Aerosol Optical Depth Retrieval by Neural Networks Ensemble with Adaptive Cost Function

Vladan Radosavljevic, Slobodan Vucetic, Zoran Obradovic
Center for Information Science and Technology, Temple University, USA
vladan@ist.temple.edu, *vucetic@ist.temple.edu*, *zoran@ist.temple.edu*

Abstract

Aerosol Optical Depth (AOD) indicates the amount of depletion that a beam of radiation undergoes as it passes through the atmosphere. In this study a data-driven approach based on training neural networks for AOD prediction was considered. To train the predictor, we used more than a thousand collocated data points whose attributes were derived from MODIS instrument satellite observations and whose target AOD variable was obtained from the ground-based AERONET instruments. In order to minimize the relative error, which is performance measure preferred by domain scientists, we trained an ensemble of neural networks with adaptive cost functions. AOD prediction accuracy of neural networks was compared to the recently developed operational MODIS Collection 005 retrieval algorithm. Results obtained over the entire globe during the first six months of year 2005 showed that neural networks were more accurate than the operational retrieval algorithm.

1. Introduction

The aerosols, small suspended liquids or solids in the atmosphere emanating from natural and man-made sources have been recognized as one of the major factors influencing the climate changes [1]. By reflecting and absorbing solar radiation, aerosols have direct influence on both cooling the surface and warming the atmosphere. Therefore, the accurate prediction of aerosols composition and their concentration is one of the biggest challenges in current climate research.

Aerosol Optical Depth (AOD) is dimensionless quantity that represents the amount of light extinction by aerosols within atmosphere [1]. In conjunction with an atmospheric model, AOD can provide estimates of atmospheric effects on transmitted and reflected solar radiation which are further used in developing climate models. The process of predicting AOD using ground [2] or satellite [3] based observations is known as *AOD retrieval*.

AEROsol robotic NETwork (AERONET) is global remote sensing network of about 540 ground-based radiometers that retrieve AOD several times an hour under clear-sky conditions. AERONET AOD retrieval is very accurate and is often considered as ground truth when validating retrieval quality of various satellite-based AOD retrieval algorithms [4]. The MODerate resolution Imaging Spectrometer (MODIS), aboard NASA's Terra and Aqua satellites, is one of the major instruments for satellite-based AOD retrieval. MODIS observes reflected solar radiation over a large spectral range with a high spatial resolution and almost daily coverage of the entire Earth.

Operational algorithms used to retrieve AOD from MODIS observations are based on matching the atmospheric component of the observed spectral reflectance at the top of the atmosphere to the simulated values stored in lookup tables. The atmospheric component is obtained after the removal of the surface effect of the observed reflectance. The lookup tables are generated by the forward simulation model that estimates observed reflectance given the aerosol type and amount.

The operational retrieval algorithms are domain-driven because they are manually tuned by domain scientists. While this guarantees that the retrievals are based on sound physical principles, it also creates problems when there is an opportunity to use ground truth data to improve the algorithm. Currently, ground truth data are used by validation studies whose goal is to reveal major sources of retrieval errors. Then, the algorithm is occasionally manually modified to address these issues and improve retrieval accuracy.

For example, a major weakness of the previous version of the MODIS AOD retrieval algorithm (C004) was found to be overestimation at small AOD and underestimation at high AOD values. In order to overcome these problems, the new retrieval algorithm known as MODIS Collection 005 (C005) has been developed [5] by including more realistic aerosol models and dynamic surface assumptions. Clearly, the algorithm improvement is a time consuming and cumbersome procedure that does not guarantee that the ground truth data is used in the most efficient manner.

As an alternative to the domain-driven approach, statistical or data-driven retrieval approaches could be used. This approach is possible when a data set is available that consists of satellite observations and collocated ground-truth labels (e.g. AERONET AOD values). Given such data a regression model can be constructed that predicts the labels from the satellite observations.

Neural networks are often a model of choice in data-driven retrieval of atmospheric properties [6, 7, 8, 9, 10]. In our previous work neural networks have been trained to predict AERONET AOD over continental US using attributes derived from satellite data [11]. Comparing to the C004 AOD predictions, operational at that time, neural network AOD predictions were significantly more accurate. Because universal predictor could not completely explain the complex aerosol spatial-temporal variability, we proposed an integration of global and local data-driven aerosol predictors [12]. There, global neural network was trained to predict AERONET AOD over US in combination with region specific neural networks. The final AOD prediction was obtained as weighted average of global and local AOD predictors. The results showed that data-driven approach could be used as complement to the traditional domain-based retrieval algorithms.

In this work, we address an additional challenge that is related to retrieval accuracy estimation. Well known accuracy measures such as Mean Squared Error (MSE) are often not informative enough because (1) retrieval error increases with AOD, (2) distribution of AOD is skewed towards small values, and (3) there are many outliers. Instead, domain scientists use an array of accuracy measures to gain better insight into the retrieval accuracy [5]. For example, the relative error makes larger absolute errors more tolerable when predicting large AOD than when predicting small AOD. Ideally, one would like to have a retrieval algorithm that provides good accuracy with respect to most of the popular accuracy measures.

