their larger values. In fact, if the Jacobians are small, we believe the single-scattering properties have little sensitivity

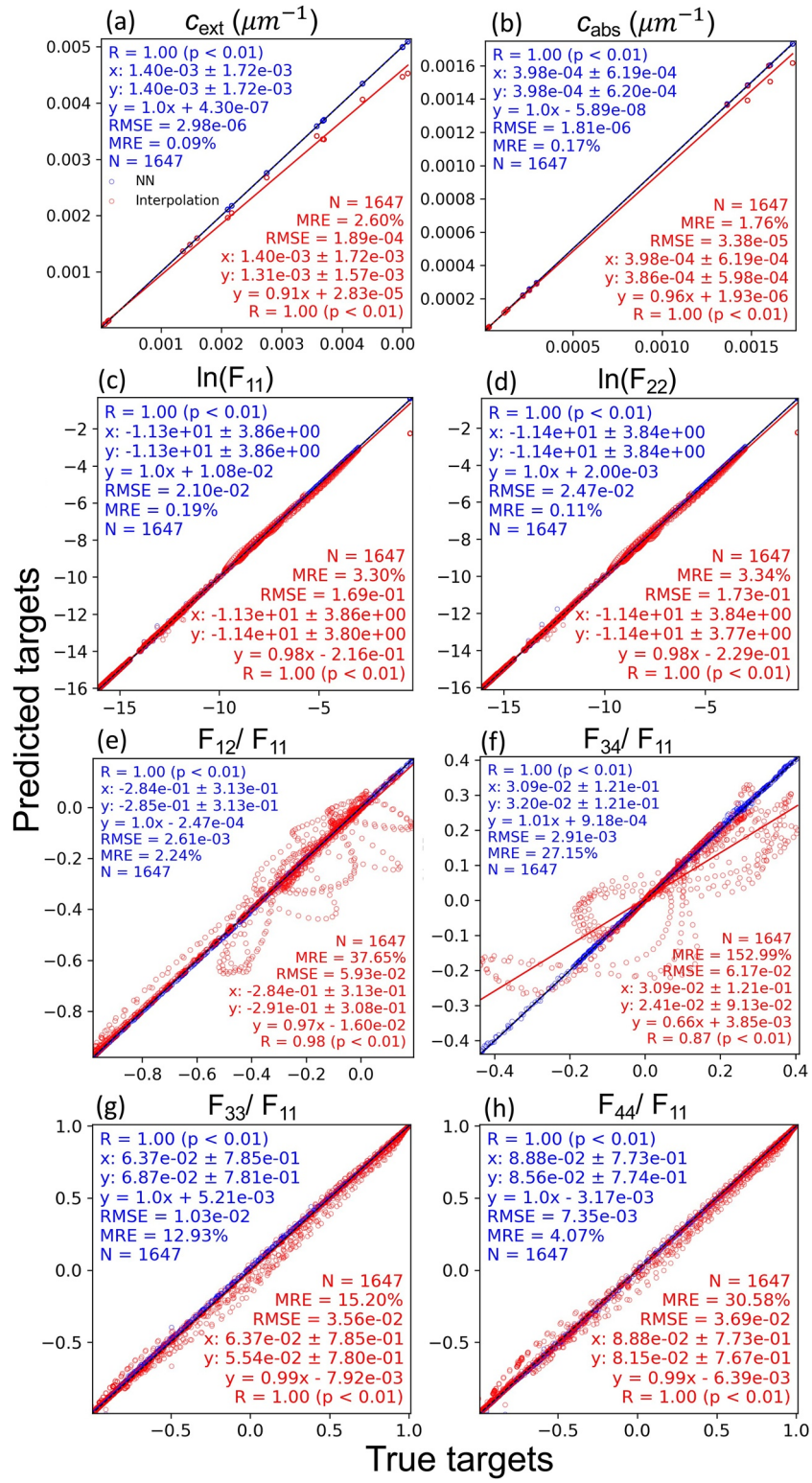


Figure 2. The scatter plots of the neural network (NN)-predicted targets (blue circle) and linear interpolated targets (red circle) with corresponding true values. In each subplot, the *x*-axis represents true targets, and the *y*-axis shows prediction from NN or interpolation. (a–h) are for $\boldsymbol{c}_{ext}$, $\boldsymbol{c}_{abs}$, ln(F_{11}), ln(F_{22}), F_{12}/F_{11}, F_{34}/F_{11}, F_{33}/F_{11}, and F_{44}/F_{11}. Similar to Figure 1, the fitting statistics are shown in each panel; texts in blue indicate results from NN prediction and red ones from interpolation method.

