A Data-Mining Technique for Aerosol Retrieval Across Multiple Accuracy Measures

Vladan Radosavljevic, Slobodan Vucetic, and Zoran Obradovic

***Abstract*—A typical approach in supervised learning is to select an accuracy measure and train a predictor that maximizes it. This can be insufficient in remote-sensing applications where predictor performance is often evaluated over multiple domain-specific accuracy measures. Here, we test the hypothesis that predictors can be trained to maximize performance over multiple accuracy measures. To do this, we evaluate several metalearning algorithms on the problem of aerosol optical depth (AOD) retrieval. The multiple accuracy measures included mean squared error, correlation, relative squared error, and fraction of satisfactory predictions. The proposed metalearning algorithms have a two-layer architecture, where the first layer consists of multiple neural networks, each trained using a different accuracy measure, and the second layer aggregates decisions of the first layer predictors. To evaluate AOD predictors, we used nearly 70 000 collocated data points whose attributes were radiances, solar and view angles, and terrain elevation from MODerate resolution Imaging Spectrometer (MODIS) instrument satellite observations and whose target AOD variable was obtained from the ground-based AEROsol robotic NETwork (AERONET) instruments. The data were collected at 221 AERONET locations over the globe in the period between 2005 and 2007. AOD prediction accuracies of neural networks were compared to the recently developed operational MODIS *C005* retrieval algorithm and to several other data-mining methods. Results showed that neural networks are better at reproducing the test data than the operational retrieval algorithm and that predictors obtained by metalearning are robust over multiple accuracy measures.**

***Index Terms*—Aerosol retrieval, metalearning, neural networks.**

I. INTRODUCTION

A DATA-MINING approach for classification and regression in remote sensing is based on learning a relationship between remotely sensed observations and the ground truth. The success of the resulting predictor is measured by its accuracy. Standard accuracy measures such as mean squared error (MSE) in regression are often selected due to their wide appeal and convenience. However, in many remote-sensing applications, standard accuracy measures can be hard to interpret or even misleading. In addition, domain scientists are often interested in multiple aspects of the predictor performance that are evaluated in various ways.

Ideally, one would like to have a predictor that provides good performance with respect to multiple accuracy measures. The complication is that predictors which perform well on one measure may not perform well on other measures. An important challenge is to train a predictor where the objective is not optimal performance on a single measure but robust performance across several measures.

An illustrative example of such a problem in remote sensing is aerosol optical depth (AOD) retrieval. Aerosols are minute particles suspended in the atmosphere originating from natural and man-made sources. AOD, as reported by surface and space-based passive remote sensing, is a measure of aerosol light extinction integrated vertically through the entire atmosphere. A regression model that retrieves AOD can be trained on a data set that consists of the satellite observations as inputs and ground-based AOD measurements as outputs. To demonstrate the need for multiple evaluation measures, let us analyze the accuracy of NASA's currently operational MODerate resolution Imaging Spectrometer (MODIS) retrieval algorithm (*C005*) [1]. A scatter plot of *C005* AOD retrieval versus ground-based AOD retrieval in a period of three years from 2005 to 2007 over the whole globe is shown in Fig. 1. The solid line represents the perfect agreement with AEROsol robotic NETwork (AERONET), while the dashed lines represent boundaries of an area within which retrievals are acceptable to domain scientists. Large absolute errors are more tolerable when retrieving large AOD than when retrieving small AOD. Therefore, a fraction of data points inside the bounded area (FRAC) is a suitable accuracy measure. MSE measure is also used for AOD retrieval assessment, but it is not as informative because of the following: 1) The retrieval error increases with AOD; 2) the distribution of AOD is skewed toward small values; and 3) there are many outliers. In addition to FRAC and MSE, domain scientists are also interested in the relative squared error (RSE) that considers larger absolute errors more tolerable when retrieving large AOD than when retrieving small AOD.

