

Dear Editor and Reviewers,

We highly appreciate the detailed valuable comments from the referees on our manuscript of “acp-2016-195”. The suggestions are quite helpful and we have incorporated them in the new version of manuscript. We have referred to literatures and papers and re-analyzed the collected data and reconstructed the paper to improve the quality of our paper.

As below, I would like to clarify some of the points raised by the Reviewers. And we hope that the reviewers and the editors will be satisfied with our responses to the ‘comments’ and the revisions for the original manuscript.

Thanks so much.

Yours Sincerely,

Yong Xue

In this study, authors validate three AATSR AOD products (ADV, ORAC and SU algorithm) provided by Aerosol_cci project over China in 2007, 2008 and 2010. It's been widely validated (compared with AERONET AOD) that these three algorithms have ability in retrieving AOD over land with high precision till 3/23/2016. However, the AERONET data has limitations as reference data that there were not enough AERONET sites built and the distribution of AERONET sites were unevenly in mainland China in 2007, 2008 and 2010 caused by large territory of mainland China. Authors introduce CARSNET data to be combined with AERONET data, making up for these limitations and improving reliability of reference data. On this basis, authors not only select common evaluation metrics, but also introduce new metrics, for example, the improved KAPPA coefficient as comprehensive evaluation metric, the DR for determination of AOD retrieved "outliers", the improved expected error envelope designed for characteristics of AATSR AOD products, etc. This study is a nice trial consisting many meaningful works and I would recommend publication if my following comments/suggestions can be adequately addressed.

Many thanks for your positive comments.

Major comments:

1. The structure and composition of manuscript should follow the requests of official website of Atmospheric Chemistry and Physics (ACP). For example, keywords, team list, etc. should add to manuscript and team list exist in this manuscript.

Response: All required structure and composition have been added in revised version of manuscript.

2. Figures in the manuscript should be clear and easily understood. The main method of this study is to validate three AATSR AOD products year by year for reason of different reference data available for authors. Readers could distinguish which sites in the Fig. 1 is from AERONET or CARSNET, but may not pick out the space distribution of ground-based data sites in same year easily. I recommend authors replot Fig. 1 of "The distribution of selected AERONET&CARSNET sites in mainland China in 2007, 2008 and 2010", using one same color or type for sites available in one year.

Response: The figures have been revised with clear and easy understood symbols and text, as Fig. 1 we have revised the symbols of sites from different networks using different colors to make it clearer.

3. Also I suggest that the paper never use the word "good" to describe the results. The coefficient of correlation (CC) as one of main evaluation metric, which indicates whether there is any linear relationship among the points. Authors could not claim which performances of products is "good" or not "good" by any values of CC or other evaluation metrics. For example, when CC is high, the performances could be viewed as "good", when CC is low, the performance is also viewed "good". The word "good" may confuse readers, leading misunderstanding of conclusions in this

study.

Response: The “good” or “well” terms have been replaced by quantitative description or comparative words. For example, in section 4 to section 6, we have refined our analysis using more detailed quantitative description to present readers easy understood analysis.

Specific comments:

Page 2 line 9, the influences of aerosol particles on cloud should cited the paper of Twomey published in 1974.

Response: This reference will be added.

Page 2 line 19, the word “because” should be replaced by other words like “including”

Response: This sentence has been revised.

Page 3 line 10, the word “more” should be removed

Response: This word has been removed.

Page 3 line 15-19, comparison of satellite retrievals with other high quality has limitations, could you illustrate it more clearly?

Response: We have added necessary illustration and cross-validation with MODIS C6 DT&SB merged datasets.

Page 3 lin26, is “Aerosol_CCI” or “Aerosol_cci” formal?

Response: The CCI official website uses “Aerosol_cci”, therefore, we’ll introduce “Aerosol_cci” in the following or revised paper.

Page 4 Tab. 1 the bottom row are same with header row, what’s it useful? And in the title abbreviation “Tab.” should avoid.

Response: Tab. 1 has been revised.

Page 9 Tab. 4 these statistics should be up to two decimal point.

Response: Relative statistics have been kept two places of decimal.

Page 9 line 3, this sentence has syntax error.

Response: This sentence has been revised in new version of manuscript.

Fig. 8 – Fig. 16, the places of titles should be same.

Response: The places of title in figures have been adjusted to the same.

Page 19 Tab. 5 these statistics should be up to two decimal points.

Response: Relative statistics have been kept two places of decimal as Tab. 4.

Page 24 line 29, in part of acknowledgements, the numbers of sites are inconsistent

with mentioned as above.

Response: The numbers of ground-based sites have been corrected in revised version of manuscript.

Dear Editor and Reviewers,

We highly appreciate the detailed valuable comments from the referees on our manuscript of “acp-2016-195”. The suggestions are quite helpful and we have incorporated them in the revised paper. We have referred to literatures and papers and re-analyzed the collected data and reconstructed the paper to improve the quality of our paper.

As below, I would like to clarify some of the points raised by the Reviewers. And we hope that the reviewers and the editors will be satisfied with our responses to the ‘comments’ and the revisions for the original manuscript.

Yours truly,

Yong Xue

Interactive comment on “Inter-comparison of three AATSR Level 2 (L2) AOD products over China” by Y. Che et al.

Anonymous Referee #2

Received and published: 25 March 2016

This paper validates 3 algorithms (SU, ADV, ORAC) for determining AOD from the European AATSR sensor against Sun photometer data in China (from AERONET and CARSNET). The topic is relevant to ACP. The work is important because these European products have not been as well-known as NASA ones, and have undergone a lot of development in the European CCI projects, so it is a good time to do some more thorough validation of these data sets. This is especially true for China since the aerosol loading is high and variable, and CARSNET has monitoring stations in some areas where AERONET is lacking.

I have read through the manuscript several times and, while it is promising, there are some things which are unclear/invalid or I think not useful, and some important things which should be added to make the analysis more complete and useful. The phrasing is odd in some places and there are a number of typos (e.g. AEROENT in some places instead of AERONET) so I think that the manuscript will need some copy editing by the production office. I appreciate that English is not the authors' first language and the writing is not bad, it is just unclear in some cases. I therefore recommend some content revisions, to address the points below. I would like to review a revised version and think that another round of reviews will be necessary because the structure/content of the manuscript might change a lot. Here are my main points:

Response: The English of the manuscript has been edited by the Elsevier's Language Services.

1. Abstract: some of the sentences could probably be shortened (e.g. first and second can be combined, as can third and fourth).

Response: The sentences have been shortened and refined in new version of abstract.

2. Introduction and start of section 2: It would be good to add a bit more information about the AATSR sensor here, like launch/end dates, swath width. A brief

discussion of the differences between the algorithms should be included, to help understand why they give different results. From the analysis, the performance and the spatial coverage are both different between the three algorithms, so some insight into what in the algorithms is responsible would be welcome.

Response: More details and information about the AATSR instrument and retrieval algorithms have been added, furthermore, a brief analysis and discussion of the differences between the retrieval algorithms have also been added in the revised version of manuscript. These information will help readers to have a deep insight about the differences of validation result we have made in this study.

3. Statistics. Some of the metrics presented here are questionable in their relevance and I think that there are simpler and clearer alternatives. Specifically, the EE envelopes quoted here for Equations 1, 2 are for the MODIS instrument, not AATSR. AATSR is quite different (two views, fewer wavelengths) so there is no reason to expect that an AOD retrieval for AATSR would have the same type of behavior. One might expect that the error formulation would be closer to that of MISR. Further, Equations 3 and 4 are basically expressing a confidence envelope around a regression line. This is not really useful since it is just the noise around the relationship and not so dependent on the actual error in the retrievals. So comparing this between algorithms does not really make sense. A well-correlated but very biased retrieval would appear 'better' by this metric than a poorer-correlated but unbiased one, while for an actual scientific application, the unbiased one may in some cases be more useful.

Further, least-squares linear regression is invalid for AOD retrievals because aerosol data violate the assumptions of this technique (see e.g. <http://people.duke.edu/~rnau/testing.htm> ; the AOD data validate assumptions 3 and 4 that linear least-squares regression makes, possibly 1 and 2 as well, and as a result the results obtained are not statistically valid). I know that a lot of people do least-squares linear regression because it is easy, but it is still wrong for this application.

So, a better alternative is just to present statistics of bias and RMS error as a function of AOD, similar to what is shown in e.g. Figure 5 and the magenta bins in Figure 2. So I suggest that the EE discussion here and linear regressions be discarded entirely, and more prominence should be given to statistics subset into different regimes (e.g. low AOD, moderate AOD, high AOD; perhaps also splits based on Angstrom exponent for the high-AOD regime) , as retrieval errors are often type-dependent as well). The kappa coefficient is probably fine. So, accounting for this comment would somewhat streamline and improve sections 2 and 3. Finally, presenting statistics to 3 significant figures is overkill and paints a picture of them being more robust than they probably are; 2 significant figures is probably enough.

Response: One objective of this manuscript is to evaluate different statistical metric for the validation of quantitative remote sensing. Different statistical metric shows different meaning and is used for different purpose. Linear regression is the most basic and commonly used statistical method that allows us to summarize and study relationships between two quantitative variables. Pearson correlation coefficient (CC) measures the fraction of the total variability in the response that is

accounted for by the retrieval and is only a measure of linear association between ground truth measurements and satellite retrieval values. Bias describes the average difference between satellite retrievals and ground AOD. For the consistency of the metric among different aerosol products, it is better to show the percent of retrievals falling within the expected error (EE) range.

4. Retrieval errors. As I understand it, the CCI project means that the data products also provide their own estimates of the uncertainty in the AOD retrieval for every pixel. This is an important point, since pixel-level uncertainties are very useful for many applications. However, it is not discussed in the paper. How do these uncertainty estimates compare to the retrieval errors observed?

Response: The satellite retrieved AOD in each collocated pair are means of retrievals in 5×5 sampling frame. On this basis, we calculate means of uncertainty estimates in sampling area for each collocated pair as sizes of circles in scatter plot. In section 4 and 5, we reanalyze validation results of different algorithms, including comparison of uncertainty estimates and retrievals error observed.

5. Figures 2-4 and discussion in Section 3. I don't see any advantage to splitting out the points by year. It would be easier to combine all points from one algorithm into one panel, not 3. This would also let you combine Figures 2-4 into one figure for a side-by-side comparison of the three algorithms. Also, as discussed before, I would delete the regression and EE lines here since they are not very meaningful. The magenta symbols and lines for the binned data are enough here. Also, the color scales used in these figures are not mentioned and can probably be removed (either show individual points without a color scale or a density plot with a color scale).

I also don't see any good reason to split the discussion of statistics out by year either. The data volume is not very large, so year to year differences are probably resulting from sampling and not statistically meaningful. Looking at the bigger picture of all data together is more statistically robust and gives a clearer picture. I don't believe any insight is gained by splitting the analysis up year by year.

Response: All points from one algorithm have been combined to make results more statistically robust and remove unnecessary plots. The colors of points in new scatter plot represent standard deviation of retrievals in sampling area for the purpose of finding influence on retrieving performance of sampling. We also keep the comparisons for each year as we would like to see the differences for each year. We added one section on the analysis of seasonal behaves of three algorithms.

6. Figures 5-7: Similar to the last comment: are the different panels the different years? It doesn't say anywhere but I infer that is the case. Again, these figures could be streamlined into one because clearly the biases are similar between years, this will be more robust, and will allow for a more direct comparison of the 3 algorithms. Additionally, I don't think the histograms (bottom panels) here are useful since they don't provide any information which is not seen clearly in

the top panels, so these could be deleted. Also, for the same reason as before, the linear regressions are invalid and should be deleted, just showing the binned values is enough.

Response: New scatter plots have been made, combining all points from one algorithm.

7. A similar plot to the bias plots could be created for RMSE. This would be a clearer way to show and compare the AOD-dependence of the retrieval error than the EE3/EE4 metrics.

Response: RMS error has been added in plot and statistic table.

8. In the discussion of the results, a lot of the time terms like “good” and “well” are used to describe performance. These are “weasel words” and should be avoided. What is “good” is only really relevant relative to a specific application (e.g. good enough to do X) or compared to the state of the art. I suggest rewording to avoid these words and be more quantitative where possible, or else stick to comparative terms (e.g. say when the data sets are similar to or better than each other). Also, some discussion of results compared to validation of other sensors (the main ones being MODIS/MISR) could be included, as these all have published validation for their aerosol products, and this would give a sense of how the AATSR data perform relative to the other available data products. Right now the paper more or less reads like AATSR is the only satellite option.