In this study we considered training of neural networks that minimize MSRE instead of MSE. In order to construct a predictor that is also accurate with respect to MSE and several other accuracy measures, we proposed an approach that builds an ensemble of neural networks, each trained with slightly different MSRE measure. The outputs of the ensemble are then used as inputs to a meta-level neural network that produces the actual AOD predictions.

AOD prediction accuracy of the proposed predictor was compared to individual neural networks trained to minimize MSE and MSRE, to an ensemble of neural networks trained to minimize MSE, as well as to the currently operational MODIS Collection 005 retrieval algorithm. Obtained results showed that our approach is more successful than the alternatives.

The rest of the paper is organized as follows. In the Section 2.1 spatial-temporal merging of MODIS and AERONET data is presented. Dataset used in our experiments is described in the Section 2.2. Different measures of AOD prediction accuracy are presented in the Section 3.1. Neural network based data-driven approach for predicting AOD is proposed in the Sections 3.2 and 3.3 and an ensemble of neural networks is introduced along with the adaptive cost functions. In the Section 4 comparative results are presented, C005 and neural network predictors are evaluated based on the defined accuracy measures. Finally, Section 5 contains our conclusions.

2. Data Sets

2.1. Aggregation of satellite and ground based AOD data

Although MODIS instrument has high spatial resolution (one pixel is as small as 250x250 m^2 at nadir), deterministic MODIS algorithms do not retrieve AOD for single pixels due to the high signal to noise ratio [5]. Instead, single pixels are aggregated to larger areas. Based on the assumption that AOD has small spatial variability, MODIS C005 algorithm retrieves AOD in 10 x 10 km^2 blocks. After discarding cloud, snow, ice, water and bright surface pixels along with 20% of the darkest and 50% of the brightest ones, the remaining pixels are aggregated as a representative for the corresponding 10 x 10 km^2 block.

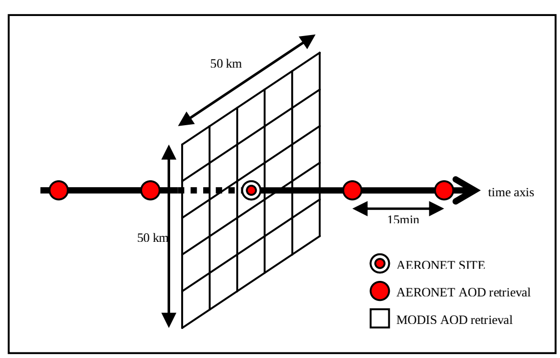


Figure 1. Spatial-temporal collocation of MODIS and AERONET AOD retrievals

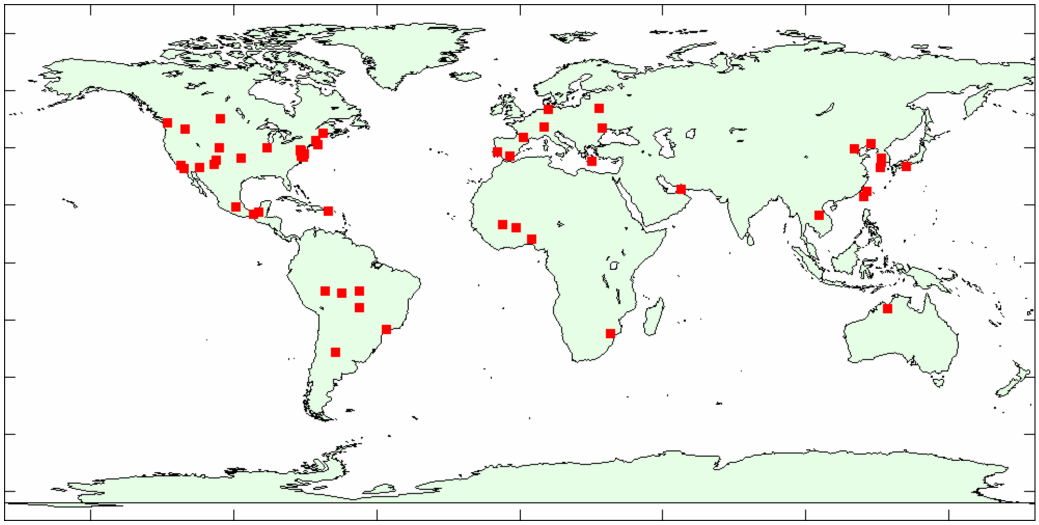


Figure 2. Location of 55 AERONET sites used in our experiments

Validation of MODIS AOD retrievals is usually performed using AERONET AOD retrievals as ground truth [4]. Whereas MODIS achieves an almost complete global coverage daily, AERONET retrievals are provided many times every day, but only over selected locations. Validation studies showed that it would be inappropriate to compare AOD from a single MODIS block directly to an AERONET point measurement [13]. Hence, the method named the "collocation" of the AERONET and the MODIS data has been proposed [13] (Figure 1). Essentially, this method involves aggregating initial MODIS blocks of 10 x 10 km size into blocks of size 50 x 50 km around each AERONET site, called spatial merging.

