Collaborative Regression on Aerosol Optical Thickness from Heterogeneous Remote Sensing Data

Bo Han

International School of Software, Wuhan University, Wuhan, 430079, China
bhan@whu.edu.cn

Abstract

Previous studies of aerosol optical thickness (AOT) estimations were generally based on observations from a single satellite sensor. Due to the limited observations from one instrument, the observations yielded AOT estimations with a system bias. In this paper, we combined two heterogeneous data sources, Moderate Resolution Imaging Spectroradiometer (MODIS) and Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP), together and proposed collaborative regression models to achieve more accurate AOT estimations than a single sensor does. These two independent remote sensors in the A-train satellite constellation both provide global AOT retrievals and they scan the same location on the Earth surface within a two-minute interval. However, the two remote sensors have different design principles respectively and their heterogeneous observation data streams pose challenges for information fusion. In our study, we proposed two types of heterogeneous collaborative regression approaches. One type of collaborative regression approach fuses information in a feature level. The other type of collaborative approach combines information in a model level. In our study, in each level, we apply a linear regression collaboration model and a neural network collaboration model. The proposed approaches are evaluated based on global observation data from MODIS and CALIOP during April 2, 2009 and April 1, 2011. The encouraging experimental results show that the regression approach collaborating in a model level achieves significantly more accurate AOT estimations than the results from the collaborative regression approach in a feature level. It also obtains significantly superior results to the deterministic AOT retrievals from any single satellite sensor.

Keywords: CALIOP, collaborative regression, linear regression, MODIS, neural network

1. Introduction

A key challenge for environmental observation satellite is inferring atmosphere parameters. One of the most important parameters is the aerosol optical thickness (AOT). It measures the degree to which aerosols prevent sun light passes through a column of atmosphere. AOT is important to study the Earth's radiation balance, environment protection and climate change. Traditional approaches of estimating AOT uses radiance data from a single satellite sensor. Due to the limited observations from one instrument, the observations yielded AOT estimations with a system bias [1-3].

In this paper, we aim to design a collaborative regression approach for improving AOT estimation accuracy by using multiple satellite sensor observations in a satellite constellation. Specifically, the A-train satellite constellation consists of multiple observation satellites running in a sun-synchronous orbit. They are spaced a few minutes apart from each other and all provide remote sensing observations of the Earth's atmosphere. MODIS (Moderate Resolution Imaging Spectroradiometer) and CALIOP (Cloud-Aerosol Lidar with Orthogonal Polarization) are two independent remote sensors aboard satellite Aqua and CALIPSO respectively in the A-train satellite constellation.

They both globally measure AOTs by deterministic approaches [4-9]. Especially, MODIS and CALIOP scan the same location on the Earth surface within a two-minute interval. The spatial-temporal synchronization provides us chance to combine the observation datasets together and make more accurate AOT estimations.

The challenge is that MODIS and CALIOP are designed with different principles and provide heterogeneous data streams. MODIS observes Earth from polar orbit in 36 wavelength bands ranging from 0.415μm to 14.5μm. Its deterministic algorithm separates the Earth surface from clouds by using 2.1μm wavelength band observations and directly derives AOTs with a forward simulation model (The estimations from a deterministic algorithm are called AOT retrievals). CALIOP, however, uses backscatter LiDAR signals to estimate the aerosol layer optical thicknesses and total column optical thickness. The different observations from the two sensors can complement with each other. Thereby, it will be very useful to build an enhanced AOT estimator by combining the two sets of heterogeneous remote sensing data.

Researchers worldwide have made efforts to estimate AOT by satellite observation attributes. Previous studies generally estimate AOT by using radiance data from a single remote sensor. Goo Jun *et al*. proposed a gaussian process regression approach to predict aerosol optical thickness from MODIS multispectral images [10]. Jhoon Kim *et al*. improved aerosol optical depth retrieval over Hong Kong from a geostationary ocean color imager satellite observations by using a critical reflectance method with background optical depth correction [11]. KostaRistovski *et al*. built a neural network model for predicting aerosol optical thickness based on MODIS observations and further analyzed its uncertainty [12]. Zhuang Wang *et al*. proposed a mixture model for multiple instance regression to estimate AOT and test the model on MISR and MODIS data independently [13]. Chul Eddy Chung globally estimate the anthropogenic aerosol direct forcing derived from MODIS information [14]. In addition, some research has compared CALIOP AOT with MODIS AOT, and found that they have different systematic biases due to different observation principles [15-17]. However, few researches have been performed to build a more accurate AOT estimator using heterogeneous satellite datasets observed from two or more remote sensors in a satellite constellation.