Response: The “weasel words” like “good” or “well” have been replaced by details of RMSE and KAPPA coefficient or comparative words. We compare and analyze AATSR AOD with “Deep Blue” and “Dark Target” 10km×10km AOD data from MODIS Collection 6 datasets which has been widely validated.

9. Do the retrievals provide other information like Angstrom exponent? From other references, I believe so. This quantity is commonly compared with AERONET measurements, so it should be easy to extend the analysis to look at this as well using the same basic approaches. This might provide more insight for the differences between the data sets, if the algorithms are making very different assumptions about what sort of aerosol is present. This would help overcome one of the weaknesses of the paper, i.e. that the comparison is presented without any sort of discussion about why the three data sets are different and how to improve them (which would be very useful information).

Response: The CARSNET dataset provides AOD and angstrom exponent (440-870) only, otherwise the ADV provides angstrom exponent (550-670) only, ORAC provides angstrom exponent (550-870) only and SU provides angstrom exponent (550-870) only. Comparison between these data may be invalid.

10. There are at least two more AERONET sites in China which provided data in the study period, but which were not used in the analysis. These are both

in Hong Kong: Hong_Kong_Hok_Tsui and Hong_Kong_PolyU. Why were these not used? If the objective (as stated) is to provide coverage over broad areas of China, then it would make sense to include them, since the data are freely available and there are no other sites used in this part of China. These sites are very close to the coast so also provide an additional type of environment to analyze, compared to the other sites presently included in the study. Additionally, it will boost the data volume. I suggest adding these sites to the analysis. There may be more, these were the main ones which sprang to mind. On a related note, Figure 1 can probably be simplified for clarity by using one symbol/color for all AERONET sites, and another for all CARSNET sites. Splitting by year isn't necessary, in my view, and just complicates things.

Response: The AERONET sites are added, including those in HongKong.

11. The title of the manuscript suggests a broader scope than the analysis, since the analysis only performs an inter-comparison in the context of AERONET/CARSNET measurements. There are various other things which could be added, at least briefly. For example, climatologies of seasonal AOD from all three algorithms (from the 1 degree products), and maps showing the available data volume (e.g. number of days per season with data), since this is another feature which is important for many applications. Otherwise, the title should be amended to reflect the scope. However I would prefer that the analysis be extended because I think that this would be quite useful (and new, to my knowledge).

Response: The seasonal validation and analysis has been added, and we also take insight into more analysis to make the scope broader.

Inter-comparison of three AATSR Level 2 (L2) AOD products over China

Y. Che^{1,6}, Y. Xue^{1,2}, L. Mei³, J. Guang³, L. She^{1,6}, J. Guo⁴, Yincui Hu⁵, H. Xu³, X. He^{1,6}, A. Di^{1,6} and C. Fan^{1,6}

5 Y. Che^{1,5}, Y. Xue^{1,2}, L. Mei³, J. Guang³, H. Xu³, L. She^{1,5}, J. Guo⁴, X. He^{1,5}, A. Di^{1,5} and C. Fan^{1,5}

¹State Key Laboratory of Remote Sensing Science, jointly sponsored by the Institute of Remote Sensing and Digital Earth of the Chinese Academy of Sciences and Beijing Normal University, Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, Beijing 100101, China

²Faculty of Life Sciences and Computing, London Metropolitan University, 166-220 Holloway Road, London N7 8DB, UK

10 ³Key Laboratory of Digital Earth Science, Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, Beijing 100094, China

⁴Centre for Atmosphere Watch and Services, Chinese Academy of Meteorological Sciences, 46, Zhongguancun South Avenue, Haidian District, Beijing 100081, China

15 ⁵Hebei Key Laboratory of Environmental Change and Ecological Construction, College of Resources and Environment Science, Hebei Normal University, Shijiazhuang, Hebei Province, China

⁶University of Chinese Academy of Sciences, Beijing 100049, China⁵University of Chinese Academy of Sciences, Beijing 100049, China

Correspondence to: Professor Y. Xue (yx9@hotmail.com)

Abstract. The Advanced Along-Track Scanning Radiometer (AATSR) aboard ~~an~~ ENVISAT is used to observe the Earth ~~by~~
20 ~~in~~ dual-view. The AATSR data can be used to retrieve aerosol optical depth (AOD) over both land and ocean, which is an
important ~~merit-parameter~~ in the characterization of aerosol properties. In recent years, aerosol retrieval algorithms have been
developed both over land and ocean ~~have been developed~~, taking advantages of the features of dual-view, which can help
eliminate the contribution of Earth's surface to top of atmosphere (TOA) reflectance. The Aerosol_cci project, as a part of the
Climate Change Initiative (CCI), provides users with three AOD retrieval algorithms for AATSR data, including the Swansea
25 algorithm (SU), the ATSR-2ATSR dual view aerosol retrieval algorithm (ADV), and the Oxford-RAL Retrieval of Aerosol
and Cloud algorithm (ORAC). The ~~v~~Validation team of the Aerosol-CCI project has validated AOD (both Level 2 and Level
3 products) and AE (Level 2 product only) against the AERONET data in a round robin evaluation using the validation tool
of the AeroCOM (Aerosol Comparison between Observations and Models) project. For the purpose of evaluating different
performances of these three algorithms ~~on-in~~ calculating AODs over mainland China, we introduce ground-based data from
30 the CARSNET (~~the~~ China Aerosol Remote Sensing Network), which ~~is-was~~ designed for aerosol observations in China.
Because China is vast in territory and ~~of-has~~ great differences in terms of land surfaces, the combination of the AEROENT
and ~~the~~ CATRNET data can validate the L2 AOD products more comprehensively. The validation results show different
performances of these products in 2007, 2008 and 2010. The SU algorithm ~~has-very-good-performance~~ performs very well over
sites with different surface conditions in mainland China from March to October, ~~but however,~~ but it slightly underestimates

AOD ~~over barren or sparsely vegetated surfaces in western China, slightly with varying~~ mean bias error (MBE) ~~ranging~~ from 0.05 to 0.10 ~~over surface of barren or sparsely vegetation in western China~~. The ADV product has ~~the~~ same precision with ~~a~~ high correlation coefficient (CC) ~~larger than 0.90 over most of sites and the same error distribution as the SU product~~. The main limits of ~~the~~ ADV algorithm are underestimation and applicability; ~~especially it occurs obvious~~ underestimation ~~is~~
5 ~~particularly obvious~~ over ~~the~~ sites of Datong, Lanzhou and Urmuchi, where the ~~dominated dominant~~ land cover is grassland, with MBE larger than 0.2, and the main ~~source of aerosol sources is are~~ coal combustion and dust. The ORAC algorithm has the ability ~~of retrieving to retrieve~~ AOD at different ranges including high AOD (larger than 1.0); ~~however, the stability will decrease significantly as with increasing AOD grows, especially when AOD > 1.0~~. In addition, ~~the~~ ORAC product ~~get matches successfully collocated is consistent~~ with ~~the~~ CARSNET ~~product~~ in winter (December, January and February), whereas other
10 validation results lack matches during winter.

1. Introduction

Aerosols ~~plays~~ a major role ~~on in~~ Earth's climate system, including intervening ~~in the~~ radiation budget and cloud processes, ~~and further to affecting~~ air quality and human health (Remer et al., 2005; Samet et al., 2000; [Tzanis and Varotsos 2008](#); Kokhanovsky and de Leeuw, 2009). The particles suspended in the troposphere ~~will~~ scatter solar radiation back to cool the
15 atmosphere or absorb solar radiation, which warms the atmosphere, causing changes in ~~the~~ net effect of aerosols. These particles ~~also could also~~ affect ~~the~~ formation and microphysical properties of clouds as cloud condensation nuclei (Andreae and Rosenfeld, 2008). The source of aerosols could be anthropogenic or natural ([Varotsos et al. 2012](#)). Particles ~~ferom~~ different sources are mixed into aerosol masses to influence AOD, reduce visibility (Kinne et al., 2003; Remer et al., 2005) and cause spatial and temporal variability of AOD; ~~therefore, the largest uncertainties in the estimation of radiative forcing are introduced~~
20 by aerosols (IPCC, 2013).

Over the past 35 years, different types of satellites have been used to obtain atmospheric information, especially ~~on~~ aerosol properties ~~with development of techniques and science~~ (Griggs, 1979; Kokhanovsky and de Leeuw, 2009). Remote sensing provides a means to obtain global and long-term observations of aerosols, especially in the widest ocean and remote regions where ground-based stations ~~cannot~~ be constructed. ~~Besides~~ ~~In addition~~, polar-orbiting satellites and geostationary satellites
25 obtain daily ~~globale~~ images, which helps to capture changes ~~of in~~ aerosol patterns and properties (Prins et al., 1998; Torres et al., 2002). There are, however, many difficulties in ~~observation of observing~~ aerosols by satellites, because ~~contribution~~ depending on ~~the~~ surface properties, ~~the contribution~~ to ~~the~~ signal received by ~~the~~ satellite ~~could be varying strikingly can vary drastically~~; aerosol components and concentrations are ~~varying~~ constantly ~~in situations varying~~, and their sources ~~ean't cannot~~ be ~~precisely~~ determined ~~exactly~~ (Levy et al., 2007).

30 The Advanced Along-Track Scanning Radiometer (AATSR) aboard ~~on~~ ENVISAT is used to observe the Earth ~~by in~~ dual-view. The data from AATSR can be used to retrieve AOD both over land and ocean, which is ~~an important merit in the important~~

for the characterization of aerosol properties (Adhikary et al., 2008). In recent years, ~~it has established~~ some aerosol retrieval algorithms ~~have been established~~ both over land and ocean, taking advantage of ~~the features~~ of dual-view, which can help eliminate ~~the~~ contribution of surface to top of atmosphere (TOA) reflectance. Aerosol_CCI, as part of ~~the~~ Climate Change Initiative (CCI) (<http://www.esa-aerosol-cci.org/>), provides users ~~with~~ three algorithms for AATSR data, including the
5 Swansea algorithm (SU) (Bevan et al. 2012), the ATSR-2/AATSR dual-view aerosol retrieval algorithm (ADV) (Kolmonen et al. 2015) and the Oxford-RAL Retrieval of Aerosol and Cloud algorithm (ORAC) (Thomas et al. 2009). The aim of this work is to evaluate different performances of these algorithms ~~on-in~~ calculating AOD over different regions of China in 2007, 2008 and 2010.

~~A~~ Ground-based sun-photometer has been used to take sun and sky measurements directly (Holben et al., 1998). The Aerosol
10 Robotic NETwork (AERONET) has ~~already~~ constructed hundreds of sites all over the world ~~til-as of~~ 2015. These stations, ~~found-operated~~ by ~~the~~ American National Aeronautics and Space Administration (NASA), are operational worldwide, providing multi-spectral channels validation data for satellite-retrieved data to complete ~~synthetical~~ measurements on a global scale.

The China Aerosol Remote Sensing Network (CARSNET) is a ground-based aerosol monitoring system ~~using-that uses~~ CE-
15 318 sun-photometers, ~~same-assimilar to~~ AERONET, and has constructed 37 sites throughout China (Che et al., 2009). It has been validated that CARSNET AOD measurements ~~has-same-accuracy-as the AERONET/PHOTONS~~ are about ~~approximately~~ 0.03, 0.01, 0.01 and 0.01 larger than measurements of AERONET at ~~the~~ 1020, 870, 670 and ~~440nm-440 nm~~ channels, respectively (Che et al., 2009). In this paper, we combine two aerosol observation datasets from AERONET and CARSNET as reference data to validate these three AATSR AOD products over China more comprehensively.