To construct a model that is accurate with respect to FRAC, MSE, and RSE, we propose to train an ensemble of neural networks, each with a different relative error measure, and combine their predictions. We explored three methods of combining ensemble predictions and compared them to neural network models optimized for a single accuracy measure as well as to the operational MODIS AOD retrieval algorithm *C005*.

II. ACCURACY MEASURES FOR AOD RETRIEVAL

There are many possible measures that could be used to assess AOD retrieval accuracy. Given vector $\boldsymbol{t} = [t_1, t_2, \ldots, t_N]$ of N target values (i.e., true AOD values) and vector $\boldsymbol{y} = [y_1, y_2, \ldots, y_N]$ of the corresponding predictions, the standard MSE is defined as

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

Manuscript received September 17, 2009; revised November 9, 2009. Date of publication February 2, 2010; date of current version April 14, 2010. This work was supported in part by the National Science Foundation under Grant IIS-0612149.

The authors are with the Information Science and Technology Center, Temple University, Philadelphia, PA 19122 USA (e-mail: vladan@ist.temple.edu; vucetic@ist.temple.edu; zoran@ist.temple.edu).

Digital Object Identifier 10.1109/LGRS.2009.2037720

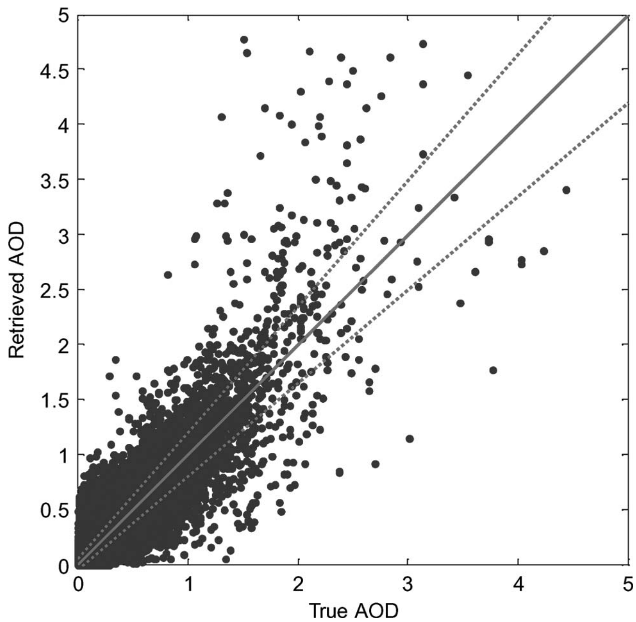

Fig. 1. Scatter plot of retrieved versus true AOD. Ideal retrievals are represented by a solid line, while dashed lines correspond to boundaries of a region of acceptable retrievals.

A related measure to MSE is the *coefficient of determination* (R^2), which is defined as

$$R^2 = 1 - \left(\sum_{i=1}^{N}(y_i - t_i)^2\right) \Bigg/ \left(\sum_{i=1}^{N}(\bar{t} - t_i)^2\right) \qquad (2)$$

where $\bar{t}$ represents the mean value of vector $\boldsymbol{t}$. The R^2 value describes a fraction of the variance that the predictor successfully explains. The highest R^2 is one, while the R^2 of the model that simply predicts the target variable mean is zero. The R^2 of some poor predictors can even be negative.

Another related measure, which is insensitive to the correctable bias, is the *correlation coefficient* (CORR)

$$CORR = \left(\sum_{i=1}^{N}(y_i - \bar{y})(t_i - \bar{t})\right) \Bigg/ \left(\sqrt{\sum_{i=1}^{N}(y_i - \bar{y})^2}\sqrt{\sum_{i=1}^{N}(t_i - \bar{t})^2}\right) \qquad (3)$$

where $\bar{y}$ represents the mean of $\boldsymbol{y}$.