In this study, we proposed two collaborative regression approaches for AOT retrievals by fusion heterogeneous information: one in a feature level and the other in a model level. Experimental results based on MODIS, CALIOP data from April 2, 2009 to April 1, 2011 observations show that for the two heterogeneous sources, the collaborative regression approach in a model level outperforms the collaborative regression approach in a feature level. The model level collaboration also achieves significantly better AOT estimation accuracy than the AOT retrieval results from a single instrument.

The paper is organized as follows. Section 2 describes the data sets and accuracy measures; Section 3 proposes heterogeneous collaborative regression models; Section 4 compares AOT retrievals and proposed regression approaches collaborating in a feature level and in a model level based on a MODIS-CALIOP-AERONET synchronization dataset; Section 5 draws the conclusion of the collaborative AOT regression approach by using two remote sensors.

2. Data Sets and Accuracy Measures

In remote sensing field, the ground-based sun photometers AOT measures from Aerosol Robotic Network (AERONET) are generally used as the "ground truth" in most validation of satellite AOT retrievals. Thereby, to evaluate the accuracy of proposed heterogeneous collaborative regression approaches, we collected the spatial-temporal synchronized data among MODIS, CALIOP and AERONET from April 2, 2009 to April 1, 2011. The three datasets are compared in visible wavelength 550nm.

2.1. AERONET Data

AERONET is a federation of ground-based remote sensing aerosol networks over multiple hundreds of global observation sites collaborated by worldwide national agencies, institutes, universities, *etc*. It provides a continuous and readily accessible public AOT database for validation of satellite retrievals [18, 19].

In our experiments, we collected cloud-screened and quality-assured Level 2.0 AERONET data over 197 observation sites globally between April 2, 2009 and April 1, 2011. These sites vary on different surface types, including land, coast, desert and marine.

2.2. MODIS Data

MODIS is a key remote sensing instrument aboard the satellite Aqua and it views the entire Earth's surface every 1 to 2 days. MODIS acquires remote sensing data in 36 spectral bands [20].

In the experiments, we used MODIS/Aqua Collection 005 product suites between April 2, 2009 and April 1, 2011, including level 2 aerosol product MYD04_L2, calibrated radiance data MYD02SSH and cloud mask product MYD35. The three MODIS datasets are collocated in a spatial coincidence square region of size 40km×40km surrounding an AERONET site.

2.3. CALIOP Data

CALIOP combines an active LiDAR remote sensing instrument with passive infrared and visible imagers to probe the vertical structure and properties of thin clouds and aerosols over the globe [16, 21].

We collected CALIOP level 2 version 3 cloud-free aerosol layer products around AERONET sites by including those observations with quality control flag QC=0.

2.4. Spatial-Temporal Synchronization Data

MODIS, CALIOP and AERONET datasets are synchronized by similar spatial-temporal coincidence criteria as Ichoku *et al*. (2002). Spatially, MODIS or CALIOP radiance attributes are averaged in a 40km× 40km box with an AERONET station in the center. Temporally, AERONET observations are averaged within ±30 minutes of MODIS overpass.

By the above coincidence criteria, we collected 6351 records covering 197 AERONET sites globally in the MODIS-AERONET synchronization data. 486 collocated data points are found covering 82 AERONET sites globally in the CALIOP-AERONET synchronization data. The MODIS-CALIOP-AERONET synchronization data contains 322 collocated data records covering 65 global AERONET sites.

2.5. Accuracy Measures

The accuracy of our heterogeneous collaborative regression models is measured by correlation coefficient (Corr), R^2, mean square error (MSE) and percentage of outliers (Outlier%). The equations of the four measures are listed in equation (1) - (4). E_AOT denotes the AOT estimation of a regression model. A_AOT denotes AERONET AOT. N is the number of AOT retrievals. $\overline{E\_AOT}$, $\overline{A\_AOT}$ present the average of AOT estimations and AERONET AOTs respectively. The std(A_AOT) denotes the standard derivation of AERONET AOT retrievals.