20 The basic method for assessment is to compare ~~the~~ retrieval results with data (AOD mainly) ~~given-obtained~~ by AERONET /CARSNET. ~~However,~~ ~~t~~his direct comparison of retrieval results with AERONET data ~~exists-is limitation-limited~~ due to different cloud removal processes (de Leeuw et al., 2013), ~~and~~ such ~~a~~ limitation could influence ~~the~~ validation reliability to some extent. To make ~~the~~ validation more reliable, comparison of ~~the validated~~ retrieval results with high quality data from MODIS or MISR is also one effective method for validation (Kahn et al., 2009). However, AERONET or other ground-based
25 networks provides accurate measurements without ~~the~~ influence of land surface reflection (Holben et al., 1998), which means ~~that~~ comparison of retrieved AOD with ground-based measurements is the basic method. The AATSR L2 products provided by Aerosol_CCI ~~have beens~~ validated by the validation team via a round robin (RR) test (de Leeuw et al., 2013). ~~On~~ this basis, we focused on assessing ~~the~~ performances of AATSR aerosol L2 products in mainland China; ~~using-the-way-of-by~~ ~~comparision-comparing-of-the~~ retrieval results with AERONET ~~and~~ CARSNET data.

2. Reference data and validation statistics

AOD is the most important parameter in terms of aerosol properties and is different from other retrieved parameters under the project of Aerosol_CCI. The Aerosol_CCI project adopts three aerosol retrieval algorithms for ATSR-2/AATSR instrument, including Swansea algorithm (SU) (Bevan et al. 2012), the ATSR-2/AATSR dual view aerosol retrieval algorithm (ADV) (Kolmonen et al. 2015) and the Oxford-RAL Retrieval of Aerosol and Cloud algorithm (ORAC) (Thomas et al. 2009b). All of these three algorithms have ability in retrieval of aerosol properties both over land and ocean. ADV algorithm was originally developed for retrieving AOD properties over land at wavelength of 0.555, 0.659 and 1.61 μm (Veefkind et al. 1998). The main advantage of ADV is the introduction of k-ratio approach to eliminate contribution of reflection to TOA reflectance, which uses the ratio of the reflectance measured in the forward and nadir views (Flowerdew and Haigh, 1995). The ORAC algorithm is designed to retrieve AOD properties at each of four AATSR short-waves channels both over land and ocean, including AOD, effective radius and surface reflectance. The build of the forward model used in ORAC algorithm is based on radiative transfer code - DISORT. A parameterized model of surface reflectance distribution is used in retrieval and combines with the AATSR dual-view to make up shortage of the need of a priori of reflectance (North et al. 1999). An iterative optimization method is employed to determinate AOD, aerosol type and surface reflectance.

AATSR L2 data (see Tab. 1) are daily products with a spatial resolution of $10 \times 10 \text{ km}^2$, and contain a quality flag or a level of confidence for each pixel (de Leeuw et al., 2013). Compared to the Level 3 (L3) product with a spatial resolution of $1^\circ \times 1^\circ$, daily L2 data have higher spatial resolution, that-which helps to capture more-greater details of aerosol properties and more-related to is further explored in our follow-up study.

~~AOD is the most important parameter in characteristic of terms of aerosol properties, and is and is different from other retrieved parameters under the project of Aerosol_CCI.~~ It has been proved-demonstrated that the ground-based observation data from the AERONET have the ability and precision to be used as reference data when users validate AOD (Holben et al., 1998). There are-re 12 AEROENT sites in China providing Level 2.0 (L2) data (cloud-screened and quality-assured) for 2008, 15 sites for 2009, & and 16 sites for 2010, from which the AOD measurement data are available on the website. However, most of these sites are distributed at-in the eastern China coastal area, as shown in Fig. 1, that-can't-which, however, does not be-meet the requirements of comprehensively validating the aerosol properties over whole-all of China-comprehensively. Substantial h-Hazardous dense-aerosol pollutions affects most regions of northern (Li, 2014) and eastern China in winter, and heavy dust aerosols from the Taklimakan desert in western China could-can be transported a-long distances to eastern China, and-even to Japan (Takahashi, 2011), showing-resulting in regional characteristics-differences.

Tab. 1. Details of AATSR AOD products.

<u>algorithm</u>	<u>version</u>	<u>sensor</u>	<u>Main parameters</u>	<u>Resolution coverage</u>
<u>ADV/ASV</u>	<u>2.3</u>	<u>AATSR</u>	<u>AOD,ANG</u>	<u>10 km, 1° global</u>
<u>SU</u>	<u>4.21</u>	<u>AATSR</u>	<u>AOD,ANG</u>	<u>10 km, 1° global</u>
<u>ORAC</u>	<u>03.04</u>	<u>AATSR</u>	<u>AOD, aerosol type</u>	<u>10 km, 1° global</u>
<u>algorithm</u>	<u>version</u>	<u>sensor</u>	<u>Main parameters</u>	<u>Resolution coverage</u>
<u>ADV/ASV</u>	<u>2.3</u>	<u>AATSR</u>	<u>AOD,ANG</u>	<u>10km,10 km, 1° global</u>
<u>SU</u>	<u>4.21</u>	<u>AATSR</u>	<u>AOD,ANG</u>	<u>10km,10 km, 1° global</u>
<u>ORAC</u>	<u>03.04</u>	<u>AATSR</u>	<u>AOD, aerosol type</u>	<u>10km,10 km, 1° global</u>
<u>algorithm</u>	<u>version</u>	<u>sensor</u>	<u>Main parameters</u>	<u>Resolution coverage</u>

AERONET&CARSNET sites in China



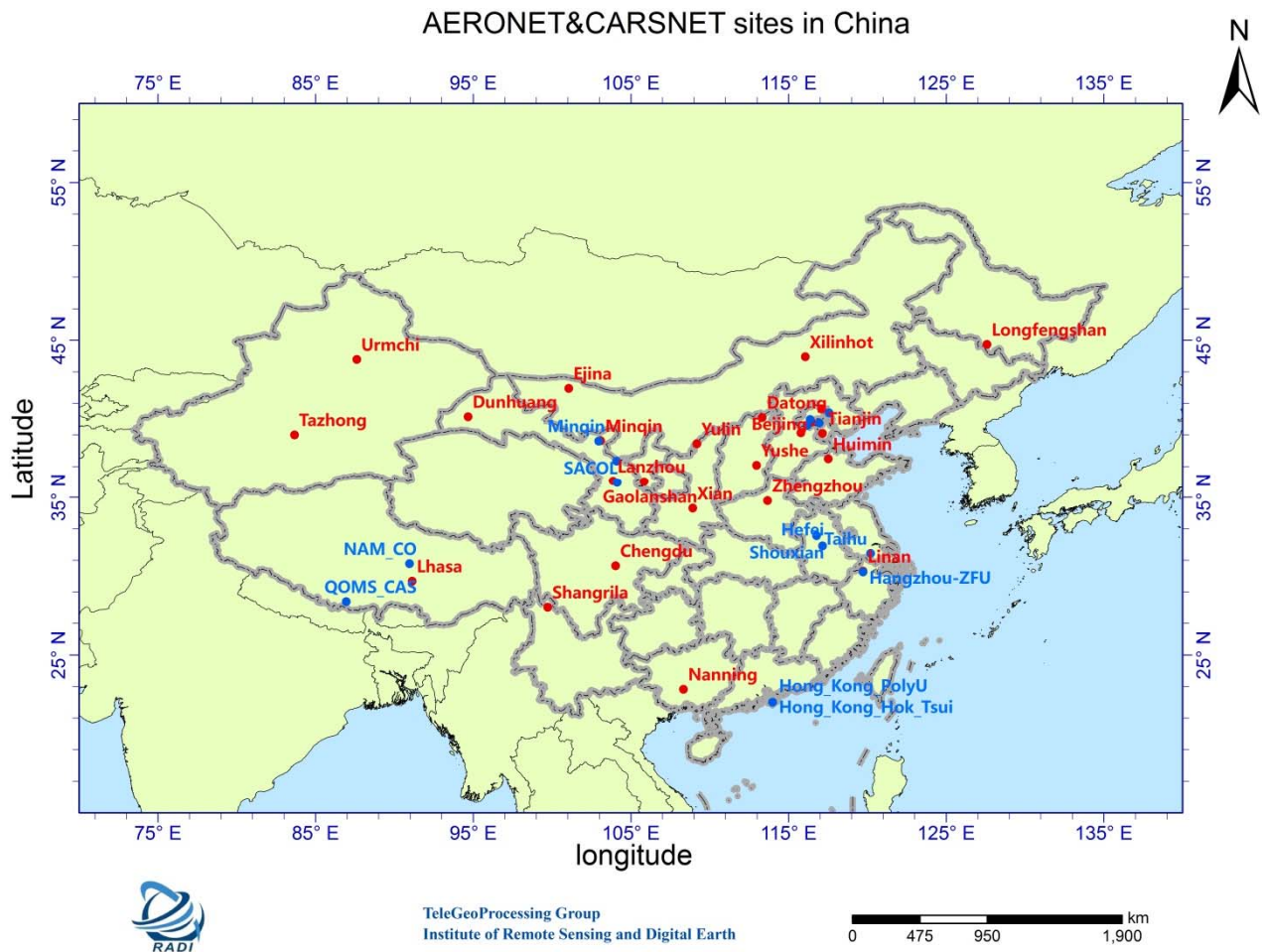


Fig. 1. The distribution of selected AERONET&CARSNET sites in mainland China in 2007, 2008 and 2010. The blue and red points represent AERONET and CARNET sites, respectively.

Fig. 1. The distribution of selected AERONET&CARSNET sites in mainland China in 2007, 2008 and 2010.

- 5 The measurements from another network, the CARSNET, equipped with calibrated CE-318 instruments, have the same accuracy as AERONET, equipping same calibrated CE 318 instruments. The CARSNET has more sites than the the AERONET-s in mainland China, and the spatial distribution of the CARSNET sites are-is distributed more evenly. Therefore, for the purpose of assessing different performances of these three AATSR L2 AOD products, we selected ground-based measurements from both of these two networks as reference data.

The AERONET provides AOD data at three data quality levels: Level 1.0 (unscreened), Level 1.5 (cloud-screened), and Level 2.0 (L2) (cloud-screened and quality-assured) (http://aeronet.gsfc.nasa.gov/new_web/index.html). Here, we selected AERONET L2 data which are screened and quality-assured. Because both of the AERONET and CARSNET data haven't are AATSR products without band-effective wavelengths, we interpolated the ground-based data to the 550nm-550-nm wavelength. Then AOD of the L2 data-sets were compared with AERONET&CARSNET observation data using scatter plots and linear-regression to of the data. The comparisons were made for collocating-located satellite and ground-based observations (Ichoku et al., 2002), i.e., AOD pixels were selected within a spatial extent of +/-50 km with-of ground-based stations in-the-middle and a time range of +/-30 min of the AATSR overpass from the ground-based measurements. At least 5 AATSR AOD retrievals and 2 AERONET/CARSNET observations are required in each collocation (Levy et al., 2010).

10 We made-conducted collocations according to years (2007 2008 and 2010) and datasets (ADV ORAC and SU). Totally-In total, twenty-20 ground-based observation sites, including 12 AERONET sites and 8 CARSNET sites, were on-in the Chinese territory in 2007, of which 6 AERONET and 8 CARSNET inland sites were selected. For 2008, we selected 8 AERONET and 24 CARSNET inland sites, for a total of 32 sites, ignoring the island sites and those near the shoreline. For 2010, only 6 CARSNET sites are available for us, and a total of 14 inland sites were selected with 8 AERONET inland sites (see Table 2).

15 **Table 2. Selected ground-based sites in China.**

	Network	inland	near shoreline	island	Total
2007	AERONET	6	6	0	12
	CARSNET	8	0	0	8
	Total	14	6	0	20
2008	AERONET	8	7	0	15
	CARSNET	24	1	0	25
	Total	32	8	0	40
2010	AERONET	8	7	1	16
	CARSNET	6	0	0	6
	Total	14	7	1	20

2.1 Statistics Metrics

Collocated pairs are analyzed using statistical methods. For the consistency of the metrics among different aerosol products, strong matches are determined using the expected error (EE) which shows the percent of retrievals falling within the expected error (EE) range. Then, good strong matches are determined using the expected error (EE). An EE envelope was put forward introduced for retrieval of MODIS AOD (Kaufman et al., 1997; Chu et al., 2002) by means of sensitivity studies, as demonstrated by Eq. (1) and Eq. (2):

$$EE1 = \pm(0.05 + 0.15\tau) \quad (1)$$

$$EE2 = \pm(0.05 + 0.20\tau) \quad (2)$$

where, τ represents the satellite-retrieved AOD. AATSR AOD retrievals are different from the MODIS AOD datasets. In this paper, we introduced the EE envelope according to based on the features of AOD underestimation and formation of the MODIS EE envelope, as demonstrated by Eq. (3) and Eq. (4):

$$EE3 = \pm(0.05 + f2 + (f1 + 0.15)\tau) \quad (3)$$

$$EE4 = \pm(0.05 + f2 + (f1 + 0.20)\tau) \quad (4)$$

where, $f1$ is the slope of the regression line of scatter points and $f2$ is the correspondent intercept. In the process of retrieving AOD, underestimation tends to be caused by systematic error. Therefore, the EE envelopes suggested by Kaufman et al. or Chu et al. are not fit for validation of the AATSR AOD. In such design of EE design, i.e., as with Eq. (3) and Eq. (4), was to take consideration of the underestimation was taken into consideration by, regarding the regression line as the center, not the 1-1 line, for determining the accidental error.