Due to the fact that MODIS and AERONET AOD retrievals may occur at different times, this gives rise to the need for temporal data merging. AERONET AOD data are acquired on average at intervals of 15 min. Assuming slow AOD variation over short time periods, the MODIS AOD retrievals are said to be temporally collocated with the corresponding AERONET AOD retrievals if there is a valid AERONET AOD retrieval within one hour time window centered at the satellite overpass time. The data collocated in this way can be obtained from the official MODIS website of NASA [5]. Each collocated data point is represented with time, date, average AERONET AOD, average MODIS reflectances and ancillary attributes. In our experiments we take into account a collocated data point if we have at least one valid out of possible 25 MODIS AOD retrievals in 50x50 km^2 spatial block and at least one valid AERONET AOD retrieval within the 30 minutes from the satellite overpass.

2.2. Data description

There are several levels of AERONET AOD measurements [2]. To avoid potential problem with outliers in ground truth data, AERONET Level 2.0 observations were considered since they were cloud screened and manually verified. For our study we collected 1,637 collocated MODIS Aqua and Terra observations with AERONET Level 2.0 points over entire globe from 55 AERONET sites (Figure 2) during the first six months of 2005. We could not consider a longer time period since AERONET Level 2.0 data were available for the years up to 2005 while C005 predictions were rarely available for the years before 2005. As shown in Figure 2, AERONET sites are not uniformly distributed over the globe. The highest density is within the US. On the other hand, continental Asia, Africa, and Australia are poorly covered. Hence, as a cautionary note, global applicability of data-driven approaches is somehow limited. During the time period, measurements from the 55 AERONET sites were almost uniformly distributed. In order to make fair comparison between the C005 algorithm and neural networks, we extracted only the satellite-based attributes that were used as

Table 1. List of attributes collected from the MODIS aggregated data

Attribute index	Description
1-7	Mean reflectance in 50 x 50 km blocks at seven wavelengths
8-14	Std. deviation of reflectance in 50 x 50 blocks at seven wavelengths
15-18	Solar Zenith, Solar Azimuth, Sensor Zenith, Sensor Azimuth angles
19	AERONET site elevation

inputs to C005 algorithm. Attributes used are listed in Table 1. The seven wavelengths were taken from the MODIS range between 440nm – 2100nm, as these are sufficient to describe aerosol properties [5].

By convention, AOD is reported at the 550nm wavelength. Since AERONET sites do not provide AOD value at that particular wavelength, based on the domain knowledge, we performed linear interpolation in the log scale of AERONET AOD measurements at 440nm and 670nm to estimate AOD at 550nm [5].

3. Methodology

3.1. Accuracy measures for AOD retrieval

Regardless of the approach used for AOD prediction, obtained predictor has to be evaluated and its accuracy adequately quantified. Considering AOD prediction as a regression problem, there are many possible measures that could be used to assess predictor performance. Given a target vector t of AERONET AOD measurements and vector y of corresponding AOD values derived from satellite observations, the appropriate measure of prediction accuracy could be *coefficient of determination* (R^2) defined as

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(y_i - t_i)^2}{\sum_{i=1}^{N}(t_i - \bar{t})^2} \tag{1}$$

where $\bar{t}$ represents the mean value of vector t, and summations are over all N collocated points. In the regression analysis, R^2 is preferred to simple quadratic distance measure *mean square error* (MSE),

$$MSE = \frac{\sum_{i=1}^{N}(y_i - t_i)^2}{N}, \tag{2}$$

since it takes into account variance in the target data. Portion of the variance that the predictor successfully models is described by R^2 value. The highest R^2 accuracy is 1, while R^2 accuracy of the predictor that simply predicts the mean of the population is 0. R^2 accuracy of some particularly poor predictors can even be negative.

Another measure that is often used is *correlation coefficient* (CORR)

$$CORR = \frac{\sum_{i=1}^{N}(y_i - \bar{y})(t_i - \bar{t})}{\sqrt{\sum_{i=1}^{N}(y_i - \bar{y})^2}\sqrt{\sum_{i=1}^{N}(t_i - \bar{t})^2}}, \tag{3}$$

where $\bar{y}$ represents the mean value of vector y, while other parameters were defined previously. CORR measure is insensitive to the prediction bias that is easily correctable [11].