We also consider several domain-specific measures. Geoscientists showed both theoretically and empirically that, taking into consideration the physical limitations of current satellite aerosol remote sensing, the desired absolute AOD retrieval error should be between 0.05 and 0.1 for small AOD and that it could increase to 15%–20% × AOD for large AOD [1] or better. Thus, the AOD retrieval can be considered successful if the absolute error is

$$|y_i - t_i| \leq 0.05 + 0.15t_i. \qquad (4)$$

We define the *fraction of successful predictions* (FRAC) as

$$FRAC = \frac{I}{N} \times 100\% \qquad (5)$$

where I is the number of predictions that satisfy relation (4).

TABLE I
C005 VERSUS AERONET GLOBAL AOD RETRIEVAL ACCURACY IN 2005–2007

Model	R^2	CORR	R_r^2	FRAC
C005	0.70	0.87	0.28	64.8%

Domain-specific RSE is defined as

$$RSE = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{y_i - t_i}{0.05 + 0.15t_i}\right)^2. \qquad (6)$$

RSE values less than one indicate that AOD retrievals are satisfactory. The closer the RSE is to zero, the better is the performance of a predictor. A related measure of accuracy is the *relative coefficient of determination*(R_r^2), which is defined as

$$R_r^2 = 1 - \left(\sum_{i=1}^{N}\left(\frac{y_i - t_i}{0.05 + 0.15t_i}\right)^2\right) \Bigg/ \left(\sum_{i=1}^{N}\left(\frac{\bar{t}_r - t_i}{0.05 + 0.15t_i}\right)^2\right) \qquad (7)$$

where $\bar{t}_r = \sum w_i t_i / \sum w_i$, $w_i = (0.05 + 0.15t_i)^{-2}$, represents the weighted mean of vector $\boldsymbol{t}$. R_r^2 is derived according to the general definition of a coefficient of determination [2]. The highest R_r^2 is one, while R_r^2 of the model that predicts the target weighted mean is zero.

The values of four accuracy metrics for the operational AOD retrieval algorithm called *C005* whose scatter plot is shown in Fig. 1 are shown in Table I. *C005* has an excellent performance based on CORR. However, R^2 tells us that there is a significant portion of variance which *C005* was unable to explain. Furthermore, domain-specific R_r^2 accuracy is small, which indicates lower than desired performance. Finally, the FRAC measure shows that more than 35% of retrievals fall outside the target agreement envelope.

III. ADAPTIVE COST FUNCTION

Neural networks are typically trained by minimizing MSE. This cost function treats all errors equally regardless of the target value. As discussed in the introduction, earth scientists prefer small *relative errors* rather than small *absolute errors* in situations where the uncertainty scales as the magnitude of the measured quantity. Hence, the MSE function is not the most appropriate cost function for this application. As a more general choice, we introduce a function defined as the *relative error* ($REL_{a,b}$) between retrieved and ground truth AODs

$$REL_{a,b} = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{y_i - t_i}{a + bt_i}\right)^2 \qquad (8)$$

where a and b are positive user-defined parameters. Here, the level of penalization of retrieval errors can be controlled by tuning parameters a and b. Note that $REL_{1,0}$ is equivalent to MSE, while $REL_{0.05,0.15}$ is equal to RSE.

We employ $REL_{a,b}$ as a cost function for training neural networks. When a is small, bt_i is dominant, and so, the emphasis is on reducing the error of retrieving small AOD. When a is large, errors for small and large AOD have similar importance. Sensitivity of a neural network optimization to t_i also depends on parameter b—for a large b, the network becomes more sensitive to the errors made when retrieving small AOD.

IV. Ensembles With Adaptive Cost Functions

Minimization of the $REL_{a,b}$ cost function, with $a = 0.05$ and $b = 0.15$, directly leads to the optimization of domain-specific measures mentioned in the introduction—maximization of FRAC and minimization of RSE. However, a neural network trained in this way would have decreased MSE accuracy. We are interested in the construction of a model that is accurate with respect to all accuracy measures.