$$Corr = \frac{\sum_{i=1}^{N}(E\_AOT_i - \overline{E\_AOT}) \cdot (A\_AOT_i - \overline{A\_AOT})}{\sqrt{\sum_{i=1}^{N}(E\_AOT_i - \overline{E\_AOT})^2 \cdot \sum_{i=1}^{N}(A\_AOT_i - \overline{A\_AOT})^2}} \tag{1}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(A\_AOT_i - E\_AOT_i)^2}{\sum_{i=1}^{N}(A\_AOT_i - \overline{A\_AOT})^2} \tag{2}$$

$$MSE = \frac{\sum_{i=1}^{N}(A\_AOT_i - E\_AOT_i)^2}{N} \tag{3}$$

If $|E\_AOT_i - A\_AOT_i| > 2 * std(A\_AOT)$

Then E_AOT_i is an outlier,

$$Outlier\% = \frac{Number\ of\ outliers}{N} \tag{4}$$

3. Heterogeneous Collaborative Regression Models

To obtain the full benefits from heterogeneous sensors, we apply collaborative regression approaches by fusion of information either in a feature level or in a model level.

In a feature level, we combine the spatial-temporal coincidence features from heterogeneous sensors together into a record and next build a regression model. Suppose the features in MODIS are described as <$MF_1, MF_2, \ldots, MF_p$>, p indicates the number of informative features which MODIS have. The features in CALIOP are listed as <CF_1, $CF_2, \ldots, CF_q$>, q suggests the number of informative features which CALIOP have. By collaborative regression in a feature level, we would like to build a model f, where

$$AOT = f(MF_1, MF_2, \ldots, MF_p, CF_1, CF_2, \ldots, CF_q) \tag{5}$$

In a model level, we apply regression on AOT retrievals from each individual sensor. AOT_m denotes the AOT retrievals from MODIS, AOT_c is the spatial-temporal synchronization AOT retrievals from CALIOP. By model collaboration, we combine the two deterministic model outputs together in the following equation,

$$AOT = f(AOT_m, AOT_c) \tag{6}$$

For the collaborative regression either in a feature level or in a model level, we explore both a linear regression estimator and a neural network estimator. For measuring its accuracy of each model, we apply cross-validation to build a separate training set and a test set for each model based on the MODIS-CALIOP-AERONET synchronization data.

3.1. Collaborative Linear Regression in a Feature Level (CLRFL)

With a linear regression method, the equation (5) can be rewritten as,

$$AOT = \beta_0 + \sum_{i=1}^{p} \beta_i \cdot MF_i + \sum_{j=1}^{q} \beta_{j+p} \cdot CF_j \tag{7}$$

Since many observation features are collected from MODIS and CALIOP, it will increase the complexity of computing a solution. Thereby, we firstly apply linear regression on MODIS features and CALIOP features respectively. Only those features with significant t-values will be kept for collaborative regression in equation (7).

3.2. Collaborative Neural Network Regression in a Feature Level(CNNRFL)

In this study, we build a feed-forward neural network (NN) with one hidden layer of neurons. It is illustrated in Figure 1. The network inputs include the attributes from both MODIS and CALIOP observations. The hidden layer consists of r hidden neurons, and the output is the AOT estimation.

The connection weights between nodes across different layers are learnt by back propagation with a gradient descent method.
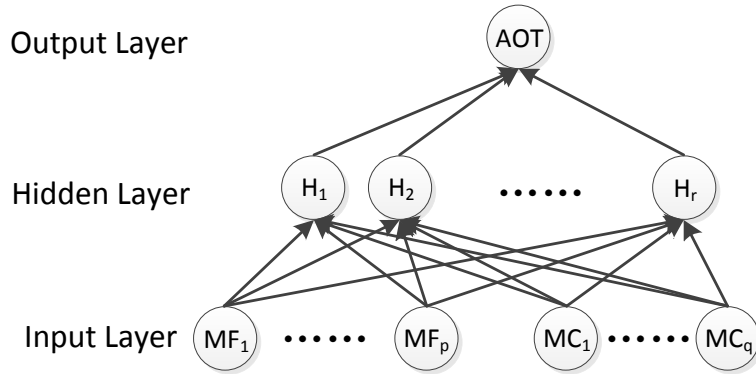

Figure 1. Neural Network Structure for CNNRFL

In our pilot study, we tried to practically apply principle component analysis (PCA) transformation to reduce dimension before the neural network training stage [22]. The experimental accuracy drops and we gave up the PCA preprocessing before a neural network modeling.