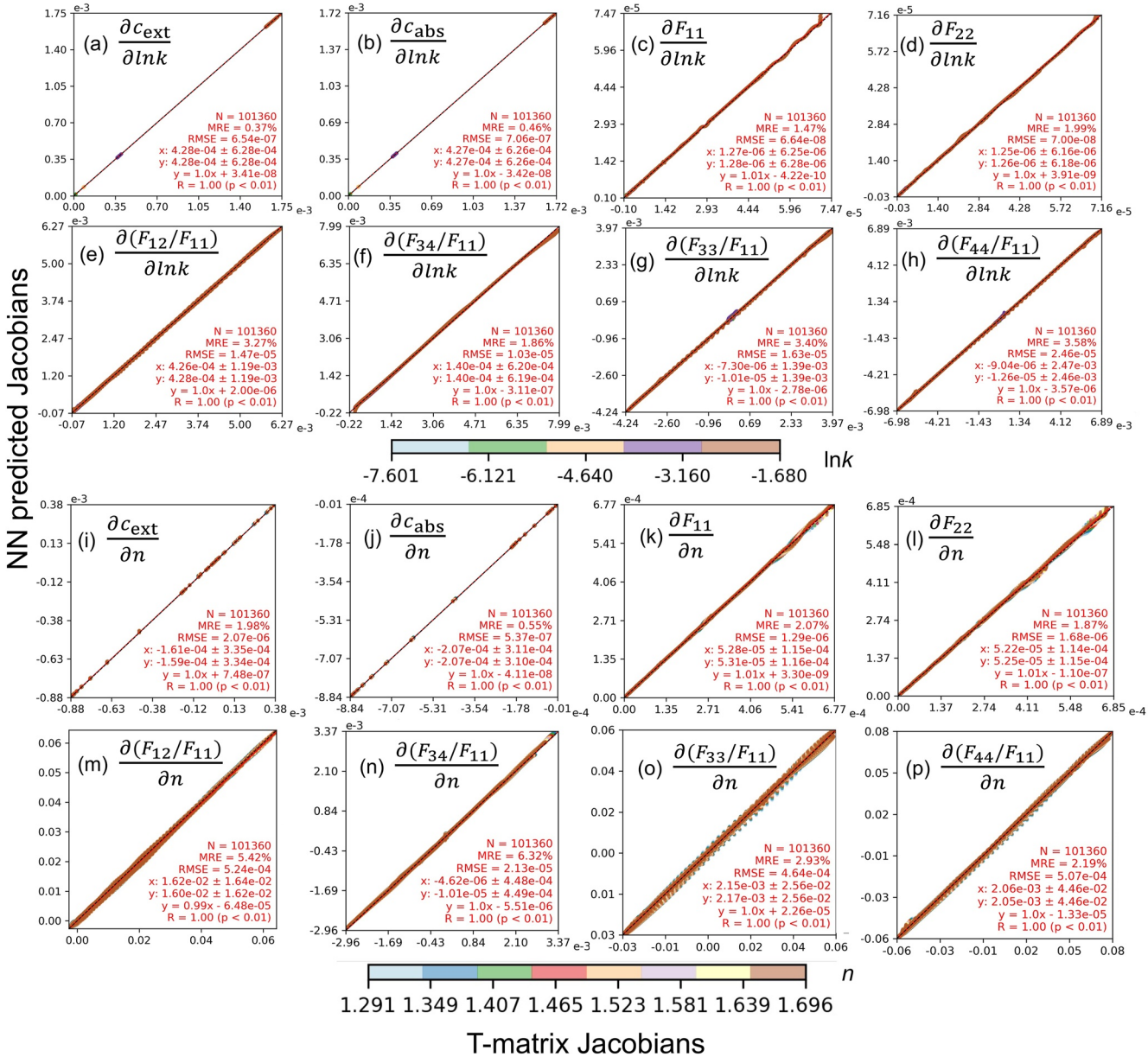


Figure 3. The scatter plots of the neural network-predicted Jacobians (*y*-axis) and true Jacobians from the linearized T-matrix model (*x*-axis). (a–h) are the Jacobians of eight targets (c_{ext}, c_{abs}, F_{11}, F_{22}, F_{12}/F_{11}, F_{34}/F_{11}, F_{33}/F_{11}, and F_{44}/F_{11}) with respect to ln*k*, while (i–p) are their Jacobians to *n*. Similar to Figures 1 and 2, the fitting results are shown in each panel. The different colors of points indicate the corresponding ln*k* or *n* values described by the color bars in the bottom.

to the microphysical parameters, which means these Jacobians affect the retrieval little. Therefore, even though the Jacobians predictions from the NN model only perform well with larger absolute values, they are still helpful in dust remote sensing by maintaining both high-speed calculation and good accuracy.