$REL_{0.05,0.15}$-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 we used the model which has the ability to decide whether the AOD value is large or small, the accurate retrieval of medium-level AOD values would still be a problem. More specifically, such a model would either overestimate or underestimate AOD depending on whether it was "classified" as large or small, respectively. To solve this problem, we propose the following two-stage approach.

1) Constructing an ensemble of $2K$ neural networks among which K networks are specialized in retrieving small AOD, while the remaining K are specialized in retrieving large AOD. This is achieved by using different values of parameters a and b. Since the distribution of AOD is skewed to the small AOD, by design, all component networks are trained to penalize errors at small AOD. However, the intensity of this penalization varies per component network.
2) Combining the outputs of the component networks to obtain an integrated AOD retrieval.

The architecture of the proposed system is shown in Fig. 2. All first-stage component networks are trained using the same data set. $O_{S1}, O_{S2}, \ldots, O_{SK}$ corresponds to networks specialized for smaller AOD, while $O_{L1}, O_{L2}, \ldots, O_{LK}$ corresponds to networks specialized for larger AOD. Those outputs are integrated at the second stage using one of the following methods.

A. Integration by Averaging

Here, the final AOD prediction is obtained as a simple average of O_S and O_L neural networks. We will refer to this approach as *AVERAGE*.

B. Integration by a Meta Neural Network

Here, predictions of O_S and O_L neural networks are used to train a second-stage meta neural network. The meta neural network is optimized to minimize $REL_{0.05,0.15}$. We will refer to this two-stage structure as *META*.

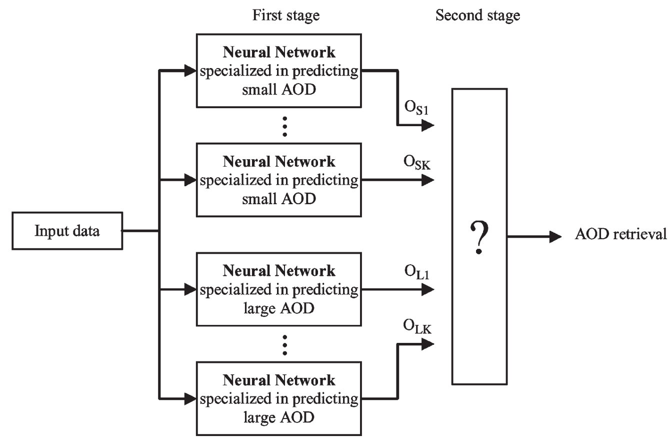

Fig. 2. Architecture of the proposed two-stage ensemble for AOD retrieval.

C. Integration by a Gating Neural Network

In the *GATING* approach, the first-stage networks are linearly combined according to the weights assigned by a gating network. A gating neural network is built as a binary classifier that predicts whether AOD is small or large. If the gating network O_G predicts a large AOD (i.e., O_G is close to one), larger weights are assigned to the O_L neural networks specialized for retrieving large AOD. On the other hand, O_G close to zero gives larger weights to O_S networks. Finally, O_G near 0.5 means that large and small AODs are likely equal, and weights of O_S and O_L are equal. To avoid bias, the sum of all the weights is set to one. The final AOD retrieval is computed as

$$y = \sum_{i=1}^{K} \left(\frac{O_G}{K} O_{Li} + \frac{1 - O_G}{K} O_{Si} \right) \tag{9}$$

where O_{Li} and O_{Si} are outputs of the first-stage networks.

To train the gating network, we assign large and small labels to AOD values. Domain knowledge suggests that AOD values that are less than 0.15 should be considered small [1]. To have balanced training data, instead of using 0.15, we use the median AOD value of 0.13 as the *threshold*. AOD values larger than the threshold are considered as large, while the remaining ones are considered as small.

V. Experiments

A. Data Set

MODIS, aboard NASA's Terra and Aqua satellites, is one of the major instruments for satellite-based AOD retrieval [3]. The MODIS instrument provides global coverage with a high spatial resolution and moderately accurate AOD retrieval. AERONET is a global network of about 250 ground-based instruments that observe aerosols [4]. AERONET instruments are densely situated in industrialized areas and are sparsely located elsewhere. AERONET AOD retrieval is often very accurate and is considered as ground truth when validating MODIS AOD retrieval quality.