In this paper we also considered several domain specific measures of AOD retrieval accuracy. Due to the inherent measurement errors of the MODIS instrument, expected boundaries for AOD retrieval error rate were proposed in [5]. Boundaries were defined as a linear function of ground truth (target) AOD value t_i

$$|y_i - t_i| \le 0.05 + 0.15t_i \tag{4}$$

Relation (4) directly implies that the errors in AOD prediction are much more tolerable for large AOD. Consequently, AOD predictor should be much more accurate in predicting small AOD. In this sense, a new accuracy measure can be defined as *mean squared relative error* (MSRE)

$$MSRE = \frac{1}{N}\sum_{i=1}^{N}\left[\frac{y_i - t_i}{0.05 + 0.15t_i}\right]^2 \tag{5}$$

where the sum is over all N collocated data points. If MSRE is closer to 0 the AOD predictor has better performance. However, having in mind the MODIS instrument uncertainty, it can be said that predictor performance is acceptable if MSRE is less than 1. In addition, we measure FRAC defined as the fraction (FRAC) of data points that are between domain expected boundaries,

$$FRAC = \frac{I}{N} \times 100\% \tag{6}$$

where I is the number of points that satisfy relation (4) and N is number of collocated points.

In order to demonstrate the need for using various kinds of measures for AOD predictor evaluation, let us analyze the accuracy of C005 AOD retrieval. Scatter plot of C005 AOD retrieval vs. AERONET AOD retrieval during the first six months of 2005 over whole globe is depicted in Figure 3 while the values of the proposed measures are shown in Table 2. In Figure 3 solid line represents an ideal, desirable AOD prediction, while dashed lines represent boundaries of an area within it data points are considered as acceptable for domain scientists. In Figure 3a the whole range of AOD values is plotted, while the zoomed-in portion of Figure 3a for small values of AOD (defined as AOD < 0.5 [5]) are presented in Figure 3b.

From Table 2 we can conclude that C005 AOD predictor has an excellent performance based on the CORR accuracy. However, R^2 accuracy tells us that there is a significant portion of variance which C005 predictor was unable to model. MSE accuracy is difficult to judge when accuracy of some simpler competing predictors is not available. Furthermore, domain specific MSRE accuracy is higher than 1

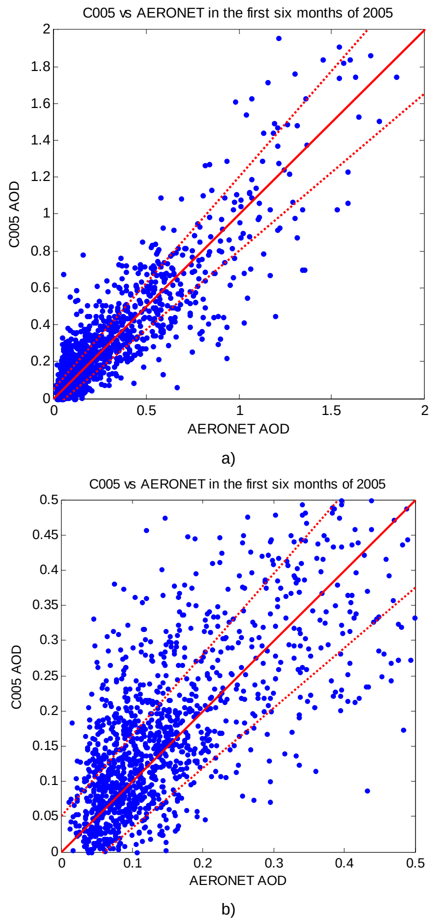


Figure 3. Scatter plot of C005 vs. AERONET AOD a) whole range of AOD b) small AOD

which indicates lower than expected accuracy. Finally, FRAC measure shows that almost 40% of predictions are of insufficient accuracy.

Table 2. C005 AOD vs. AERONET AOD accuracy

Time period	# points	C005 AOD retrieval				
		MSE	R^2	CORR	MSRE	FRAC
2005 Jan-Jun	1637	0.02	0.76	0.89	1.95	61%

3.2. Relative error as a cost function

Since AERONET AOD predictions are considered as highly accurate [2], they can be used as target values in data-driven approaches for AOD prediction. Construction of the neural network AERONET AOD predictor based on the MODIS observed parameters will be explored here.

Standard approach in building neural networks uses the MSE minimization as the optimization objective. This kind of cost function treats all errors equally regardless of the level of target value. As discussed in the previous section, domain scientists prefer small *squared relative errors* rather than small *squared errors*. Hence, in this application using an MSE function as the cost function in the neural network training process is not the most appropriate.

To address this issue, we introduced a new neural network cost function defined as *relative error* (REL) between predicted and ground truth AOD:

$$REL = \frac{1}{N}\sum_{i=1}^{N}\left[\frac{y_i - t_i}{b + at_i}\right]^2 \qquad (7)$$

where a and b are positive user defined parameters. Note that this measure is equivalent to the MSE if a=0 and b=1 and is equivalent with MSRE from equation (5) if a=0.15 and b=0.05. Influence of the various errors in the training process can be controlled by tuning parameters a and b. Different combination of a and b values put different importance to the errors with small and large AOD. For backpropagation algorithm, neural network weights have been updated in the training process proportionally to the derivative of the cost function

$$\frac{\partial REL}{\partial y_i} = \frac{2}{N}\frac{(y_i - t_i)}{(b + at_i)^2} \qquad (8)$$

By analyzing equation (8), we can conclude that if b approaches 0, the term at_i becomes more important and so does error when predicting small AOD. On the other hand, if b is large, term at_i for some small a can be neglected. Hence, errors for small and large AOD become of similar importance. Parameter a in conjunction with parameter b defines sensitivity to the ground truth value. As long as parameter a increases, neural network becomes more sensitive to the errors made in predicting small AOD. Otherwise, all errors become equivalent in the training process.