3.3. Collaborative Linear Regression in a Model Level (CLRML)

By comparing MODIS and CALIOP AOT retrievals with AERONET AOT records, we observe that the two satellite datasets show different system bias due to different design principles. In this heterogeneous context, we intuitively propose a collaborative regression model by linearly correcting their bias in equation (8).

$$AOT = \beta_0 + \beta_1 \cdot AOT_m + \beta_2 \cdot AOT_c \qquad (8)$$

3.4. Collaborative Neural Network Regression in a Model Level (CNNRML)

For a model level collaboration, we would like to build a feed-forward neural network (NN) with one hidden layer of neurons (Figure 2). It is similar to the neural network structure illustrated in Figure 1. The difference is that the input layer here includes only two attributes: the MODIS AOT retrievals M_AOT and the CALIOP AOT retrievals C_AOT.

4. Experimental Results

4.1. Experimental Settings and Optimization

The spatial-temporal synchronized data among MODIS, CALIOP and AERONET from April 2, 2009 to April 1, 2011 is used in our experiments. The learning regression target is the AOT retrieved by AERONET. For the collaborative regression in a feature level, the driving attributes collected from MODIS and CALIOP are used as learning variables. They are designed according to MODIS/CALIOP data descriptions and our previous research results [23, 24]. The detailed name and explanation of all driving

attributes are listed in Table 1. For the collaborative regression in a model level, the AOT retrievals by MODIS and CALIOP are used as inputs.
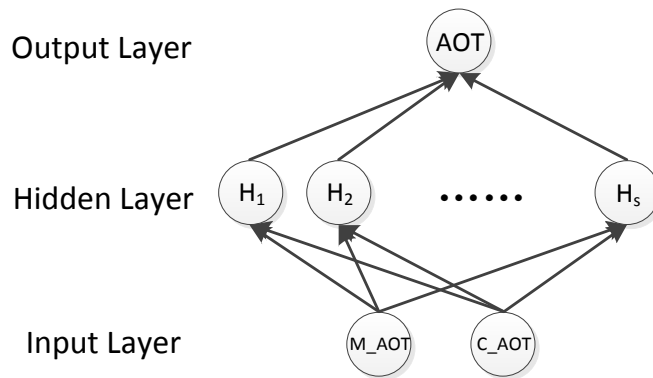

Figure 2. Neural Network Structure for CNNRML

For all methods, we apply 10-fold cross validation to access their accuracy. That is, the MODIS-CALIOP-AERONET synchronization dataset is randomly partitioned into 10 equal sized subsets. Among them, 9 subsets are used as training data for constructing a model for a method, and the remaining single subset is retained as the validation data for testing the method. The cross-validation process is then repeated 10 times, with each of the 10 subsets used exactly once as the validation data. The 10 test results then are combined to produce a single estimation. For each method, we repeated the 10-fold cross validation for 10 rounds. In each round, the synchronization dataset is randomly partitioned. All the results reported below are the average accuracy from 10 rounds experiments.

Table 1. Driving Attributes from MODIS and CALIOP

Source	Attribute Index	Name and Explanation
MODIS	1	AOT retrieval by MODIS
	2	NDVISwir– a parameter suggesting surface greenness
	3-9	Surface Reflectance at 7 different wavelength
	10	Solar zenith angle
	11	Sensor zenith angle
	12	Scattering angle
CALIOP	13	AOT retrieval by CALIOP
	14	IGBP surface type
	15	Aerosol subtype analyzed by CALIOP
	16	Solar zenith angle
	17	Scattering angle
	18	DEM surface elevation
	19-20	The layer effective multiple scattering factors at 532nm and 1064nm wavelength
	21-22	Pressure at the geometric midpoint of the thickest/thinnest layer
	23-24	Temperature at the layer base altitude for the thickest/thinnest layer
	25-26	Layer top altitude for the thickest/thinnest layer

For linear regression in a feature level or in a model level, we apply a linear regression method with least squares fit to construct a model.

For collaborative neural network regression in a feature level, we explored different types of neural transfer functions in the hidden layer and the output layer. By experiments, the best functions for the two layers are both set to 'purelin'. Table 2 summarizes the achieved accuracy for different number of hidden nodes in a neural network. By the four accuracy measures, the neural network with 6 hidden nodes shows to be an appropriate choice.

Table 2. Optimization of Neural Network for CNNRFL

No. of Hidden Nodes	Corr	R^2	MSE	Outlier%
3	0.759	0.557	0.014	7.27%
6	**0.807**	**0.651**	**0.011**	**7.14%**
9	0.776	0.597	0.013	7.82%
12	0.759	0.571	0.014	7.57%
15	0.738	0.540	0.015	7.05%
18	0.670	0.112	0.031	8.38%

Similarly, we optimized the collaborative neural network in a model level. We practically set both transfer functions in the two layers to 'purelin' by experiments. The regression accuracies with different number of hidden nodes are reported in Table 3. It shows the hidden layer with four nodes achieved the best performance.