The modalities of MODIS- and AERONET-based AOD retrieval are very different. MODIS achieves an almost complete global coverage daily, while AERONET retrievals are provided many times every day but only over selected locations. The collocation of the AERONET and the MODIS data involves aggregating MODIS observations into blocks of 50 km × 50 km around each AERONET site. 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 a one-hour window centered at the satellite overpass time. The data collocated in this way are obtained from the official MODIS Web site of NASA [1].

To avoid potential problems with outliers in ground truth data, we considered AERONET Level 2.0 observations since they were cloud screened and manually verified. For our study, we collected 68 935 collocated observations distributed globally over 221 AERONET sites from 2005 to 2007.

We extracted satellite-based attributes by consulting inputs to the MODIS operational retrieval algorithm. The radiances at four wavelengths were taken from the MODIS range 440–2100 nm, as these are sufficient to describe aerosol properties. We used average and standard deviation of radiances within 50 km × 50 km as attributes. We also used solar and sensor angles and surface elevation. *C005* retrieves AOD at 550 nm. Since AERONET instruments do not provide AOD values at that wavelength, we performed linear interpolation in the log scale of AERONET AOD at 440 and 670 nm [1].

B. Evaluation Methods

Spatial–temporal cross-validation was applied in all experiments. First, we split AERONET locations into five subsets A_i, $i = 1, \ldots, 5$, and created data sets D_i, $i = 1, \ldots, 5$, each with data points from one of the AERONET subsets. Then, we split each D_i into D_i^{56}, containing data from 2005 and 2006, and D_i^7, containing data from 2007. We reserved one of the D_i^{56} data sets for testing and merged data from the remaining four data sets D_j^{56}, $j \neq i$, for training. The trained predictor was tested on three data sets:

(TEST1) D_i^{56}—data collected in 2005 and 2006 at the locations unobserved during training;

(TEST2) $\{D_j^7, j \neq i\}$—data collected in 2007 at the locations observed in 2005 and 2006;

(TEST3) D_i^7—data collected in 2007 at the locations unobserved during training.

The procedure was repeated five times, for values $j = 1, \ldots, 5$, and the average accuracy over the five runs was reported. It is expected that TEST3 is the most challenging for prediction.

C. Benchmark Methods

1) Operational Retrieval Algorithm C005: The primary benchmark for comparison with our predictors was the most recent version of the MODIS operational algorithm called *C005*. The operational algorithms that retrieve AOD from MODIS observations rely on the domain knowledge of aerosol properties and are based on lookup tables representing the most common atmospheric conditions.

2) Single Neural Networks: As a baseline data-mining algorithm, we used single neural networks trained to predict AERONET AOD from MODIS attributes. Two different single neural network models were evaluated. The first network is trained by minimizing a standard MSE cost function (*SingleMSE*), while the second network minimized our novel $REL_{a,b}$ measure (*SingleREL*). Parameters a and b were fixed to $a = 0.05$ and $b = 0.15$.

TABLE II
SATELLITE-BASED VERSUS AERONET AOD RETRIEVAL ACCURACY ON TEST3 (UNSEEN LOCATIONS AND UNSEEN TIME)

Model	R^2	CORR	R_r^2	FRAC
C005	0.65	0.86	0.13	63.9%
SingleMSE	0.74	0.87	0.40	66.2%
SingleREL	0.68	0.85	0.55	69.3%
EnsembleMSE	**0.76**	**0.88**	0.45	68.9%
EnsembleREL	0.67	0.86	**0.56**	70.6%
DIFFREG	0.65	0.84	0.07	66.8%
AVERAGE	0.75	**0.88**	0.54	70.5%
META	0.75	0.87	0.50	69.5%
GATING	**0.76**	**0.88**	0.53	**70.9%**

3) Simple Ensembles of Neural Networks: We also compared the proposed methods to two ensemble algorithms. Each ensemble consisted of ten neural networks. The outputs of these ten networks were used as inputs to the second-level neural network. In the *EnsembleMSE* approach, all networks were trained using MSE cost function. In *EnsembleREL*, the cost function for all networks was $REL_{0.05,0.15}$.