3.3. Ensembles with adaptive cost functions

REL cost function (equation (7)) with a=0.15 and b=0.05 directly leads to maximization of MSRE accuracy from equation (5). However, neural network trained in this way would have decreased MSE accuracy (equation (2)). We were interested in construction of an AOD predictor that is accurate with respect to both accuracy measures. To achieve this, we used the observation that MSRE-optimized neural networks will be more accurate when AOD is small, while MSE-optimized networks will work better when AOD is large.

However, the problem arises because it is not known in advance whether the AOD value is small or large. If the predictor which has the ability to decide whether AOD value is large or small has been used, the accurate prediction of medium level AOD values would still be the problem. More specifically, such a predictor would either overestimate or underestimate medium level AOD depending on whether it was "classified" as large or small, respectively. To solve this problem we proposed a two-stage approach:

1. construct an ensemble of neural networks among which some would be specialized in predicting small AOD while others would be specialized in predicting large AOD. This could be done by using different values for the parameters *a* and *b*.
2. apply the meta-learning method previously defined as "stacked generalization" approach [14], more specifically use outputs of the ensemble networks as inputs to a second-level neural network. This neural network is optimized to minimize MSRE and is used as AOD predictor (Figure 4).

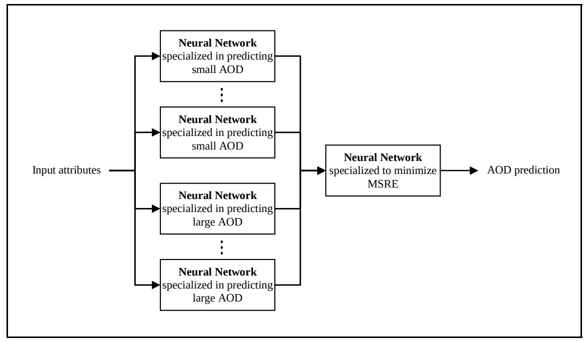


Figure 4. Architecture of the proposed two-stage system for AOD prediction

4. Experimental results

It is well known that most aerosols have temporal correlation below one week [1]. Therefore, we can neglect AOD time correlation if we divide our six months data set in weekly intervals. Thanks to this result, in all of the following experiments 4-cross-validation method was applied in the following manner. For each month, data from one week were left out and they were used as a test set while the remaining three weeks data were merged and used as a training set.

Neural networks with ten neurons in the hidden layer and one neuron in the output layer were used throughout all experiments. Neural networks had 19 inputs as there were 19 attributes in feature vector. Sigmoid activation function has been used for all hidden neurons while the linear activation function was used for the output neuron.

4.1. Experiments using single neural network

First, we examined single neural network predictors trained to predict AERONET AOD using aggregated MODIS attributes from Table 1. Two different neural network predictors were evaluated. One predictor used standard MSE measure as a cost function, while the other used our novel REL measure defined in equation (7). According to the analysis from the Section 3, parameters *a* and *b* were fixed as $a=0.15$ and $b=0.05$ in order to give an emphasis to the errors attained in predicting small AOD values. Results achieved by deterministic C005 algorithm are presented in the Table 3. Accuracies of neural networks are presented in Tables 4 and 5. We report results on 5 accuracy measures from Section 3.1 – MSE, R^2, CORR, MSRE, FRAC.

Based on the obtained results we can conclude that both neural networks were more accurate in predicting AOD than the operational C005 algorithm for all 5 accuracy measures. Neural network trained using an MSE cost function has better performance regarding MSE, R^2, and CORR (Table 4), while neural network trained using REL as a cost function achieves better performance considering MSRE and FRAC measures (Table 5).

4.2. Experiments using an ensemble of neural networks

To evaluate the proposed ensemble approach, we built two ensembles of neural networks which consisted of ten networks followed by a filter network trained to make the final decision. Each network in the ensemble was trained separately on the whole training data set. Second-level network was trained on the outputs of ensemble networks. Two approaches were evaluated. First, all networks in the ensemble along with the filter network were trained using MSE cost function. Obtained results are presented in the Table 6. Second, neural networks in the ensemble were trained in the following way. Half of them were specialized for the small AOD prediction by using REL as the cost function with parameter *a* changing from $a=0.05$ to $a=0.25$ in the steps of 0.05 and parameter *b* fixed to the value $b=0.05$. Another half of networks were specialized in prediction of large AOD values which was achieved by setting parameter *a* from $a=0.05$ to $a=0.25$ in the steps of 0.05 and by fixing parameter *b* to $b=1$. Having in mind domain requirements, a second-level neural network was trained to be more accurate in small AOD prediction. Hence, the parameters of this network are set to $a=0.15$ and $b=0.05$. Results are presented in the Table 7.