Table 3. Optimization of Neural Network for CNNRML

No. of Hidden Nodes	Corr	R^2	MSE	Outlier%
1	0.818	0.652	0.011	1.05%
2	0.825	0.670	0.011	0.87%
3	0.839	0.699	0.010	0.59%
4	**0.848**	**0.714**	**0.009**	**0.31%**
5	0.836	0.689	0.010	0.66%
6	0.827	0.677	0.010	0.80%

4.2. Experimental Results

For comparing the collaborative regression results with deterministic retrievals from a single sensor, we illustrate their scatterplots in Figure 3. Figure 3(a) and Figure 3(b) intuitively present the comparisons between AERONET AOT with AOT retrievals from MODIS and CALIOP respectively. They show MODIS AOT and CALIOP AOT both have two retrieval outliers with AOT>1.4, though all AERONET AOT are less than 1.2. We checked the outliers' details. The two outliers for MODIS are a point with MODIS AOT=2.461, CALIOP AOT=0.860 and AERONET AOT=1.065, and another point with MODIS AOT=1.455, CALIOP AOT=0.536 and AERONET AOT=0.996. The two outliers for CALIOP are a point with CALIOP AOT=2.178, MODIS AOT=0.438 and AERONET AOT=0.663, and another point with CALIOP AOT=1.804, MODIS AOT=0.901 and AERONET AOT=0.826. All four outliers suggest that MODIS and

CALIOP derive outliers significantly differently. When one overestimates outlier AOT, the other underestimates its AOT. The differences show that, potentially we can combine both sensor measurements together to provide more robust and accurate AOT retrievals.
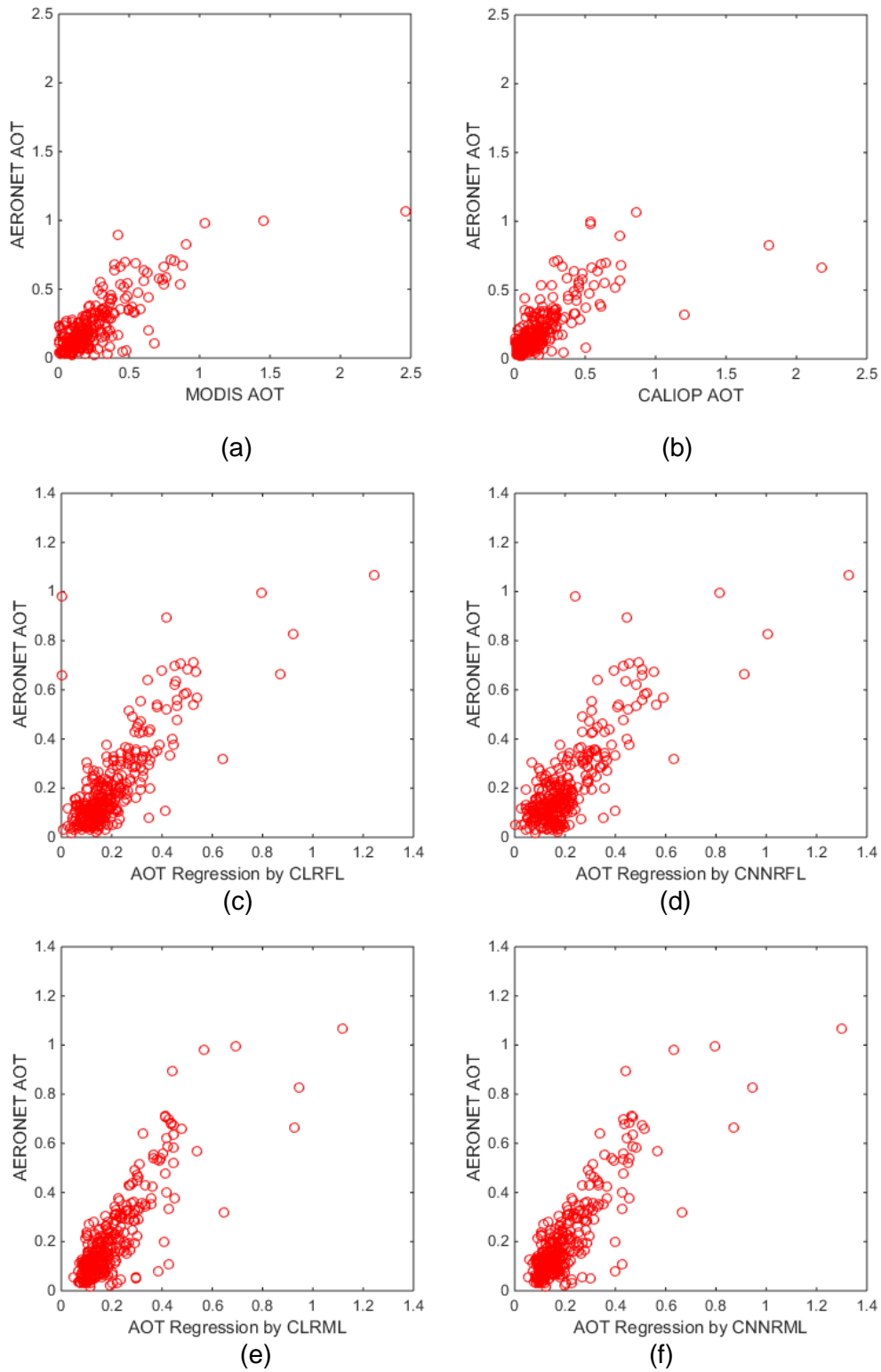