4. Conclusions and Discussions

We developed a feed-forward multi-layer fully connected NN model to replace the LUT approach to obtain the bulk single-scattering optical properties of an ensemble of dust particles assumed to be spheroids, as well as their Jacobians with respect to dust microphysical parameters, simultaneously. The NN was trained from the LUT database developed by Dubovik et al. (2002, 2006) in which the particle radius (sphere volume-equivalent), scattering angle, and real and imaginary parts of the refractive index are regarded as input features of NN, and

the corresponding particle extinction coefficient, absorption coefficient, and six elements in the scattering matrix are defined as targets to be learned. Based on the definition of neuron output in NN, the derivatives of targets to features are derived from NN model parameters and an activation function by the chain rule. To improve the Jacobians prediction by the NN model, we added the analytical Jacobians from a linearized EBCM T-matrix model as true values in the cost function to constrain the training process.

The validation illustrates that c_{ext}, c_{abs}, $\ln(F_{11})$, and $\ln(F_{22})$ can be predicted from NN with <0.5% MRE, while the other four phase matrix elements are more difficult to train, with 3%–10% MRE. Furthermore, the comparison of NN predictions and the results interpolated from LUT also underscores the advantage of replacing the LUT with the NN model due to its smaller MRE and higher correlation coefficients for all eight targets, especially for F_{12}/F_{11} and F_{34}/F_{11}, whose interpolated results have large errors. When validating NN-derived Jacobians, the comparison before and after adding T-matrix Jacobians in the training indicates the improvement in MRE of our method, even though the training of Jacobians used only limited datapoints. The accuracy of NN-derived Jacobians could be <7% after adding analytical Jacobians from linearized Tmatrix in the training.

Based on a NVIDIA GeForce RTX 2080 Ti GPU compute node and CUDA tool, the training of a NN model without true Jacobians in this study needs ~6.5 hr. When adding Jacobians in the training, 30–40 hr are necessary. Once the NN training is accomplished, the evaluation process can be achieved within 1 min even for 2 million data including the model loading and predictions of Jacobians. This efficiency becomes the biggest advantage of the NN model for applications in remote sensing where a huge amount of data needs to be processed.

This study provides a new application of NN in non-spherical aerosol optical properties simulation to replace traditional LUT, where an inaccurate interpolation method needs to be used. The expression of NN Jacobians and the improved method of Jacobians training in this study also provide a possibility to compute the derivatives of aerosol optical properties with respect to aerosol size and index of refraction without any approximation compared with the finite difference method. Since the true Jacobians used are only for small particles where a linearized EBCM T-matrix could be used, this NN model deserves more training in the future to include more particle sizes and shapes. This framework of this method, however, can be generalized beyond the remote sensing problem to many inversion problems that use gradient descent for optimization.

Data Availability Statement

The data used for the training can be obtained via the website of our co-author Dr. Dubovik's group by free registration (https://www.grasp-open.com/products/spheroid-package-release/). The Jacobians derived from the linearized T-matrix model and used in training, validation, and test process are available at zenodo via https://doi.org/10.5281/zenodo.5770687. Version 0.0.2 of the python codes used to make NN predictions and compare with test data from T-matrix is preserved at https://doi.org/10.5281/zenodo.5770738.

Acknowledgments
This work is supported in part by NASA's Remote Sensing Theory program (Grant No.: 80NSSC20K1747) and in part by the James E. Ashton Professorship, the Iowa Technology Institute's computational funds, and High-Performance Computing at the University of Iowa.

References

Bi, L., & Yang, P. (2014). Accurate simulation of the optical properties of atmospheric ice crystals with the invariant imbedding T-matrix method. *Journal of Quantitative Spectroscopy and Radiative Transfer*, *138*, 17–35. https://doi.org/10.1016/j.jqsrt.2014.01.013

Chen, N., Wang, W., Gatebe, C., Tanikawa, T., Hori, M., Shimada, R., et al. (2018). New neural network cloud mask algorithm based on radiative transfer simulations. *Remote Sensing of Environment*, *219*, 62–71. https://doi.org/10.1016/j.rse.2018.09.029

Draine, B. T., & Flatau, P. J. (1994). Discrete-dipole approximation for scattering calculations. *Journal of the Optical Society of America. A*, *11*, 1491–1499. https://doi.org/10.1364/josaa.11.001491