Linear regression is the most basic and commonly used statistical method that allows us to summarize and study relationships between two quantitative variables. Pearson correlation coefficient (CC) measures the fraction of the total variability in the response that is accounted for by the retrieval and is only a measure of linear association between ground truth measurements and satellite retrieval values. Bias describes the average difference between satellite retrievals and ground AOD. After that Then, to determine how well the satellite data match the ground-based observation data, exploring what the relationship between them is explored. A regression equation and some basic statistics are put on shown on the scatter plot, including the correlation coefficient (CC) and, root mean square error (RMSE):

$$CC = \frac{\sum_{i=1}^n (\tau_{aero,i} - \bar{\tau}_{aero})(\tau_{sat,i} - \bar{\tau}_{sat})}{\sqrt{\sum_{i=1}^n (\tau_{aero,i} - \bar{\tau}_{aero})^2} \sqrt{\sum_{i=1}^n (\tau_{sat,i} - \bar{\tau}_{sat})^2}} \quad (5)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\tau_{sat,i} - \tau_{aero,i})^2} \quad (6)$$

Where, $\tau_{aero,i}$ represents the ground-based observation data and, τ_{sat} represents the satellite retrievals.

Mean satellite-retrieved AOD (MSA) and mean AERONET&CARSNET AOD (MAA) represent the central tendency of the

data. Relative mean bias (RMB) is used to determine under- or overestimation of the AOD retrievals; it is the ratio of MSA to MAA as Eq. (5):

$$RMB = MSA/MAA \quad (7)$$

Mean bias error (MBE) is the mean value of difference between the satellite retrievals and AATSR AODs, and the mean absolute error (MAE) is the absolute value of the mean value of bias error. Together with RMB, the MBE and MAE are used to determine the magnitude of the difference between the two datasets.

2.2 KAPPA Statistics

In the scatter plot of the collocated pairs, the retrieved data and the corresponding collocated ground-based observation data could be considered as two arrays, and the main purpose of KAPPA is to explore how these two arrays match each other. For retrieval of aerosol properties, the performances of most algorithms will turn down decrease in effectiveness with increase of increasing AOD, i.e., difficulties in retrieving AOD will be increased as AOD grows increases. Obviously, when only using $|bias|$, the absolute value of the difference between ground-based data and AATSR AOD data in each collocation pair, as an assessment standard for different AODs, is insufficient and lack of persuasion. Therefore, the combination of $|bias|$ and $|bias|/Ground$, i.e., the ratio of $-|bias|$ to the value of the reference data in each collocation pair, used in the KAPPA coefficient will make up account for this shortage and provides a new statistic for assessing the agreement between two arrays, taking advantage of the KAPPA coefficient.

The KAPPA coefficient was originally proposed as a descriptive statistic indicating the degree of beyond-chance agreement between two ratings per subject in a dichotomous form (Bloch and Kraemer, 1989). KAPPA coefficients with various forms also could be used to measure the accuracy for of thematic classifications (Rosenfield and Fitzpatrick-Lins, 1986). KAPPA is, in short, a measure of “true” agreement (Cohen, 1960). The pairs collocated by matching ground-based data with AATSR L2 AOD data could be regarded as two different arrays so that we introduced the KAPPA coefficient to assess agreement between these two arrays. According to Based on the concept about of the KAPPA coefficient proposed proposed by Cohen (1960), an appropriate modification with a two-category nominal scale has been made is shown in Table 3.

Table 3. Design of the KAPPA coefficient.

		Criterion 2		Total
		Relevant (highly)	Relevant (low)	
Criterion 1	Relevant (highly)	a	b	G1
	Relevant (low)	c	d	G2
Total		F1	F2	n

To estimate the KAPPA coefficient, ~~it one~~ needs to determine ~~which which pairs are~~ “true” or ~~which which pairs are~~ “relevant”. However, ~~if~~ only given matched collocation pairs, we cannot²⁴ determine which pair is relevant ~~or which retrieved AOD in the collocation pair is on behalf of high quality~~. Therefore, the design of ~~criteria~~ 1 and ~~criteria~~ 2 needs to be reasonable and fit for the purpose of validation.

- 5 For ~~criteria~~ 1, if $|bias|$ is greater than the mean of $|bias|$, then ~~it is~~ marked as “lowly relevant”, and if not, ~~it is~~ marked as “highly relevant”. Here, the bias was ~~assessed~~ from ~~the~~ first quartile to ~~the~~ third quartile for eliminating possible “outliers”. The $|bias|_{only}$ indicates the absolute error of ~~the~~ retrieved AOD, and it still needs another statistic for ~~criteria~~ 2, ~~the i.e.~~ $|bias|/Ground$, ~~which~~ indicates the relative error of ~~retrieval~~-AOD ~~retrieval~~. For ~~criteria~~ 2, if $|bias|/Ground|$ is greater than 0.2 (according to EE4), then ~~it is~~ marked as “lowly relevant”, and if not, ~~it is~~ marked as “highly relevant”. For the
- 10 conventional formula of calculating the KAPPA coefficient:

$$K = \frac{P_0 - P_c}{1 - P_c} \quad (8)$$

~~where~~ P_0 is the proportion of observed agreements and P is the proportion of chance agreements.

$$P_0 = \frac{(a + d)}{n} \quad (9)$$

$$P_c = \frac{\left(\frac{F_1 \times G_1}{n}\right) + \left(\frac{F_2 \times G_2}{n}\right)}{n} \quad (10)$$

- Algorithms for AATSR AOD retrieval used to underestimate AOD over different regions in China ~~include including the~~ ADV ORAC and SU algorithms, ~~otherwise, on~~ this basis, ~~it is good~~ the agreement between ground-based observation data and satellite retrievals ~~for is assessed based on the~~ ADV and SU algorithms (Che et al., 2015). The main aim of this new KAPPA
- 15 coefficient is to evaluate ~~the~~ comprehensive performance of ~~these~~ algorithms, ~~its~~ function is to represent not only ~~the~~ degree of underestimation, ~~but also but also~~ the level of agreement between different datasets.

3. Validation results and analysis

- We ~~have~~ collected different validation reference data of AERONET and CASNET in 2007, 2008 and 2010. Only 14 ground-based observation sites are available ~~for us~~ in 2007, of which some are ~~located~~ close to each other. Most ~~of them~~ are located
- 20 in different provinces, ~~but however, the total numbers of sites are is~~ small and the space distribution is not uniform. Therefore, ~~the numbers of matches are is~~ relatively small for ~~all the all of the~~ algorithms. More AERONET/CARSNET data are available in 2008, with ~~a total of~~ 32 sites including 8 AERONET sites and 24 CARSNET sites. There ~~are re~~ 14 AERO&CARS sites ~~giving providing~~ data for validation in 2010.

The focus of this paper is to ~~find-determine the~~ differences between ~~the~~ ADV, ORAC, and SU L2 AOD products. Therefore, we calculated statistics and analysed ~~the~~ validation results separately by year (see Tab. 4).

5 **Table 4. Main statistics of ~~the~~ validation results.**

		N	MSA	MAA	MBE	MAE	RMSE	RMB	CC	KAPPA	Within EE3	Within EE4
AATSR ADV	2007	94	0.244	0.347	-0.103	0.115	0.095	0.704	0.885	0.473	74.5%	77.7%
	2008	307	0.211	0.361	-0.151	0.157	0.124	0.583	0.790	0.329	70.0%	77.5%
	2010	136	0.150	0.293	-0.143	0.148	0.089	0.512	0.785	0.180	91.2%	93.4%
AATSR ORAC	2007	137	0.324	0.270	0.054	0.137	0.206	1.200	0.708	0.474	44.5%	50.4%
	2008	612	0.271	0.330	-0.060	0.160	0.209	0.819	0.472	0.439	40.4%	47.4%
	2010	282	0.254	0.274	-0.020	0.141	0.170	0.925	0.665	0.367	42.2%	46.8%
AATSR SU	2007	94	0.330	0.404	-0.074	0.092	0.124	0.816	0.933	0.409	77.7%	87.2%
	2008	435	0.293	0.412	-0.118	0.137	0.140	0.713	0.822	0.484	71.5%	80.2%
	2010	167	0.270	0.375	-0.105	0.119	0.131	0.720	0.888	0.520	77.3%	84.4%

3.1 The ADV algorithm

For 2007, ~~the~~ RMS error is 0.095, ~~minimal-the lowest~~ of all results, the CC is 0.885, and the distribution of collocated pairs in ~~the~~ scatter plot ~~are-is~~ concentrated near the regression, as shown in Fig. 2a. ~~Most of the~~ collocated pairs ~~are~~ within EE3 (~~about-approximately~~ 74.5%), indicating that ~~the~~ satellite retrievals ~~agree-are consistent~~ with AERONET/CARSNET data ~~well~~.

10 The RMB is 0.704, and the regression line is $y = 0.77x - 0.02$, which reflects ~~one-the~~ tendency of underestimation. This ~~kind~~ ~~type~~ of underestimation ~~will-be~~ more severe ~~as-growth-of~~ with increasing AOD ~~value~~. Low dispersion and slight underestimation make ~~the~~ KAPPA coefficient high (0.473), ~~showing-demonstrating~~ that ~~the~~ ADV algorithm ~~has-good~~ ~~performance-performs well~~ in calculating ~~the~~ AOD over China in 2007. ~~The~~ ADV algorithm is appropriate for ~~the~~ retrieval of low AODs, especially for those less than 1.0; ~~so-thus~~, the MSA for 2007 is 0.244.

15 For 2008, ~~the~~ lower RMB (0.621) ~~means-suggests~~ more severe underestimation, and ~~the~~ lower CC (0.776) and higher RSE (0.130) ~~mean-indicate~~ lower accuracy. Similar with 2007, the MSA of ~~the~~ ADV is 0.211. Therefore, ~~the~~ KAAPA coefficient, which ~~is-on-behalf~~ ~~measures the~~ overall performance, is 0.329, lower than ~~result-that~~ of 2007. For 2010, ~~the~~ lowest RMS (0.089) and largest proportion of matches ~~are~~ located in EE3 (91.2%) and EE4 (93.4%), ~~respectively~~, ~~with the lowest in-three-years~~ ~~mean-small~~ accidental error ~~of the three years~~. However, the KAPPA coefficient is 0.180, ~~also~~ the lowest ~~in-of the~~ three years.

The most obvious feature of the ADV algorithm is underestimation, as shown demonstrated in Fig. 2. The mean $\pm 2\sigma$ lines in the different ranges are almost within the EE4 lines for these three years. The highest MSA is 0.250 in 2007, and the lowest MSA is 0.173 in 2010 in three years. The ADV algorithm has ability in can retrieving retrieve low AOD values well with high accuracy. Actually this This “ability” is systematic for either high AODs or low AODs. This also limits the range of application of the ADV algorithm, especially in calculating AODs in range of high value rangess.