(a) (b)

(c) (d)

(e) (f)

Figure 3. AOT Regression Scatterplots

Figure 3(c) and Figure 3(d) show the collaborative regression results in a feature level by a linear regression method and a neural network method respectively. Figure 3(e) and 3(f) present the collaborative regression results in a model level by the two approaches. They suggest clearly that all four collaborative methods have produced the regression results without outliers (AOT>1.4). It suggests that the two sensors can compensate with each other: when one have bad performance, the other one can excel. This practically proves one advantage of collaborative models. Specifically, multiple points' AOTs are estimated by a linear regression model in a feature level with negative values. By environment science theory, they are truncated with a minimum value 0 in Figure 3(c).

Visually from the figure comparison, we see that the AOT estimations collaborated in a model level are more closely to the minus diagonal line than collaboration in a feature level. This can be explained that the remote sensing attributes collected by the two sensors are heterogeneous. Meanwhile, the deterministic AOT retrievals from MODIS and CALIOP have different system bias, which can be effectively reduced by collaboration in a model level. Thereby, a feature level combination cannot easily lead to better results than model level fusion.

The detailed regression results are summarized in Table 4. From the results, we see that both collaborative regression methods in a feature level can achieve the similar AOT retrieval accuracy as MODIS or CALIOP. The two collaborative regression approaches in a model level can achieve significantly more accurate AOT retrievals in all four measures.

Hence, both visualization results and measurement results confirm the effectiveness of collaboration models. From atmospheric theory, the physical and chemical processes involved in the lower atmosphere and multiple surface types in the Earth are very complicated, one satellite sensor designed by some specific theories may be strong in some cases for AOT estimations, but weak in other cases. A collaboration model can use data-driven approach to smooth the weakness and make the overall model more accurate.

Table 4. AOT Retrieval Accuracy by Different Methods

Method		Corr	R^2	MSE	Outlier%
MODIS AOT		0.804	0.647	0.019	2.17%
CALIOP AOT		0.719	0.517	0.024	2.80%
Collaborative Regression in a Feature Level	CLRFL	0.779	0.605	0.013	0.93%
	CNNRFL	0.807	0.651	0.011	7.14%
Collaborative Regression in a Model Level	CLRML	0.838	0.691	0.010	0.62%
	CNNRML	**0.848**	**0.714**	**0.009**	**0.31%**

5. Conclusions

In this paper, we explored two types of collaborative regression approaches by utilizing heterogeneous data sources MODIS and CALIOP in a satellite constellation to improve AOT retrieval accuracy. One type of collaborative regression approach fuses information in a feature level. It aims to integrate informative and potentially complement observation features from two sources for regression. The other type of approach combines heterogeneous information in a model level. It supposes MODIS and CALIOP have powerful deterministic AOT retrievals, but they have different system bias. The model level combination can complement their system bias and achieve better regression results. In each level of collaboration, we apply a linear regression method and a neural network regression method.