Dubovik, O., Sinyuk, A., Lapyonok, T., Holben, B. N., Mishchenko, M., Yang, P., et al. (2006). Application of spheroid models to account for aerosol particle nonsphericity in remote sensing of desert dust. *Journal of Geophysical Research*, *111*, D11208. https://doi.org/10.1029/2005jd006619

Dubovik, O., Holben, B. N., Lapyonok, T., Sinyuk, A., Mishchenko, M. I., Yang, P., & Slutsker, I. (2002). Non-spherical aerosol retrieval method employing light scattering by spheroids. *Geophysical Research Letters*, *29*, 541–545. https://doi.org/10.1029/2001gl014506

Huang, J., Minnis, P., Chen, B., & Huang, Z. (2008). Long-range transport and vertical structure of Asian dust from CALIPSO and surface measurements during PACDEX. *Journal of Geophysical Research*, *113*, D23212. https://doi.org/10.1029/2008jd010620

Hughes, M. J., & Hayes, D. J. (2014). Automated detection of cloud and cloud shadow in single-date Landsat imagery using neural networks and spatial post-processing. *Remote Sensing*, *6*, 4907–4926. https://doi.org/10.3390/rs6064907

Johnson, B. R. (1988). Invariant imbedding T matrix approach to electromagnetic scattering. *Applied Optics*, *27*, 4861–4873. https://doi.org/10.1364/ao.27.004861

Kane, Y. (1966). Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media. *IEEE Transactions on Antennas and Propagation*, *14*, 302–307. https://doi.org/10.1109/tap.1966.1138693

Meng, Z., Yang, P., Kattawar, G. W., Bi, L., Liou, K. N., & Laszlo, I. (2010). Single-scattering properties of tri-axial ellipsoidal mineral dust aerosols: A database for application to radiative transfer calculations. *Journal of Aerosol Science*, *41*, 501–512. https://doi.org/10.1016/j.jaerosci.2010.02.008

Mishchenko, M. I., & Travis, L. D. (1994). T-matrix computations of light scattering by large spheroidal particles. *Optics Communications*, *109*, 16–21. https://doi.org/10.1016/0030-4018(94)90731-5

Mishchenko, M. I., Travis, L. D., Kahn, R. A., & West, R. A. (1997). Modeling phase functions for dustlike tropospheric aerosols using a shape mixture of randomly oriented polydisperse spheroids. *Journal of Geophysical Research*, *102*, 16831–16847. https://doi.org/10.1029/96jd02110

Mishchenko, M. I., Travis, L. D., & Lacis, A. A. (2002). *Scattering, absorption, and emission of light by small particles*. Cambridge university press

Mohajerani, S., Krammer, T. A., & Saeedi, P. (2018). *Cloud detection algorithm for remote sensing images using fully convolutional neural networks*. *arXiv preprint arXiv:1810.05782*

Moulin, C., Coauthors, C. E., Dayan, U., Masson, V., Ramonet, M., Bousquet, P., et al. (1998). Satellite climatology of African dust transport in the Mediterranean atmosphere. *Journal of Geophysical Research*, *103*, 13137–13144. https://doi.org/10.1029/98jd00171

Nanda, S., De Graaf, M., Veefkind, J. P., Ter Linden, M., Sneep, M., De Haan, J., & Levelt, P. F. (2019). A neural network radiative transfer model approach applied to the Tropospheric Monitoring Instrument aerosol height algorithm. *Atmospheric Measurement Techniques*, *12*, 6619–6634. https://doi.org/10.5194/amt-12-6619-2019

Pachauri, R. K., Leo, M., & Core Writing Team (2014). *IPCC, 2014: Climate Change 2014: Synthesis Report* (p. 151).

Rodgers, C. D. (2000). *Inverse methods for atmospheric Sounding: Theory and Practice*. World Scientific Publishing Co. Pte. Ltd.

Saito, M., Yang, P., Ding, J., & Liu, X. (2021). A comprehensive database of the optical properties of irregular aerosol particles for radiative transfer simulations. *Journal of the Atmospheric Sciences*, *78*, 2089–2111. https://doi.org/10.1175/jas-d-20-0338.1

Saponaro, G., Kolmonen, P., Karhunen, J., Tamminen, J., & de Leeuw, G. (2013). A neural network algorithm for cloud fraction estimation using NASA-Aura OMI VIS radiance measurements. *Atmospheric Measurement Techniques*, *6*, 2301–2309. https://doi.org/10.5194/amt-6-2301-2013