Comparing ensembles of neural networks to the single neural network predictors, we noticed that all measures were either similar or significantly improved

Table 3. C005 AOD vs. AERONET AOD

Year	Week out	# Points	C005 AOD				
			MSE	R2	CORR	MSRE	FRAC
2005 First six months	1	436	0.018	0.76	0.88	1.75	61%
	2	405	0.019	0.81	0.91	1.74	65%
	3	405	0.022	0.66	0.83	2.25	56%
	4	391	0.021	0.79	0.91	2.04	63%
Overall mean		**1637**	**0.020**	**0.76**	**0.89**	**1.95**	**61%**

Table 4. Neural network AOD (MSE) vs. AERONET AOD

Year	Week out	# Points	Neural network AOD, cost function MSE				
			MSE	R2	CORR	MSRE	FRAC
2005 First six months	1	436	0.014	0.81	0.90	1.54	66%
	2	405	0.014	0.86	0.93	1.37	68%
	3	405	0.016	0.75	0.87	1.74	60%
	4	391	0.015	0.85	0.92	1.67	63%
Overall mean		**1637**	**0.015**	**0.82**	**0.91**	**1.58**	**64%**

Table 5. Neural network AOD (REL with *a=0.15* and *b=0.05*) vs. AERONET AOD

Year	Week out	# Points	Neural network AOD, cost function REL				
			MSE	R2	CORR	MSRE	FRAC
2005 First six months	1	436	0.019	0.74	0.88	1.30	68%
	2	405	0.018	0.82	0.92	0.93	73%
	3	405	0.018	0.72	0.86	1.26	68%
	4	391	0.016	0.83	0.92	1.19	65%
Overall mean		**1637**	**0.018**	**0.78**	**0.89**	**1.17**	**68%**

Table 6. Ensemble of neural networks AOD (MSE) vs. AERONET AOD

Year	Week out	# Points	Neural network AOD, cost function MSE				
			MSE	R2	CORR	MSRE	FRAC
2005 First six months	1	436	0.013	0.83	0.91	1.23	68%
	2	405	0.013	0.87	0.93	1.03	74%
	3	405	0.014	0.79	0.89	1.39	69%
	4	391	0.012	0.88	0.94	1.21	75%
Overall mean		**1637**	**0.013**	**0.84**	**0.92**	**1.22**	**71%**

Table 7. Ensemble of neural networks AOD (adaptive REL) vs. AERONET AOD

Year	Week out	# Points	Neural network AOD, cost function REL				
			MSE	R2	CORR	MSRE	FRAC
2005 First six months	1	436	0.014	0.81	0.91	1.02	74%
	2	405	0.012	0.88	0.94	0.83	80%
	3	405	0.013	0.80	0.90	1.10	72%
	4	391	0.013	0.86	0.94	0.95	74%
Overall mean		**1637**	**0.013**	**0.84**	**0.92**	**0.97**	**75%**

by applying ensembles instead of single neural network as AOD predictor. Therefore, we conclude that ensembles of neural networks are much more accurate than single neural network AOD predictors.

By comparing results shown at Table 6 and Table 7, we conclude that an ensemble of neural networks trained with adaptive cost functions based on REL function preserves MSE, R^2 and CORR accuracies, while it significantly improves MSRE (from 1.22 to 0.97) and FRAC (from 71% to 75%).

Improvement in the AOD prediction can be seen in the Figure 5 where comparative scatter plots of C005 vs. AERONET AOD and ensemble of neural networks (with adaptive REL) AOD vs. AEROENT AOD are presented. By inspecting these plots, we can conclude that the ensemble is equally successful for both small and large AOD. High accuracy of the proposed method in predicting small AOD can be seen in the zoomed-in plots in the Figure 6. Comparing to the C005, bias in predicting small AOD values is significantly reduced using ensemble of neural networks in conjunction with the adaptive cost function.

5. Conclusion

An ensemble-based data-driven approach for AOD prediction was presented. Neural networks from the ensemble were trained using collocated data points whose attributes were derived from MODIS instrument satellite observations and whose target AOD variable was obtained from the ground-based AERONET instruments. Instead of relying on MSE minimization criterion for neural network training, we proposed use of the relative error REL, which can be considered as generalization of MSE.

We observed that REL criterion allowed us to achieve increased accuracy over certain ranges of AOD values. In an attempt to provide a predictor that is accurate over the whole range of AOD values for each of the 5 commonly used accuracy measures, we proposed an ensemble of neural networks with adaptive cost functions. Some networks in the ensemble were specialized in predicting small AOD while others were specialized in predicting large AOD.

The experiments showed that the proposed ensemble outperformed an ensemble that used standard MSE optimization; it managed to achieve as high MSE, R^2 and CORR accuracies while it significantly improved MSRE and FRAC accuracies. In addition, AOD prediction accuracy of the proposed ensemble was compared to the recently developed operational MODIS Collection 005 retrieval algorithm. Results obtained over the entire globe during the first six months of year 2005 showed that the proposed ensemble of neural networks was significantly more accurate for all the considered accuracy measures.