We evaluated the two types of collaborative regression approaches in the context of MODIS, CALIOP remote sensing data globally from April 2, 2009 to April 1, 2011. In our experiments, both AOT regression methods in a model level collaboration were clearly superior to both AOT retrievals from a single observation sensor and collaborative AOT regression in a feature level. The most accurate results were obtained through a neural network regression in a model level collaboration. It achieves MSE=0.009. In the feature level collaboration, a neural network regression is more accurate than a linear regression. It obtains AOT regression accuracy with MSE=0.011, 22% bigger than the results from a model level neural network collaboration results. For deterministic AOT retrievals from a single satellite sensor, MODIS and CALIOP achieves MSE=0.019 and MSE=0.024 respectively, which are 111% and 167% larger than the results from a model level neural network collaboration. These encouraging results suggest that in the context of heterogeneous remote sensing data sources in a satellite constellation, a model level collaborative neural network regression approach can significantly improve AOT regression accuracy. Exploring the spatial-temporal collaboration of more remote sensors in a satellite constellation for further improving AOT regression accuracy is a subject of our next steps research.

Acknowledgments

References

[1]   L. A. Remer, Y. J. Kaufman and D. Tanré, "The MODIS aerosol algorithm, products, and validation", Journal of the atmospheric sciences, vol. 62, no. 4, (2005), pp. 947-973.

[2]   R. C. Levy, L. A. Remer and S. Mattoo, "Second- generation operational algorithm: Retrieval of aerosol properties over land from inversion of Moderate Resolution Imaging Spectroradiometer spectral reflectance", Journal of Geophysical Research: Atmospheres, vol. 112, no. D13, (2007).

[3]   G. L. Schuster, M. Vaughan and D. MacDonnell, "Comparison of CALIPSO aerosol optical depth retrievals to AERONET measurements, and a climatology for the lidar ratio of dust", Atmospheric Chemistry and Physics Discussions, vol. 12, no. 5, (2012), pp. 11641-11697.

[4]   A. H. Omar, D. M. Winker, J. L. Tackett, D. M. Giles, J. Kar, Z. Liu, M. A. Vaughan, K. A. Powell and C. R. Trepte, "CALIOP and AERONET Aerosol Optical Depth Comparisons: One Size Fits None", Journal of Geophysical Research: Atmospheres, vol. 118, no. 10, (2013), pp. 4748-4766.

[5]   J. M. Livingston, J. Redemann, Y. Shinozuka, R. Johnson, P. B. Russell, Q. Zhang, S. Mattoo, L. Remer, R. Levy, L. Munchak and S. Ramachandran, "Comparison of MODIS 3 km and 10 km Resolution Aerosol Optical Depth Retrievals over Land with Airborne Sunphotometer Measurements during ARCTAS Summer 2008", Atmospheric Chemistry and Physics, vol. 4, no. 4, (2014), pp. 2015-2038.

[6]   M. Kacenelenbogen, J. Redemann, M. A. Vaughan, A. H. Omar, P. B. Russell, S. Burton, R. R. Rogers, R. A. Ferrare and C. A. Hostetler, "An Evaluation of CALIOP/CALIPSO's Aerosol-above-cloud Detection and Retrieval Capability over North America", Journal of Geophysical Research: Atmospheres, vol. 119, no. 1, (2014), pp. 230-244.

[7]   P. R. Colarco, R. A. Kahn, L. Remer and R. Levy, "Impact of Satellite Viewing-swath Width on Global and Regional Aerosol Optical Thickness Statistics and Trends", Atmospheric Measurement Techniques, vol. 7, no. 7, (2014), pp. 2313-2335.

[8]   R. Levy, L. Remer, R. Kleidman, S. Mattoo and C. Ichoku, "Global Evaluation of the Collection 5 MODIS Darktarget Aerosol Products over Land", Atmospheric Chemistry & Physics, vol. 10, no. 21, (2010), pp. 10399-10420.

[9]   Z. Liu, D. Winker, A. Omar, M. Vaughan, J. Kar, C. Trepte, Y. Hu and G. Schuster, "Evaluation of CALIOP 532-nm Aerosol Optical Depth Over Opaque Water", Atmospheric Chemistry and Physics, vol. 15, (2015), pp. 1265-1288.

[10] G. Jun, J. Ghosh, V. Radosavljevic and Z. Obradovic, "Predicting Ground-based Aerosol Optical Depth With Satellite Images Via Gaussian Process", International Conference on Knowledge Discovery and Information Retrieval (KDIR), (2010).