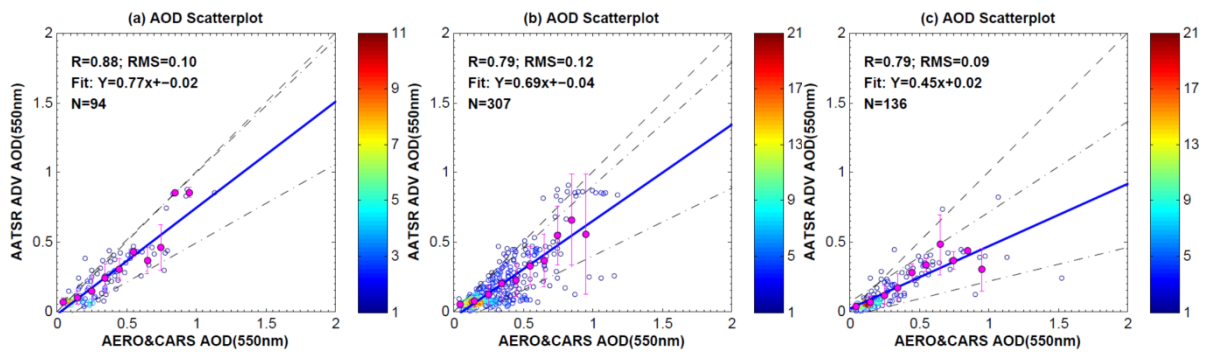


Fig. 2. Scatter plots of AATSR ADV L2 AOD products with ground-based data in China in 2007, 2008 and 2010. The dashed, dotted and blue solid lines represent the 1-1 line, EE4 line and regression line, respectively. The magenta points are means for specific ranges of AERO&CARS AOD, and the magenta lines are the mean $\pm 2\sigma$ for a certain range.

3.2 The ORAC algorithm

The ORAC algorithm had good performance-performed well for in 2007, getting-achieving a KAPPA coefficient of 0.474. However, the distribution of those matches is dispersed in Fig. 3b, implying low CC (0.708) and high RMSE (0.206). From the angle of In terms of the degree of fitness, its performance is not good-effective. However, there is no obvious trend of underestimation or overestimation, and the regression line is close to the 1-1 line. Only 50.4% of collocated pairs are within EE4, and most of the mean $\pm 2\sigma$ lines are out of the EE4 lines, showing-suggesting that accidental errors influence the accuracy of the ORAC algorithm. The MSA of the ORAC is 0.324.

From the number of matches, ORAC has the most matches of the three algorithms (see Fig. 3). Different from 2008, no there's obvious underestimation occurs from-in the results of 2007 and 2010, as demonstrated by the regression lines shown in Fig. 3b and 3c. For 2008, the RMB is 0.829, showing-suggesting a trend-of-slight underestimation trend. The applicability of ORAC is good-high, with MSA of 0.271. The collocated pairs are relatively dispersed, and almost all mean $\pm 2\sigma$ lines are out of EE4 lines, influencing the RMSE and CC. For 2010, the same dispersion of points in the scattered plot and low KAPPA coefficient are observed.

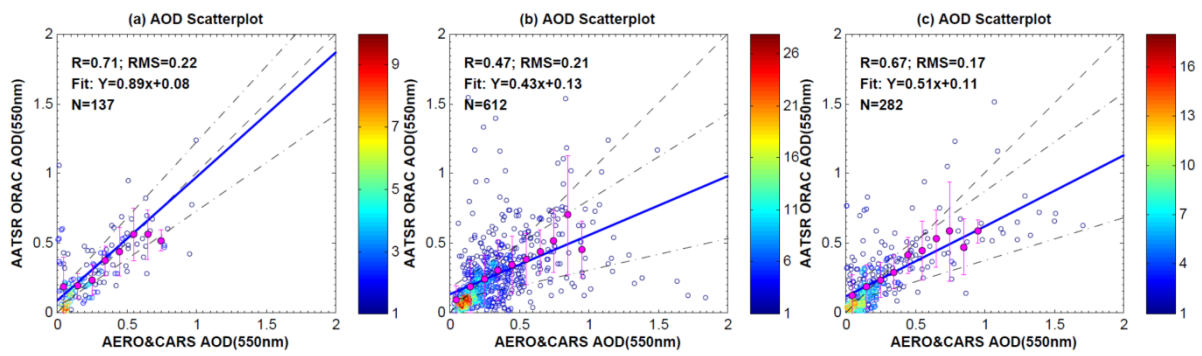


Fig. 3. Scatter plots of AATSR ORAC L2 AOD products with ground-based data in China in 2007, 2008 and 2010.

Overall, the ORAC algorithm tends to retrieve AODs unstably for either high AODs or low AODs and with a slight underestimation in 2007. The results of 2008 and 2010 have share common features in common, even though the regression lines are below the 1-1 lines, influences indicating that of accidental error are is larger than systematic error.

3.3 The SU algorithm

The SU algorithm had good performances performed well for all in three years, getting achieving KAPPA coefficients of 0.409, 0.484 and 0.520, respectively. Large proportions of matches are within EE3 and EE4, and almost all the all of the mean $\pm 2\sigma$ lines are within the EE4 lines, both showing suggesting that the matches are concentrated in small regions around the regression line. The RBMs are 0.816, 0.713 and 0.720 respectively for 2007, 2008 and 2010, respectively, showing demonstrating the underestimation of the SU product. The applicability of -SU is good high with MSA of 0.293 for 2008.

The most obvious feature of the SU algorithm is its stability in retrieving AOD for different years or different regions (Fig. 4). The MSA are ranges from 0.270 for 2010 to 0.330 for 2007, and the KAPPA coefficient is from 0.520 to 0.409, which means suggests that the SU algorithm had performed better performance in retrieving low AODs. The SU algorithm has the best performance in retrieving AOD terms of AOD retrieval, as it has the with highest KAPPA coefficient (0.520).

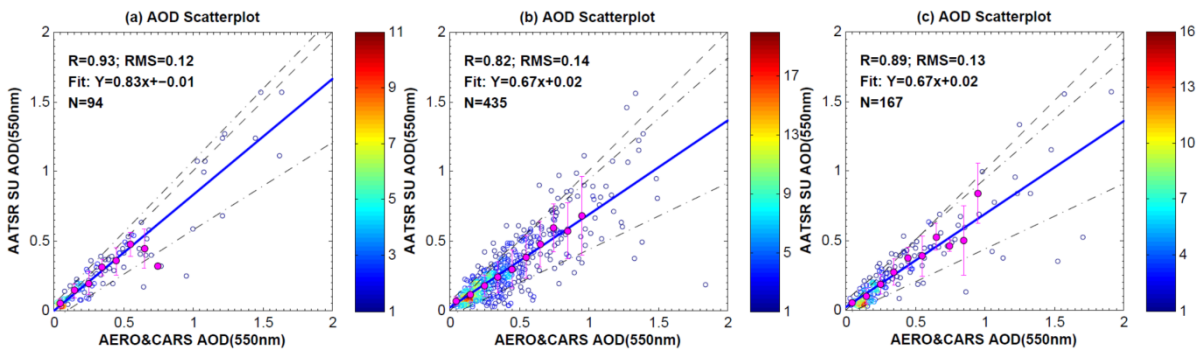


Fig. 4. Scatter plots of AATSR SU L2 AOD products with ground-based data in China in 2007, 2008 and 2010.

Overall, the SU algorithm could be applied to retrieve AOD in different ranges with high precision. Factors influencing the performances of the SU algorithm includes small systematic error and even smaller accidental error.

4. Uncertainty analysis based on aerosol loading

In the previous section, we have validated all three AOD products over mainland China in 2007, 2008 and 2010, discovering that all these three products tend to be exhibiting underestimation to some extent. For the purpose of ascertaining the causes of the underestimation, in this section, we focus on analysing the AOD uncertainties which are the leading to differences between retrieved AODs and ground-based AODs in special conditions. Collocated pairs are divided into three groups according to aerosol loading, including light loading ($\tau < 0.15$), heavy loading ($\tau > 0.4$), and moderate loading (Levy et al., 2010). It is obvious that the AOD bias become greater increases with the growth of increasing AOD for all three products. These products have one feature in common, that is, the AOD bias tends to be negative, which means indicates that the underestimation becomes more significant with the growth of increasing aerosol loading. The ADV and SU algorithms have good performance in estimating AOD even, i.e., with little underestimation, when aerosol loading is low (light loading) (Fig. 5).

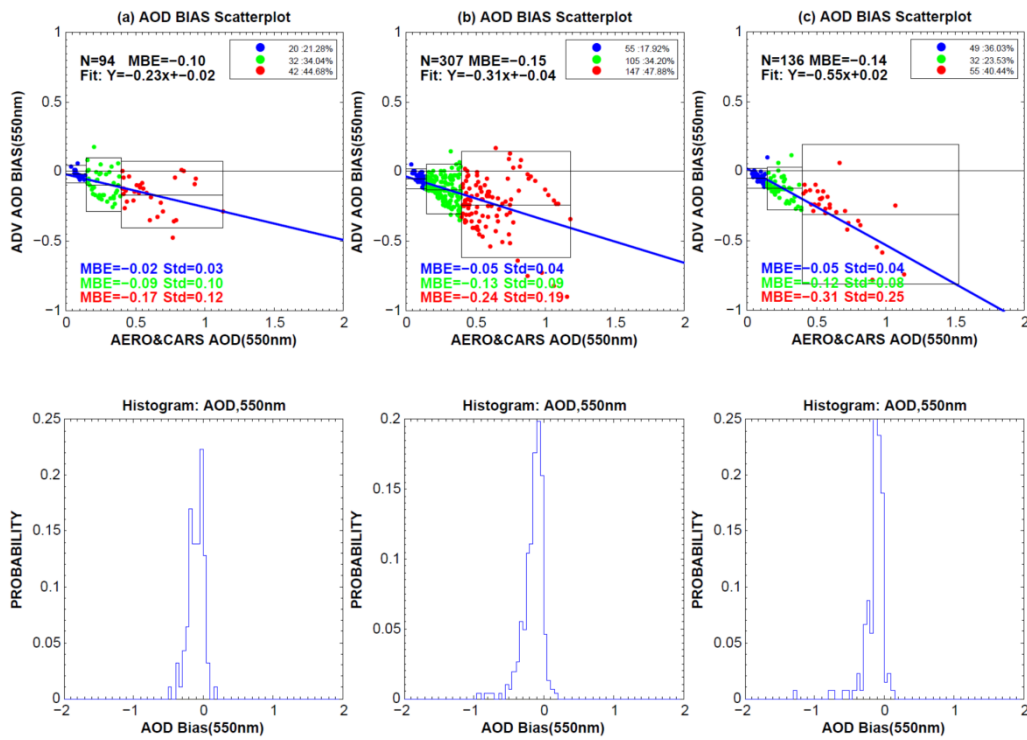
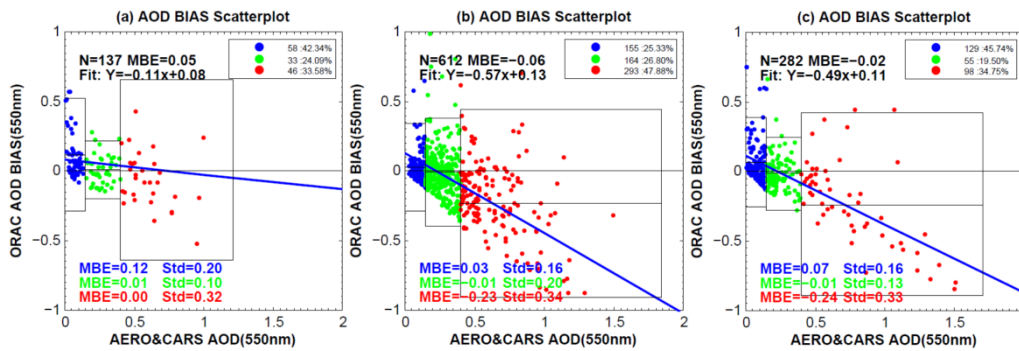


Fig. 5. Scatter plot of AERONET&CARSNET AODs with ADV AOD bias or uncertainties, and histogram of AOD bias. Colours represent different groups; blue for denotes light loading, green for denotes moderate loading, and red for denotes heavy loading. Basic statistics are ~~texted displayed on in~~ the top left corner, including the number of scattered points, MBE and linear regression equation (Fit). The text on the bottom with different colours ~~are describes the~~ basic statistics of each group. Each group has one box, the ~~bottom and top borders bottom and borders top borders~~ of which represent $MBE + 2\sigma$ and $MBE - 2\sigma$, respectively, containing 96% of scattered points from each group. The center line of each box represents the MBE of each group. The blue line is the regression line of all scattered points.