Acknowledgement

This work is funded in part by NSF grant IIS-0612149.

6. References

[1] Y. J. Kaufman, D. Tanre, and O. Boucher, “A satellite view of aerosols in the climate system”, Nature, 419: 215-223, 2002.

[2] B. N. Holben, T. F. Eck, I. Slutsker, T. Tanre, J. P. Buis, A. Setzer, E. Vermote, J. A. Reagan, Y. J. Kaufman, T. Nakajima, F. Lavenu, I. Jankowiak, and A. Smirnov, “AERONET: a federated instrument network and data archive for aerosol characterization”, Remote Sens. Environ., 37 : 2403 – 2412, 1998.

[3] M. D. King, Y. J. Kaufman, P. W. Menzel, and D. Tanreacute, “Remote sensing of cloud, aerosol, and water vapor properties from the Moderate Resolution Imaging Spectrometer (MODIS)”, IEEE Trans. on Geosci. and Remote Sens., 30: 2-27, 1992.

[4] R. C. Levy, L. A. Remer, et al., “Evaluation of the MODIS aerosol retrievals over ocean and land during CLAMS”, J. Atmos. Sci., 62(4): 974-992, 2005.

[5] L. A. Remer, D. Tanré and Y. J. Kaufman, “Algorithm for Remote Sensing of Tropospheric Aerosol from MODIS for Collection 005”,
http://modis-atmos.gsfc.nasa.gov/atbd02.pdf, 2006.

[6] F. Aires, C. Prigent, W. B. Rossow, and M. Rothstein, “A new neural net-work approach including first guess for retrieval of atmospheric watervapor, cloud liquid water path, surface temperature, and emissivities overland from satellite microwave observations”, J. Geophys. Res., 106,14,887–14,908, 2001.

[7] V. M. Krasnopolsky, “Neural networks for standard and variational satelliteretrievals”, OMB Contrib. 148, NOAA Environ. Model. Cent., Ocean Model. Branch, Washington, D.C., 1997.

[8] F. Chevallier, F. Chéruy, N. A. Scott, and A. Chédin, “A neural networkapproach for a fast and accurate computation of a longwave radiativebudget”, J. Appl. Meteorol., 37, 1385–1397, 1998.

[9] M. D. Muller, A. K. Kaifel, M. Weber, S. Tellmann, J. P. Burrows and D. Loyola, “Ozone profileretrieval from Global Ozone Monitoring Experiment (GOME) data using a neural network approach (neural network ozone retrieval system (NNORSY))”, J. Geophys. Res., 108(D16),4497, 2003.

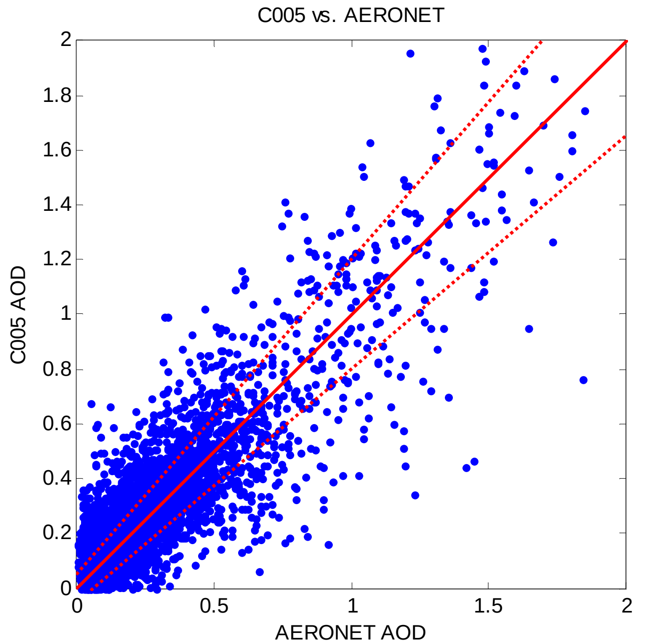
[10] C. Jamet, C. Moulin, and S. Thiria, “Monitoring aerosol opticalproperties over the Mediterranean from SeaWiFS images usinga neural network inversion”, Geophys. Res. Lett., 31, 2004.

[11] B. Han, Z. Obradovic, Z. Li and S. Vucetic, “Data Mining Support for Improvement of MODIS Aerosol Retrievals”, Proc. IEEE Int’l Geoscience and Remote Sensing Symp., Denver, CO, Aug. 2006.