[11] J. Kim, M. S. Wong, J. Yoon, J. Lee and D. Wu, "Improvement of Aerosol Optical Depth Retrieval over Hong Kong from a Geostationary Meteorological Satellite Using Critical Reflectance with Background Optical Depth Correction", Remote Sensing of Environment, vol. 142, no. 02, (2014), pp. 176–187.

[12] K. Ristovski, S. Vucetic and Z. Obradovic, "Uncertainty Analysis of Neural Network-Based Aerosol Retrieval", IEEE Transactions on Geoscience and Remote Sensing, vol. 50, no. 2, **(2012)**, pp. 409-414.

[13] Z. Wang, L. Lan and S. Vucetic, "Mixture Model for Multiple Instance Regression and Applications in Remote Sensing", IEEE Transactions on Geoscience and Remote Sensing, vol. 50, no. 6, **(2012)**, pp. 2226 – 2237.

[14] C. Eddy Chung, V. Ramanathan, D. Kim and I. A. Podgorny, "Global Anthropogenic Aerosol Direct Forcing Derived from Satellite and Ground-based Observations", Journal of Geophysical Research, vol. 110, **(2005)**.

[15] C. Kittaka, D. M. Winker, M. A. Vaughan, A. Omar and L. A. Remer, "Intercomparison of Column Aerosol Optical Depths from CALIPSO and MODIS-Aqua", Atmospheric Measurement Techniques, vol. 4, **(2011)**, pp. 131–141.

[16] D. M. Winker, J. Pelon, J. A. CoakleyJr, S. A. Ackerman, R. J. Charlson, P. R. Colarco, P. Flamant, Q. Fu, R. M. Hoff, C. Kittaka, T. L. Kubar, H. Letreut, M. P. Mccormick, G. Mégie, l. Poole, K. Powell, C. Trepte, M. A. Vaughan and B. A. Wielicki, "The CALIPSO MISSION: A Global 3D View of Aerosols and Clouds", American Meteorological Society, **(2010)**, pp. 1211-1229.

[17] X. Ma, K. Bartlett, K. Harmon and F. Yu, "Comparison of AOT between CALIPSO and MODIS: Significant Differences over Major Dust and Biomass Burning Regions", Atmospheric Measurement Techniques, vol. 6, **(2013)**, pp. 2391-2401.

[18] B. N. Holben, T. F. Eck, I. Slutsker, D. 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 Sensing Environment, vol. 66, **(1998)**, pp. 1–16.

[19] O. Dubovik, A. Smirnov, B. N. Holben, M. D. King, Y. J. Kaufman, T. F. Eck and I. Slutsker, "Accuracy Assessments of Aerosol Optical Properties Retrieved from Aerosol Robotic Network (AERONET) Sun and Sky Radiance Measurements", Journal of Geophysical Research , **(2000)**.

[20] L. A. Remer, D. Tanre´, Y. J. Kaufman, C. Ichoku, S. Mattoo, R. Levy, D. A. Chu, B. Holben, O. Dubovik, A. Smirnov, J. V. Martins, R. R. Li and Z. Ahmad, "Validation of MODIS Aerosol Retrieval over Ocean", Geophysical Research Letters, **(2002)**.

[21] A. H. Omar, D. M. Winker and C. Kittaka, "The CALIPSO Automated Aerosol Classification and Lidar Ratio Selection Algorithm", Journal of Atmosphere and Oceanic Technology, vol. 26, **(2009)**.

[22] A. Sharma and K. K. Paliwal, "Fast Principal Component Analysis Using Fixed-point Algorithm", Pattern Recognition Letters, vol. 28, **(2007)**, pp. 1151–1155.

[23] B. Han, X. Gao and X. Cui, "Model-Based Sensitivity Analysis on Aerosol Optical Thickness Prediction", International Journal of Distributed Sensor Networks, vol. 2015, **(2015)**, Article ID 326132.

[24] S. Vucetic, B. Han, W. Mi, Z. Li and Z. Obradovic, "A Data Mining Approach for the Validation of Aerosol Retrievals", IEEE Geoscience and Remote Sensing Letters, vol. 5, no. 1, **(2008)**, pp. 113-117.

Authors

Bo Han, he received the B.S. degree in computer science and its applications from Wuhan University of Technology, Wuhan, China in 1993, and the M.S. degree in computer science from Wuhan University, Wuhan, China in 1996, and the Ph.D. degree in computer and information science from Temple University, Philadelphia, USA in 2007. He is currently an Associate Professor with the International School of Software in Wuhan University, Wuhan, China. His research interests are in spatio-temporal data mining, text mining and machine learning.