Under complex conditions, the ORAC overestimates AOD in regions of light loading and moderate loading compared with the AEROENT, as shown in Fig. 6a-6b. Compared to the CARSNET data, it also appears overestimation occurs for light loading, and this overestimation is mainly due to two points with large error. In the moderate loading region, the MBE tends to be positive in Fig. 6a, probably because the distribution of AEROENT sites is uneven, that as most of sites are located in eastern China.

The top and bottom borders of the box we draw represent the interval of $[-2\sigma, 2\sigma]$, which contains most of the data (about approximately 95%) for a given group. The data outside the box are “possible outliers” based on the due to largest error contained in each group. Those “possible outliers” have one feature in common in that the corresponding points in the bias scatter plot are far away from other points. Otherwise, the points below or above the box are different. If a points are is above the box, which means indicates that the satellite-retrieved AOD are is larger than the ground-based observed AOD, those this “outliers” tends to be caused by a residual cloud. Because The ground-based network measures AOD just from only one point; but however, the satellite-retrieved AODs in each collocated pairs are an average of 25 pixels. Any one of these 25 pixels with a cloud residual will lead to an increase of AOD in a collocated pair. Therefore, we make a conclusion conclude that the “outliers” above the box are possibly caused by a cloud residual. From this view, there’s one point above the box of corresponds to each aerosol loading respectively for the ADV product. This kind of The “outliers” are concentrated in the up light loading region and moderate loading region for the SU product (Fig. 7). The situation of the ORAC is relatively complex; it exists “outliers” occur in the light loading region, which makes the box of the light loading much larger than box that of the moderate loading region in 2007 and 2010 as shown in (Fig. 6a and 6c).



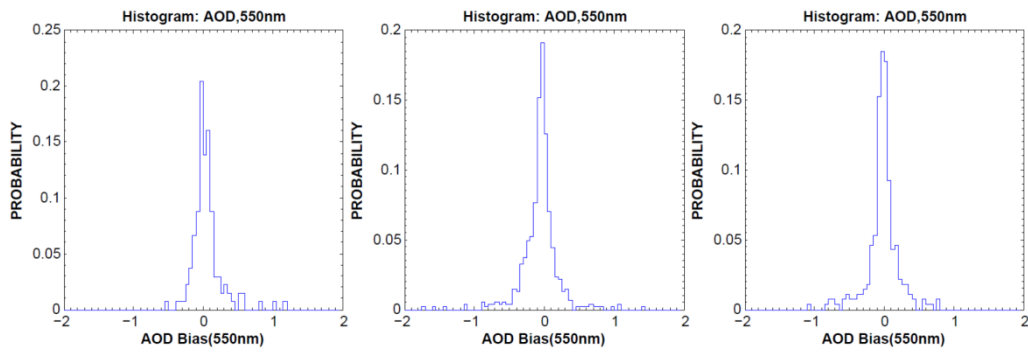


Fig. 6. Scatter plot of AERONET&CARSNET AODs with ORAC AOD bias or uncertainties and histogram of AOD bias.

The points below the box are different from those above the box; most of them are only below the box for due to heavy loading, indicating that the ability of estimating AOD will decrease with the increase of increasing aerosol loading, especially in region of the heavy aerosol loading region.

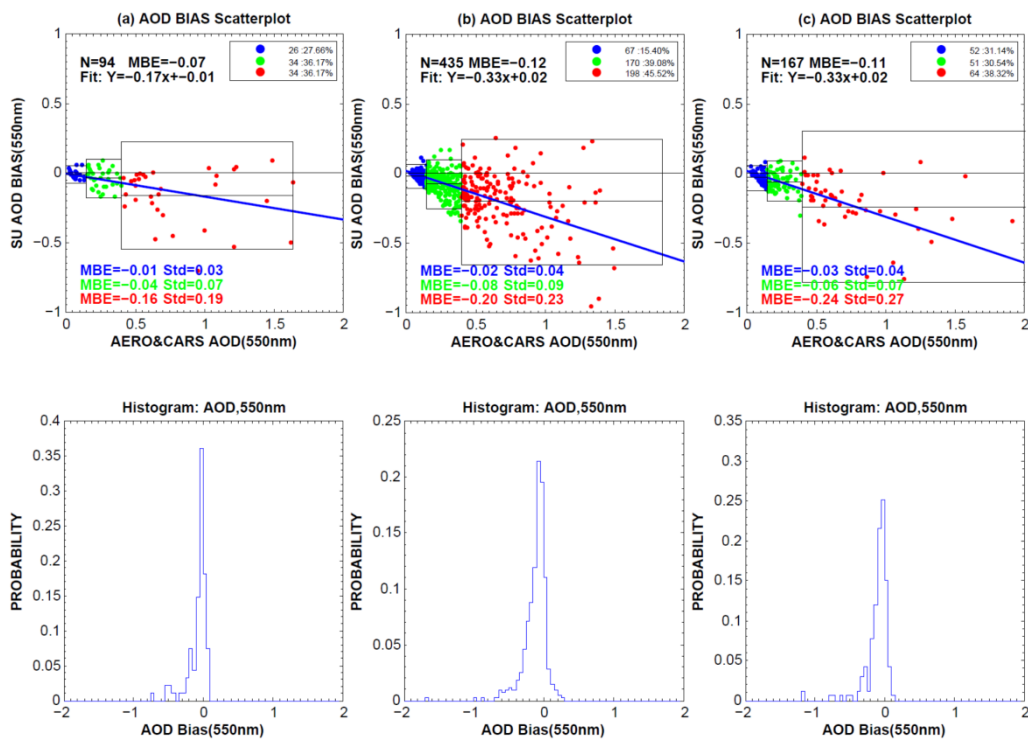


Fig. 7. Scatter plot of AERONET&CARSNET AODs with SU AOD bias or uncertainties and histogram of AOD bias.

We make these groups because aerosols have exhibit different natures-behaviours with different loading conditions. In general, the bias or uncertainty of satellite-retrieved AOD will increase with the increase of increasing AOD or aerosol loading. As

discussed above, all of these algorithms underestimate AOD at different levels; similarly, it is's worth noting that underestimation get-becomes more severe with the increase of/increasing AOD or aerosol loading.

5. Uncertainty analysis of individual ground measurement sites

5 For the purpose of further evaluating the different performances of these three algorithms on-in estimating AOD over mainland China, we validate these products on a site-by-site basis. It is significant to explore what-the roles of different factors in estimating AOD. There are several factors maybe-that may have impacts on AOD calculationing, including land cover, aerosol type, elevation, etc. Therefore, we analyse different validation results of each site to study how these factors work (see Table 5).

5.1 Inter-comparison of algorithms site by site

10 In this section, we pick-select five representative AERONET&CARSNET sites in-whichwith it-allocated more than 30 successful matches -suecessfully in 2007, 2008 and 2010 to guarantee an appropriate statistical sample size. These selected sites are located in different regions where the land cover and climatic pattern are different and of-representative strongly in of mainland China. Two AERONET sites and three CARSNET sites were selected, including SACOL and; XiangHe from AERONET, and Linan, Shangdianzi and Xilinhot from CARSNET. Most matches of ADV and SU products collocated with
15 ground-based data occurred in/distributed at March to October, lost data at winter time in 2007, 2008 and 2010, as shown in Fig. 8 to Fig. 12. The matches of the ORAC product were distributed at-in each month over most sites.

Linan is located at 119.73°E , 30.3°N , 119.73°E , northwest of Zhejiang province. A total of 80% of the $50\text{ km} \times 50\text{ km}$ surrounding area is covered by green vegetation, and the other 20% is covered with urban land. The ADV and ORAC algorithm underestimated AOD, with $\text{MBE} = 0.13$ and 0.12 in 2010, respectively. The SU has good performance performed well in Linan, with slight underestimation. The underestimation of the ADV algorithm is more severe than that of SU and ORAC. Even though. Although the ORAC algorithm has the most matches in Linan, its performance was unstable, which means that the level of underestimation were was different in different years.

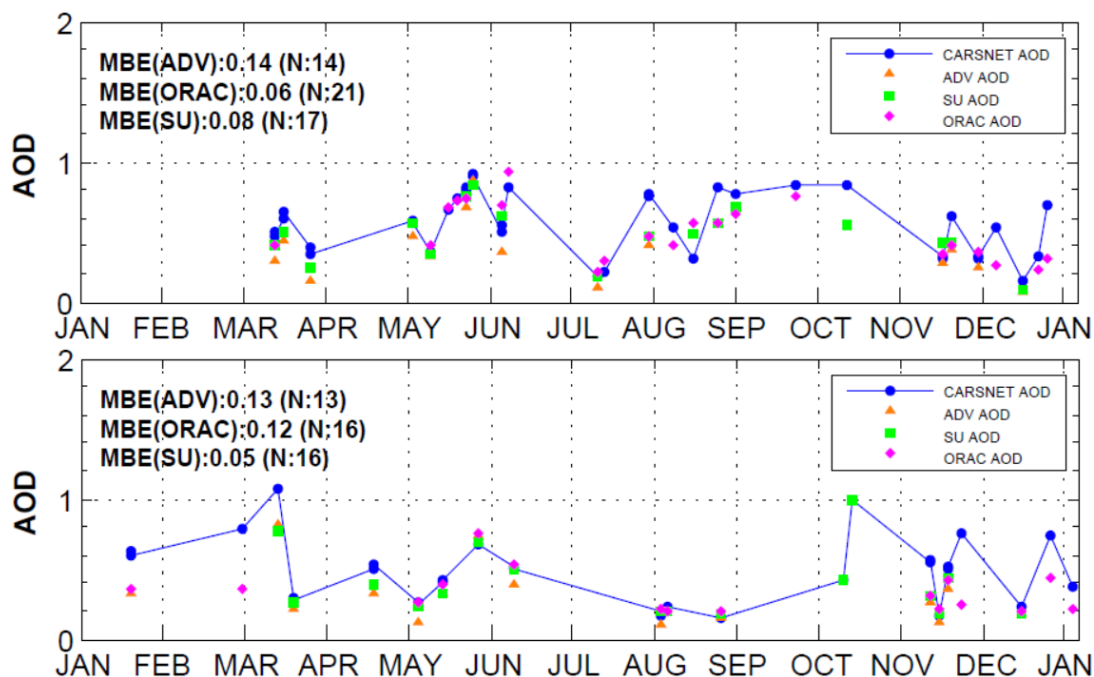
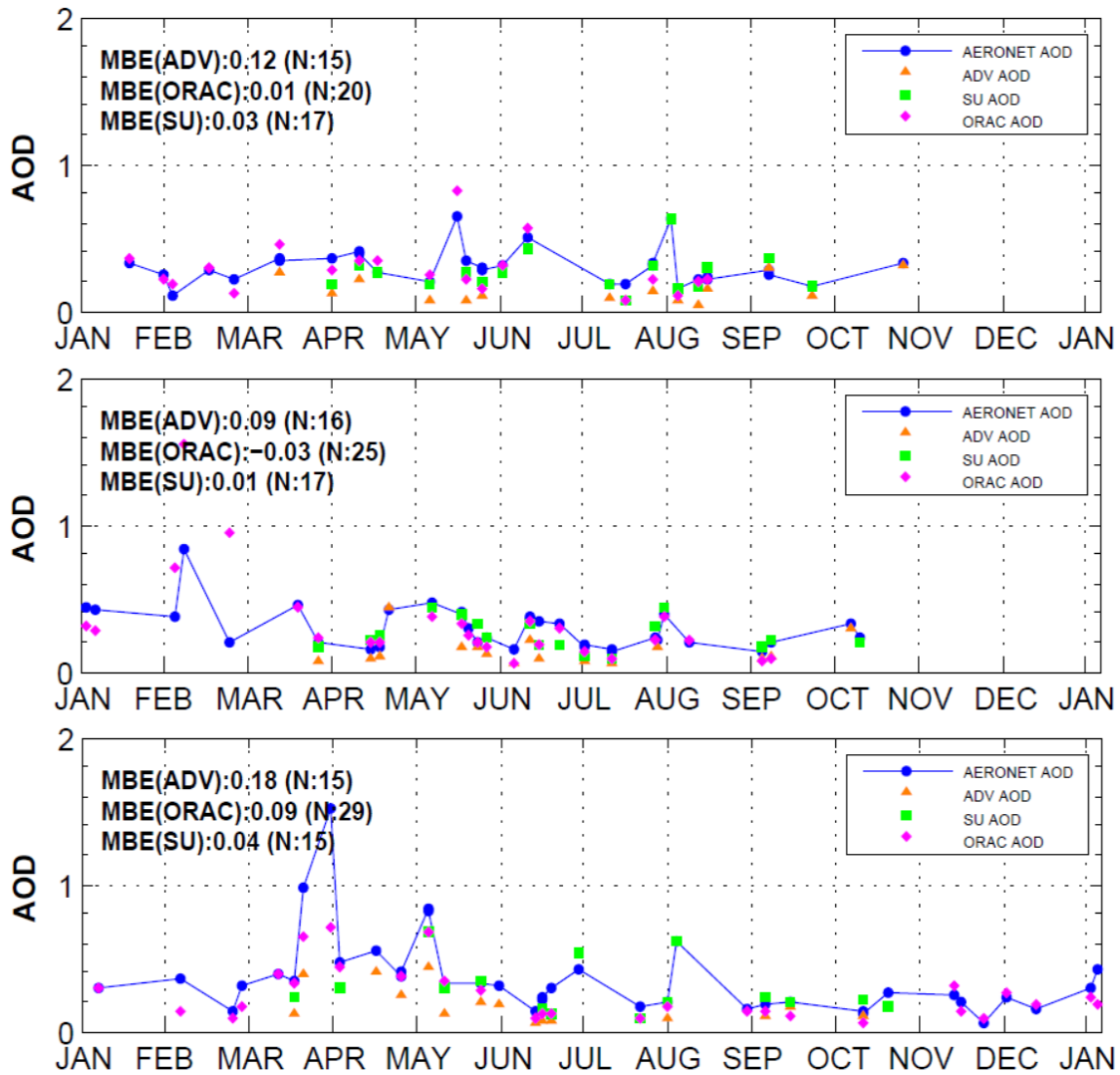