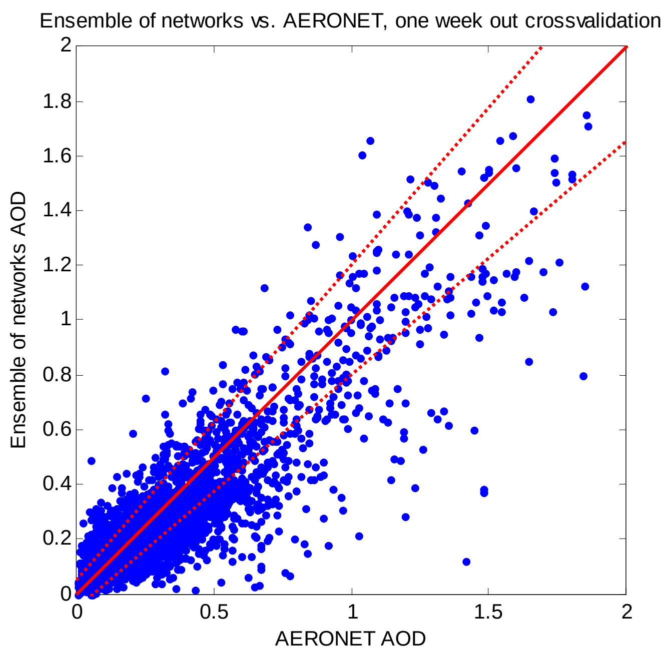
[12] B. Han, S. Vucetic, A. Braverman and Z. Obradovic, “A Statistical Complement to Deterministic Algorithms for Retrieving Aerosol Optical Thickness from Radiance Data”, Engineering Applications of Artificial Intelligence, vol. 19, no. 7, pp. 787-795, 2006.

[13] C. Ichoku, D. A. Chu, et al., “A spatio-temporal approach for global validation and analysis of MODIS aerosol products”, Geophys. Res. Lett. 29(12): no.616, 2002.

[14] D. Wolpert, “Stacked generalization”, Neural Networks 5: 241-259, 1992.

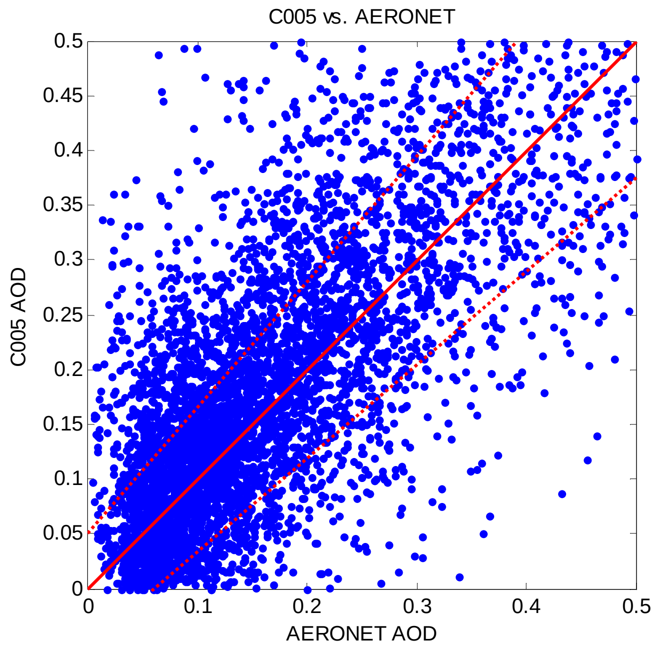


a)

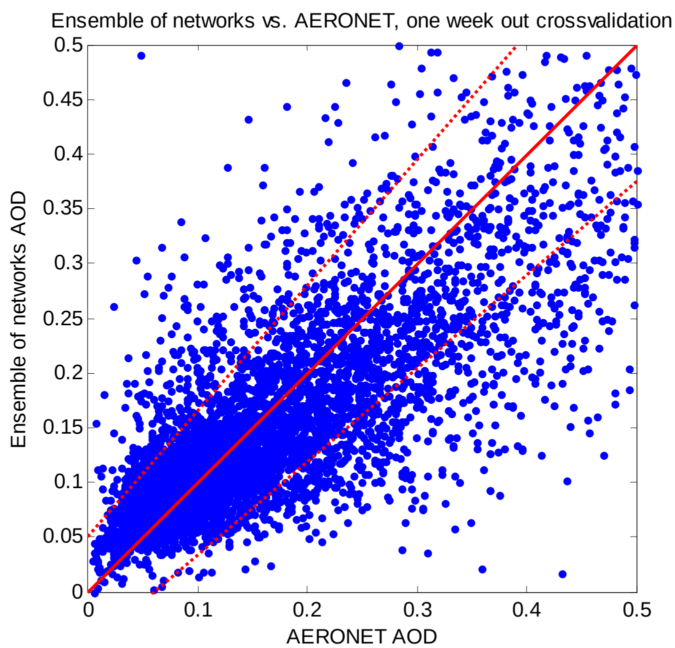


b)

Figure 5. Scatter plots of: a) C005 vs. AERONET and b) Ensemble of neural networks (adaptive REL) AOD predictions vs. AERONET; solid line – ideal prediction, dashed lines – boundaries of acceptable error



a)



b)

Figure 6. Zoomed scatter plots of: a) C005 vs. AERONET and b) Ensemble of neural networks (adaptive REL) AOD predictions vs. AERONET; solid line – ideal prediction, dashed lines – boundaries of acceptable error