4) Ensemble of Networks Specialized for Low and High AOD: In the *DIFFREG* approach, $K = 5$ neural networks were trained using a portion of the training data with small AOD, while another K networks were trained using data with large AOD. To permit smooth transition in attribute space, overlapping between two training data sets was allowed. Small AOD was defined as AOD $< threshold + \varepsilon$, while large AOD was defined as AOD $> threshold - \varepsilon$, where ε was 0.05. All networks were trained to minimize MSE, and the two sets of networks were integrated using the gating neural network described in Section IV-C.

D. Results on TEST3

Ensemble neural networks having 13 inputs and 10 neurons in a single hidden layer and 1 in the output layer were used in all experiments. The sigmoid activation function was used in hidden neurons, while the linear activation function was used for the output neuron.

The average accuracies of the proposed *AVERAGE*, *META*, and *GATING* predictors and of six benchmark algorithms using R^2, CORR, R_r^2, and FRAC measures are shown in Table II. These results were obtained on the most challenging TEST3 data. We note that the averaging of the coefficient of determination measure over five different cross-validation experiments might be misleading since those measures depend on standard deviation of a particular test set. However, the variation of R^2 in five sets used in these experiments was negligible, and so, we decided to also report the average R^2.

1) Operational Retrieval Algorithm C005: C005 accuracies are shown in the first row of Table II. As discussed in Section II, *C005* has an excellent performance based on CORR, but R^2 accuracy reveals that it was not able to explain a large portion of variance. Also, domain-specific R_r^2 and FRAC measures indicate that *C005*-based retrievals are of insufficient accuracy.

2) Single Neural Networks: SingleMSE and *SingleREL* accuracies are shown in rows 2 and 3 in Table II. Both single neural networks achieve closer agreement with AERONET than the MODIS *C005* values, based on all four metrics shown in Table II. However, their performance was quite different over individual accuracy measures: *SingleMSE* was more accurate with respect to R^2 and CORR, while *SingleREL* was a better choice with respect to R_r^2 and FRAC measures.

3) Simple Ensembles of Neural Networks: EnsembleMSE and *EnsembleREL* accuracies are listed in rows 4 and 5 of Table II. Both predictors outperformed *C005* in all accuracy measures. Also, they were more accurate than single neural networks. However, neither ensemble achieved consistently high performance on all four measures: *EnsembleMSE* achieved better accuracy than *EnsembleREL* with respect to R^2 and CORR, while *EnsembleREL* was better according to R_r^2 and FRAC measures.

4) Ensemble of Specialized Neural Networks: DIFFREG accuracies are listed in row 6 of Table II. This benchmark method was quite unsuccessful, with accuracies below *SimpleMSE* and just slightly better than *C005*.

5) Ensembles With Adaptive Cost Neural Networks: In *AVERAGE*, *META*, and *GATING* predictors, five neural networks of the ensemble were specialized for the retrieval of small AOD. This was achieved by using the $REL_{a,b}$ cost function, with $a = 0.05$ and b changing from $b = 0.03$ to $b = 0.15$ in steps of 0.03. Another five networks in the ensemble were specialized for the retrieval of large AOD by using $a = 1$ and b changing from $b = 0.03$ to $b = 0.15$ in steps of 0.03.

Results for *AVERAGE*, *META*, and *GATING* adaptive cost ensembles are presented in the last three rows of Table II. All three predictors were robust across all accuracy measures. *GATING* ensemble with a second-level gating neural network was slightly more accurate than the alternatives. On standard measures (R^2 and CORR), *GATING* was as good as the most successful benchmark method on these measures (*EnsembleMSE*), and it had similar accuracy with the best benchmark method (*EnsembleREL*) on domain-specific measures (R_r^2 and FRAC). This result shows that it is possible to simultaneously achieve high accuracy on disparate measures using a two-level ensemble neural network architecture.

