Journal of Quantitative Spectroscopy & Radiative Transfer 153 (2015) 119–130

Contents lists available at ScienceDirect

Journal of Quantitative Spectroscopy & Radiative Transfer

journal homepage: www.elsevier.com/locate/jqsrt

Transfer learning used to analyze the dynamic evolution of the dust aerosol

Yingying Ma [1], Wei Gong *, Feiyue Mao [2]

State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan, Hubei 430079, China

ARTICLE INFO

Article history:
Received 2 June 2014
Received in revised form
13 September 2014
Accepted 29 September 2014
Available online 13 October 2014

Keywords:
CALIPSO
CloudSat
Dust aerosol
TrAdaboost
SVM

ABSTRACT

To keep the advantage of Support Vector Machine (SVM) in analyzing the dynamic evolution of the dust aerosol, we introduce transfer learning as a new method because transfer learning can utilize knowledge from previously collected data and add dozens of new samples, which can significantly improve dust and cloud classification results. It can also reduce the time of sample collection and make learning efficient. In this paper, we receive significant improvement effect using SVM as the basic learner in TrAdaBoost during four consecutive dust storm days, and correct one error classification in PDF. As a result, dust aerosol in high altitude can even spread to stratosphere. Moreover, in the process of dust aerosol transportation, it is highly affected by anthropogenic aerosol, for example, the color ratio (CR) changes from 0.728 to 0.460 and finally reaches 0.466, while depolarization ratio (DR) changes from 0.308 to 0.081 and finally reaches 0.156. It is indicated that the big size and non-spherical aerosol particles reduce obviously after dust aerosol deposition, but small size and spherical anthropogenic aerosol also produce a certain effect, and on March 22, 2010 had a small recovery above the ocean following the reduction of DR and CR. Due to the MODIS resolution not meeting the observation requirement and layer identification being different between CALIPSO and CloudSat, a problem such as stratocumulus cloud in low altitude still exists in aerosol and cloud classification. Lack of ground-based auxiliary data is the main problem which hinders our validation and quantitative analysis. It is pressing for a solution in future.

1. Introduction

As the principal Asian dust sources, there are several dust storms happening in Northwest China and Mongolia every spring [1,2]. Generally, it not only influences the local climate, but also impacts on global climate change after the long distance transportation from west to east, even realizing transpacific transport and research North America [3–5]. Further, Asian dust aerosol contributes to most of the dust aerosol loading in the troposphere over the mid-latitude regions from East Asia to western North America during springtime [6]. The primary ways in which aerosols influence climate change include (1) direct effect, influence the radiative energy budget by scattering and absorbing solar radiation; (2) indirect effect, altering cloud droplet size distribution and concentration; (3) semi-indirect effect, absorbing aerosols could enhance cloud evaporation by contributing large diabatic heating in the atmosphere [7–9]. Evidently, observing and analyzing the Asian dust characteristics are important for us to

* Corresponding author. Tel.: +86 13517241678.
E-mail addresses: yym863@gmail.com (Y. Ma), weigong@whu.edu.cn (W. Gong), maofeiyue@aliyun.com (F. Mao).
[1] Tel.: +86 15827142071.
[2] Present address: School of Remote Sensing and Information Engineering, Wuhan University. Tel.: +86 15994220024.

http://dx.doi.org/10.1016/j.jqsrt.2014.09.025
0022-4073/

understand the mechanism of how a dust aerosol influences climate change [10,11].

The most popular method used in big scale observation is satellite remote sensing, and passive remote sensor still holds a dominant position. However, both Alpert and Jeffrey [12,13] pointed out that information about the vertical distribution of mineral desert dust is required for different climatological applications. Therefore, lidar observation is irreplaceable.

The Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP) is a space-based lidar system onboard the Cloud Aerosol Lidar and Infrared Pathfinder Satellite Observations (CALIPSO) satellite launched in 2006 [14]. Precisely, it is a two-wavelength (532 nm and 1064 nm), polarization sensitive (532 nm) elastic backscatter lidar which could obtain depolarization ratio and color ratio of the atmospheric layer. Through this ability, it can describe the particles size and shape features [15,16], it even has unique capabilities for measuring dust aerosol [17]. At the beginning, Vaughan et al. [18] introduced the overall retrieval algorithms and Liu et al. [19] introduced the discrimination method of cloud and aerosol by probability distribution functions (PDFs). But the classification accuracy is not very good. Afterwards, other methods [20–22] by combining passive remote sensor and multi-satellite data emerged, for example, Chen's method [20] used brightness temperature information to compensate the weakness of lidar, by adding a large number of ground station data, they validated their method quantitatively and received the significantly improved results, and Liu's new method [23] greatly improved the classification accuracy between cloud and dust aerosol. Based on their outstanding work, people can apply these valuable data to carry out their research. For example, by using this space-borne lidar, we know that Tibet also suffered the inversion of dust [24], the spatial distribution features of dust layer in this area [25]; by combining the other radiation model with CALIOP data, more full utilization of dust aerosol radiative forcing and heating rate are achieved [7].

However, if we only limit the observation to dust source area it is not enough. In recent years, more and more researches have focused on dust transportation, particles characteristics change and it has different effects on radiation in the whole transport process. Initially, the study was only developed by model prediction [26], then CALIOP helped them to analyze the spatial distribution of the dust layer; these regions almost focus on North Africa [27–29], but there are obvious differences between African dust and Asian dust: firstly, Sahara dust usually happens in summer and Asian dust happens in spring [30,31]; secondly, their formation mechanisms are different, Asian dust is generally associated with Mongolian cyclones and the passage of cold fronts in Northern China and Southern Mongolia which makes the suspension height of Asian dust higher and life time longer [6,26,32]; thirdly, evolution results are also different, African dust can remain largely uncontaminated by other types of aerosols when it crosses the Atlantic Ocean, but Asian dust is usually influenced by the anthropogenic aerosol above East China, Korea and Japan, when it transports westward [33–35]. Consequently, a study of the transportation of Asian dust which is based on CALIPSO has been carried out. Huang's study [35] shows CALIOP can distinguish the different effects between Taklimakan and Gobi, including their frequency and intensity. In addition, Liu focused on the changes of dust optical properties during transportation, by combining multi-datas and models, he gave the preliminary analysis and explanation of PDR and AOD (aerosol optical depth) change character over the entire process [36]. To ensure the quality of these works, the correction identification of dust layer cannot be left. How to improve the classification accuracy while reducing manual intervention and increasing computing speed?

In this paper we introduce a new classification algorithm, which is based on transfer learning. By comparing the difference of CALIOP Level 2 data VFM (vertical feature mask), the MODerate resolution Imaging Spectro radiometer (MODIS) remote sensing image and CloudSat Level 2 data cloud classification, we detect some controversial cloud and aerosol classification, and then give the experimental result by transfer learning, and finally, we describe the advantage and disadvantage of this algorithm. The purpose of our research is to just increase the application of CALIPSO and make it more accurate when used as input data in global climate models.

2. Algorithms

There are two stages promoting the development of cloud–aerosol discrimination (CAD), first is the 3D probability distribution functions (PDFs) [19], the feature spatial include the height of the atmospheric layer, attenuated backscatter coefficient, and color ratio. But the classification results of dust and cloud are not very good [20]. Because both ice cloud and dust particle have the non-spherical feature, if the algorithms add depolarization ratio as the new feature hastily, the result will still be unpredictable; if the algorithms abandon this feature, the distinction between ice cloud and dust aerosol will not be apparent. Consequently, to solve the problem, 5D PDFs has been proposed [23]. This new algorithm adds the geographical information and depolarization ratio simultaneously, and makes the classification result between cloud and dust aerosol improve effectively.

The prophase of our research which is based on Support Vector Machine (SVM) also focuses on cloud and dust aerosol discrimination, and has achieved initial results [37,38] and gained some recognition [39–41]. Although it has the advantage of requiring less sample amount, achieving quick calculation and not needing to simulate the sample distribution, an obvious shortage also exists, for example, if we extend the study area from Northwest China to North, Middle and East China, this method loses classification effect instantly, this is because the dust characteristics have changed. How to improve the scalability of our former research is the chief content of this paper. Below we will review our former algorithm and analyze the limitation, and then introduce the new method.

2.1. Status of SVM application in dust aerosol identification

In this section, as already mentioned [38], we do not intend to introduce mathematical formulation and theory but mainly introduce the application of the algorithm. In the beginning, considering the error ratio in CAD version1,