Segal-Rozenhaimer, M., Miller, D. J., Knobelspiesse, K., Redemann, J., Cairns, B., & Alexandrov, M. D. (2018). Development of neural network retrievals of liquid cloud properties from multi-angle polarimetric observations. *Journal of Quantitative Spectroscopy and Radiative Transfer*, *220*, 39–51. https://doi.org/10.1016/j.jqsrt.2018.08.030

Spurr, R. J. D. (2006). Vlidort: A linearized pseudo-spherical vector discrete ordinate radiative transfer code for forward model and retrieval studies in multilayer multiple scattering media. *Journal of Quantitative Spectroscopy and Radiative Transfer*, *102*, 316–342. https://doi.org/10.1016/j.jqsrt.2006.05.005

Spurr, R., Wang, J., Zeng, J., & Mishchenko, M. (2012). Linearized T-matrix and Mie scattering computations. *Journal of Quantitative Spectroscopy and Radiative Transfer*, *113*, 425–439. https://doi.org/10.1016/j.jqsrt.2011.11.014

Sun, W., Fu, Q., & Chen, Z. (1999). Finite-difference time-domain solution of light scattering by dielectric particles with a perfectly matched layer absorbing boundary condition. *Applied Optics*, *38*, 3141–3151. https://doi.org/10.1364/ao.38.003141

Takenaka, H., Nakajima, T. Y., Higurashi, A., Higuchi, A., Takamura, T., Pinker, R. T., & Nakajima, T. (2011). Estimation of solar radiation using a neural network based on radiative transfer. *Journal of Geophysical Research*, *116*, D08215. https://doi.org/10.1029/2009jd013337

Wang, J., Liu, X., Christopher, S. A., Reid, J. S., Reid, E., & Maring, H. (2003). The effects of non-sphericity on geostationary satellite retrievals of dust aerosols. *Geophysical Research Letters*, *30*, 2293. https://doi.org/10.1029/2003gl018697

Wang, J., Nair, U. S., & Christopher, S. A. (2004). GOES 8 aerosol optical thickness assimilation in a mesoscale model: Online integration of aerosol radiative effects. *Journal of Geophysical Research*, *109*, D23203. https://doi.org/10.1029/2004jd004827

Wang, J., Xu, X., Ding, S., Zeng, J., Spurr, R., Liu, X., et al. (2014). A numerical testbed for remote sensing of aerosols, and its demonstration for evaluating retrieval synergy from a geostationary satellite constellation of GEO-CAPE and GOES-R. *Journal of Quantitative Spectroscopy Radiatiative Transfer 2014*, *146*, 510–28. Retrieved from https://www.sciencedirect.com/science/article/pii/S0022407314001368

Xu, X., Wang, J., Wang, Y., Henze, D. K., Zhang, L., Grell, G. A., et al. (2017). Sense size-dependent dust loading and emission from space using reflected solar and infrared spectral measurements: An observation system simulation experiment. *Journal of Geophysical Research: Atmospheres*, *122*, 8233–8254. https://doi.org/10.1002/2017jd026677

Yang, P., Ding, J., Panetta, R. L., Liou, K.-N., Kattawar, G. W., & Mishchenko, M. (2019). On the convergence of numerical computations for both exact and approximate solutions for electromagnetic scattering by nonspherical dielectric particles. *Electromagnetic Waves (Cambridge, Mass*, *164*, 27–61. https://doi.org/10.2528/pier18112810

Yang, P., & Liou, K. (1996a). Finite-difference time domain method for light scattering by small ice crystals in three-dimensional space. *JOSA A*, *13*, 2072–2085. https://doi.org/10.1364/josaa.13.002072

Yang, P., & Liou, K. N. (1996b). Geometric-optics–integral-equation method for light scattering by nonspherical ice crystals. *Applied Optics*, *35*, 6568–6584. https://doi.org/10.1364/ao.35.006568

Yang, P., Liou, K., Mishchenko, M. I., & Gao, B.-C. (2000). Efficient finite-difference time-domain scheme for light scattering by dielectric particles: Application to aerosols. *Applied Optics*, *39*, 3727–3737. https://doi.org/10.1364/ao.39.003727

Yu, J., Bi, L., Han, W., & Zhang, X. (2022). Application of a neural network to store and compute the optical properties of non-spherical particles. *Advances in Atmospheric Sciences*, *111*, D11208. https://doi.org/10.1007/s00376-021-1375-5

Yurkin, M. A., Maltsev, V. P., & Hoekstra, A. G. (2007). The discrete dipole approximation for simulation of light scattering by particles much larger than the wavelength. *Journal of Quantitative Spectroscopy and Radiative Transfer*, *106*, 546–557. https://doi.org/10.1016/j.jqsrt.2007.01.033