E. Results on TEST1 and TEST2

Accuracies on TEST1 and TEST2 experiments were fully consistent with TEST3 results shown in Table II. These results are omitted due to lack of space, but supplementary tables with complete TEST1 and TEST2 results are provided at www.ist.temple.edu/~zoran/research/measures.pdf. Our experiments showed that, if a certain method was more accurate than an alternative method on TEST3, it was most often also more accurate on TEST1 and TEST2. In particular, in all three tests, the *GATING* method was the most accurate over all four measures. The results of the *GATING* method over three types of tests are compared in Table III.

TABLE III
GATING VERSUS AERONET AOD ACCURACY ON DIFFERENT TEST SETS. TEST1: UNOBSERVED LOCATION AND OBSERVED TIME. TEST2: OBSERVED LOCATION AND UNOBSERVED TIME. TEST3: UNOBSERVED LOCATION AND UNOBSERVED TIME

Test set	R^2	CORR	R_r^2	FRAC
TEST1	0.76	0.88	0.55	71.4%
TEST2	0.79	0.89	0.61	73.5%
TEST3	0.76	0.88	0.53	70.9%

Experiments over three types of tests showed that all methods were most accurate when tested on data at unobserved time but over previously seen locations (TEST2). Predicting AOD at unseen locations during the same two years (TEST1) was a more challenging objective but not as difficult as predicting AOD at unseen locations and in an unseen year (TEST3). These results suggest that, in our data, temporal correlation was stronger than spatial correlation and that both kinds of correlation could be exploited to improve the quality of AOD retrievals.

VI. CONCLUSION

To provide a predictor that is accurate over the standard and domain-specific accuracy measures, we have proposed developing an ensemble of neural networks with adaptive cost functions. In the proposed ensembles, some neural networks were specialized in predicting small AOD, while others were specialized in predicting large AOD. Instead of relying on MSE minimization criterion for neural network training, we have proposed using the relative error $REL_{a,b}$, which can be considered as a generalization of MSE. In experiments over the entire globe during the period of three years from 2005 to 2007, the proposed ensemble achieved R^2 and CORR accuracies as high as an ensemble relying on standard MSE optimization, while it significantly improved domain-specific R_r^2 and FRAC accuracies. In addition, AOD prediction using the proposed ensemble produces AOD values significantly closer to those of AERONET than the MODIS *C005* results.

Our statistical models were trained and evaluated using data collected at AERONET sites. Uncertainty analysis of such retrievals at other locations is a topic of our work in progress, and results on this aspect will be reported in a separate article.

REFERENCES

[1] L. A. Remer, D. Tanré, Y. J. Kaufman, R. Levy, and S. Mattoo, "Algorithm for remote sensing of tropospheric aerosol from MODIS: Collection 005," MODIS Algorithm Theoretical Basis Document ATBDMOD-04, 2006. [Online]. Available: http://modis-atmos.gsfc.nasa.gov/_docs/MOD04:MYD04_ATBD_C005_rev1.pdf

[2] N. J. D. Nagelkerke, "A note on a general definition of the coefficient of determination," *Biometrika*, vol. 78, no. 3, pp. 691–692, Sep. 1991.

[3] M. D. King, Y. J. Kaufman, W. P. Menzel, and D. Tanré, "Remote sensing of cloud, aerosol, and water vapor properties from the Moderate Resolution Imaging Spectrometer (MODIS)," *IEEE Trans. Geosci. Remote Sens.*, vol. 30, no. 1, pp. 2–27, Jan. 1992.

[4] B. N. Holben, T. F. Eck, I. Slutsker, D. Tanré, 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.*, vol. 66, no. 1, pp. 1–16, Oct. 1998.