Fig. 8. is the time series comparison of AATSr AOD with CARSNET AOD over the site of Linan in 2008 and 2010.

SACOL is situated at along the southern bank of the Yellow River in Lanzhou city, Gansu province. Lanzhou city is has a temperate continental climate, having with four clearly distinctive seasons. The dominated dominant land cover is grassland, coverings about approximately 95% at of the spatial extent of the $50\text{ km} \times 50\text{ km}$ area from the MODIS MCD12C1 land cover data. A total of 30% of the surface is arid and semi-arid areas, which can be a source of dust aerosols. SU has good performance on performs well in retrieving AOD over SACOL, with a high CC (0.849) and low RMSE (0.072). The a Accidental error in the retrievals of using the ORAC algorithm is obvious, leading to a low CC (0.683) and high RMSE (0.170). Most of the retrievals (91.3%) of the ADV algorithm are within EE4. However, as discussed above, the ADV algorithm severely underestimated AOD severely in SACOL. The ADV algorithm tended to severely underestimate the AOD of different ranges, severely except for a small number of matches with high quality matches. The matches of the SU product are of high

quality in for the three years. The ORAC has collocated matches in January, February, November and December (winter time), when there're no matches of unlike the ADV and SU products. However, the accuracies accuracy of ORAC in winter time are is of great highly uncertainty, as shown in Fig. 9.



5 Fig. 9. is the time Time series comparison of AATSR AOD with AERONET AOD over the site of at SACOL in 2007, 2008 and 2010.

Shangdianzi is situated at 94.68°E , 40.15°N , 94.68°E , with complex land cover about of approximately 45% of croplands, 30% of mixed forest, 18% of closed shrublands, 5% of grasslands, 1% of water and 1% of evergreen needleleaf forest. The SU algorithm has high precision of AOD calculation over this site from March to October, when most of the land cover is green. The ADV algorithm has good performances on also performs well in calculating AOD over these three sites, with slight underestimation. The performance of the ORAC algorithm in Shangdianzi is unstable, good with strong agreement with ground-based data from March to October and severe underestimation in winter time, as shown in Fig. 10.

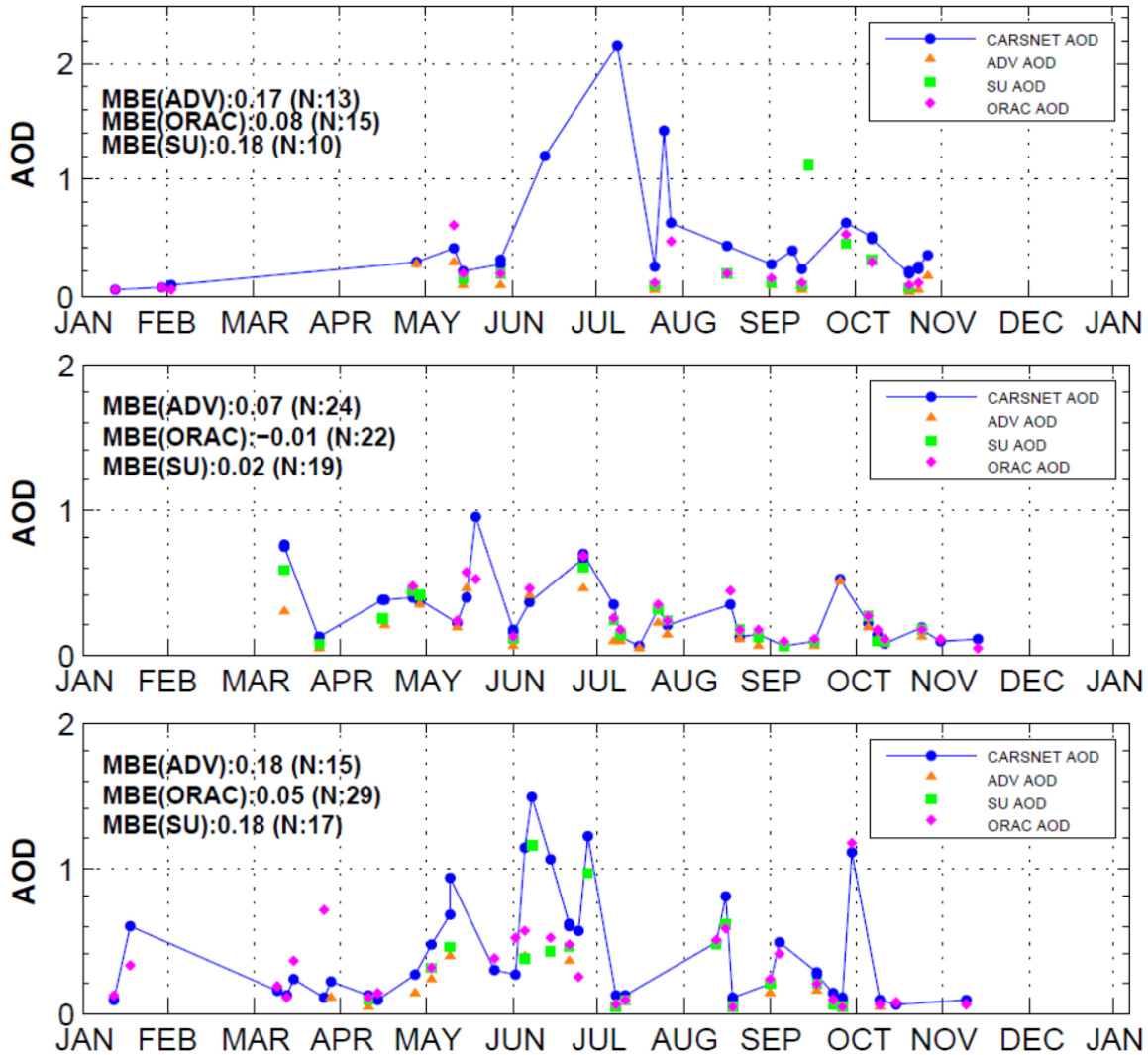


Fig. 10. is the time series comparison of AATSR AOD with CARSNET AOD over the site of Shangdianzi in 2007, 2008 and 2010.

Xianghe is located ~~at to the~~ southeast of Beijing, ~~having and has the same climatic pattern conditions~~ as Beijing. ~~About~~ ~~Approximately~~ 98% of ~~the~~ surface is covered with urban land ~~from according to the~~ MCD12C1 data ~~at extent of a~~ 50 km × ~~50km-50 km~~ area. The performances of these three algorithms are at ~~the same high quality~~ level ~~with high quality~~. ~~However,~~ ~~the~~ ADV algorithm still underestimated AOD at ~~a~~ level of MBE = 0.12 in 2007 and 0.10 in 2008.

5

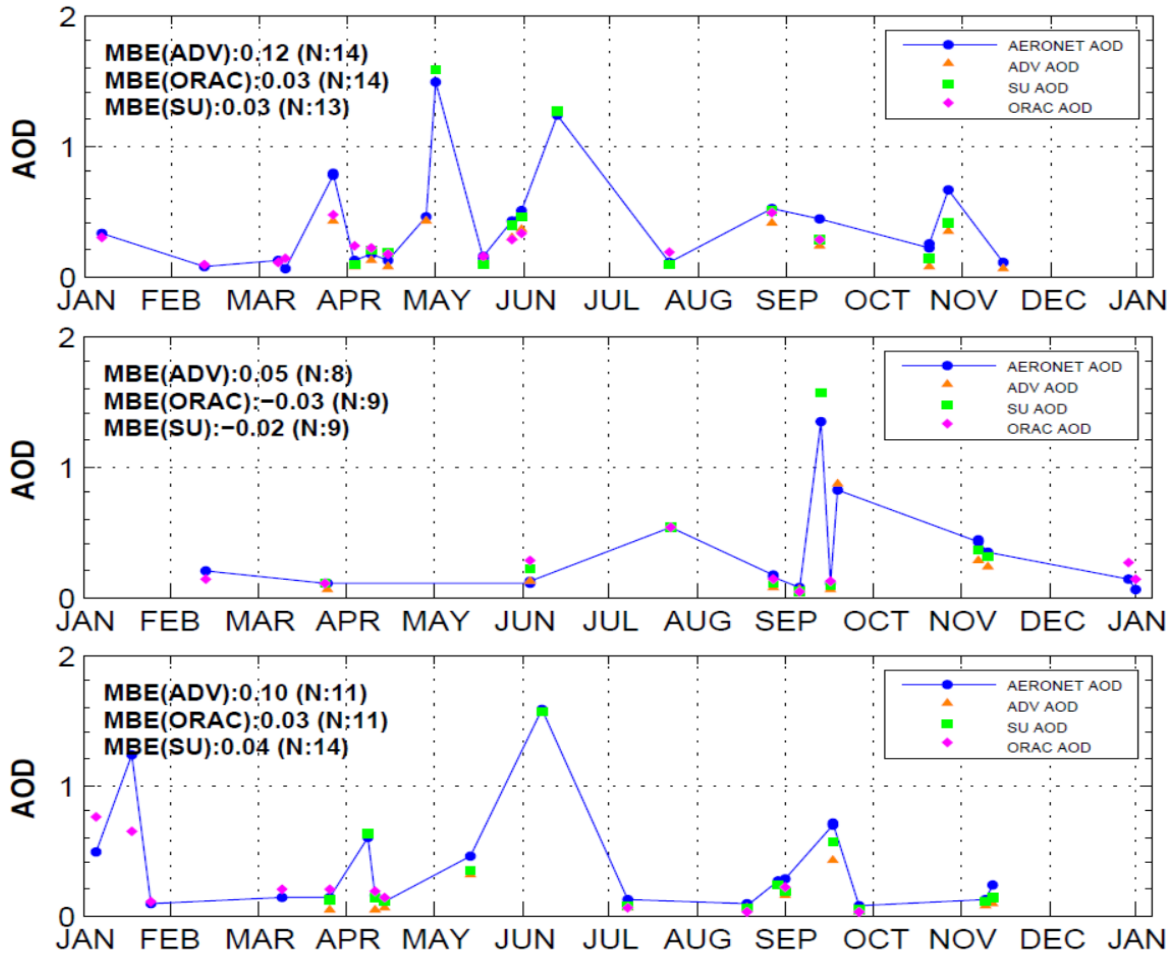


Fig. 11. ~~is the time~~Time series comparison of AATSR AOD with AERONET AOD ~~over the site of at~~ XiangHe in 2007, 2008 and 2010.

Xilinhot is situated at ~~116.07°E, 43.95°N_116.07°E_~~ at the centre of ~~the~~ Xilinguole grassland. The main land cover is grassland (100%) ~~from based on the~~ MODIS MCD12C1 data, ~~at with a~~ spatial extent of 50 km × ~~50km-50 km~~. The surface ~~circumstance~~ and climate features of Xilinhot are ~~much likes~~ similar to those of SACOL's, and the performances of the SU

algorithm ~~on~~ at these two sites ~~are~~ the same, i.e., both with high R and low RMSE ~~both~~. The ADV algorithm ~~slightly~~ underestimated AOD ~~slightly~~ with MBE of 0.10~0.13. The ORAC AOD ~~had not showed~~ ~~good~~ ~~weak~~ agreement with the Xilinhot data, mainly because ~~it exists~~ possible “outliers” ~~exist~~ in March to June 2008 and March 2010.

5

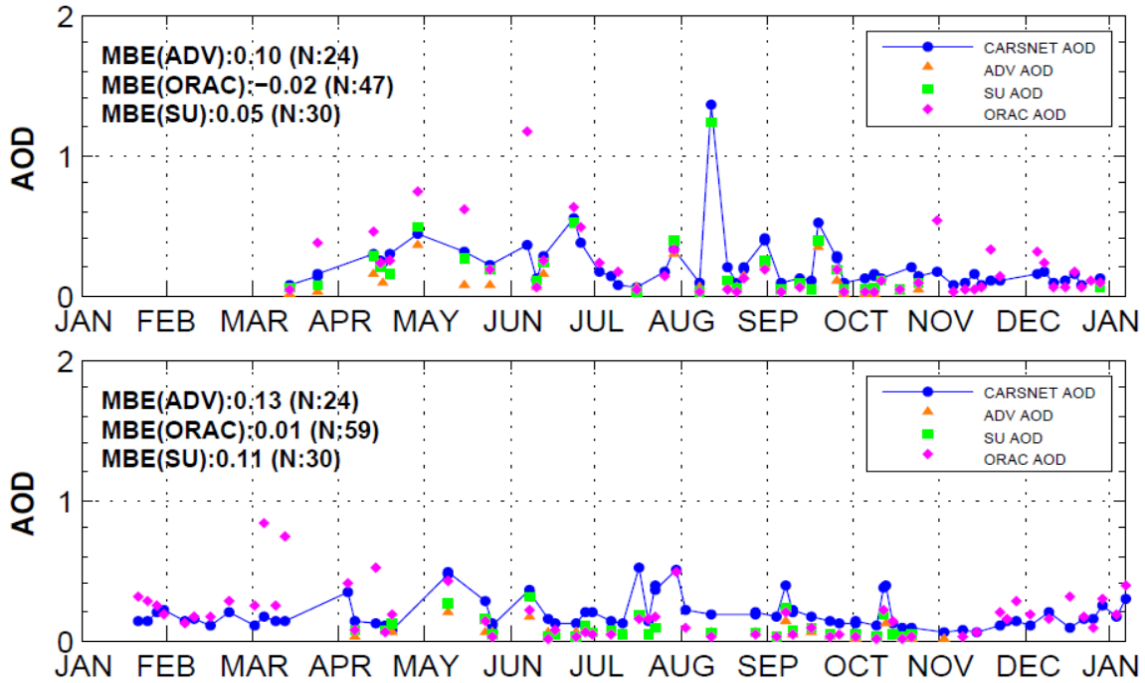


Fig. 12. is the ~~C~~omparison of SU AOD with CARSNET AOD ~~over the site of at~~ Xilinhot in 2008 and 2010.

Table 5. Statistics of validation results of different products over different sites.

Site	Algorithm	N	MSA	MAA	MBE	MAE	RMSE	RMB	CC	KAPPA	Within EE3	Within EE4
Linan	ADV	33	0.346	0.462	-0.116	0.122	0.088	0.748	0.916	0.341	84.9%	90.9%
	ORAC	48	0.426	0.470	-0.044	0.131	0.144	0.906	0.647	0.668	58.3%	70.8%
	SU	40	0.430	0.484	-0.054	0.082	0.093	0.889	0.917	0.650	85.0%	90.0%
SACOL	ADV	46	0.156	0.285	-0.129	0.132	0.068	0.547	0.763	0.283	89.1%	91.3%
	ORAC	74	0.286	0.314	-0.028	0.102	0.170	0.910	0.683	0.595	67.6%	73.0%
	SU	49	0.265	0.291	-0.027	0.062	0.072	0.908	0.849	0.878	77.6%	83.7%

Shangdianzi	ADV	52	0.172	0.297	-0.125	0.131	0.087	0.578	0.780	0.339	78.9%	90.4%
	ORAC	66	0.267	0.304	-0.037	0.107	0.134	0.879	0.781	0.407	54.6%	65.2%
	SU	46	0.285	0.402	-0.117	0.128	0.101	0.710	0.924	0.457	82.6%	87.0%
XiangHe	ADV	33	0.184	0.284	-0.100	0.102	0.070	0.649	0.921	0.169	97.0%	97.0%
	ORAC	34	0.227	0.240	-0.013	0.091	0.096	0.946	0.825	0.577	73.5%	76.5%
	SU	36	0.368	0.392	-0.024	0.058	0.077	0.939	0.984	0.444	88.9%	91.7%
Xilinhot	ADV	49	0.082	0.198	-0.116	0.117	0.046	0.414	0.814	0.148	95.9%	95.9%
	ORAC	110	0.190	0.182	0.008	0.109	0.166	1.043	0.517	0.389	42.7%	48.2%
	SU	61	0.140	0.220	-0.081	0.085	0.063	0.634	0.937	0.444	88.9%	91.7%

For To guarantee of statistical reliability, there must be more than 30 collocated pairs in at one site. The determination of the surface cover at each site is is according to based on the proportion (> 80% for one land type) of each land cover type from the MCD12C1 data at a spatial extent of 50 km × 50 km. If there's no one land cover type accounts for a's proportion larger than 80% in at a given site, it will be identified as mixed; then, we select two or more (sum > 80%) land types with the largest proportions as the main land cover.

5.2 Analysis of algorithm performances in western China

Because it lacks enough sufficient ground-based data in western China are lacking for the from AERONET measurements, only data from CARSNET sites are used in 2008. We picked selected six CARSNET sites which are located in western China and in which there are're more than 25 matches.

Urumchi is, situated at 87.62°E, 43.78°N, 87.62°E, serves as the provincial capital of Xinjiang Uyghur Autonomous Region, and is and is the most remote city in China from in terms of distance to any sea in the world. The dominant land cover at the spatial extent of 50 km × 50 km is grassland, which accounts for about approximately 85%. The ADV, ORAC and SU algorithms all severely underestimated AOD severely, with MBE = 0.22, 0.12 and 0.17, respectively. The MBE is lowest mainly because of the "outlier" in April, which decreases its the MBE.

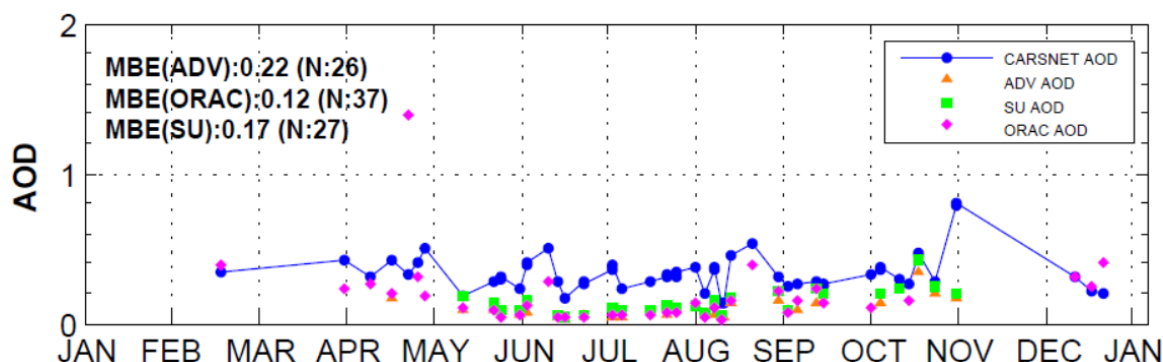


Fig. 13. ~~is the time~~ Time series comparison of AATSR AOD with CARSENT AOD ~~over the site of at~~ Urmchi in 2008.

Ejina is situated at ~~101.07°E, 41.95°N, 101.07°E,~~ and its main land cover ~~of is~~ barren ground (84%) ~~ground~~. The performances of ORAC and SU are at the same high quality level ~~with high quality, the with~~ MBEs ~~are of~~ 0.02 and 0.09, respectively.

5 Another reason why we chose this site is that there ~~are're~~ no matches of ADV products ~~collocated~~ successfully collocated with ground-based data. ~~As result from~~ Based on Fig. 15-, the ORAC algorithm has strong applicability in Ejina and high accuracy in retrieving AOD. The SU algorithm ~~had good performance too~~ also performed well. This ~~explains~~ demonstrates that another limitation of the ADV algorithm is its applicability in calculating AOD in China. Dunhuang is situated at ~~94.68°E, 40.15°N, 94.68°E,~~ and is and is surrounded by barren ground (85%). The same situation is ~~like true for~~ Ejina, that it arises a little

10 ~~bit~~ which causes slight underestimation ~~on at~~ each point but high R and low RMSE for the ORAC algorithm. The performance of the SU algorithm was not as good as that of the ORAC because of its underestimation with MBE = 0.10. The limits of underestimation and applicability of the ADV were more obvious ~~in at~~ this site, as it only had 6 matches ~~on demand and and~~ showed severe underestimation with MBE = 0.17. Tazhong is situated at ~~83.67°E, 39°N, 83.67°E,~~ and is and is surrounded by barren or sparsely ~~vegetation~~ vegetated surface. Almost all land cover is barren ground ~~from according to the~~ MODIS

15 MCD12C1 data. Similar ~~with to~~ the former two sites, the ADV product did ~~not haven't~~ collocated any successful matches ~~in~~ at this site. Both ~~of the~~ ORAC and SU algorithms ~~had exhibited~~ severe underestimation of retrievals, with MBE = 0.17 and 0.20, respectively. The outliers of the ORAC product in February are much higher than the observation data, ~~which makes~~ causing the lower MBE ~~lower~~.

The ~~dominated~~ prevailing climatic ~~ice~~ pattern in western China is a temperate continental climate with ~~clearly~~ four distinct

20 seasons and less precipitation in winter and spring. In conclusion, compared to eastern China, the applicability of the ADV algorithm is not strong, and the underestimation is more severe. In the four selected sites in western China, the performance of the ORAC algorithm is best, even though ~~it exists~~ severe underestimation occurs at ~~in~~ some sites. The accuracy of the SU algorithm is not ~~good~~ as high as the ORAC product, with more severe underestimation and ~~the lower~~ applicability ~~is not strong~~ as ORAC.

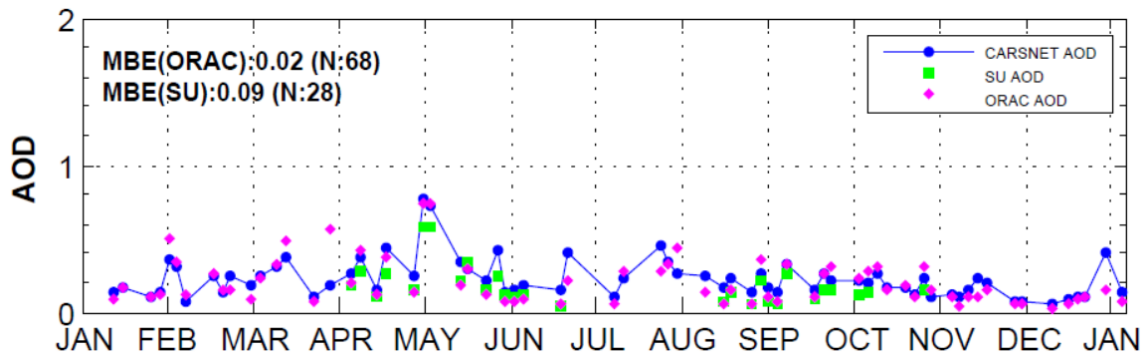


Fig. 14. is the time series comparison of AATSR AOD with CARSENT AOD over the site of Ejina in 2008.

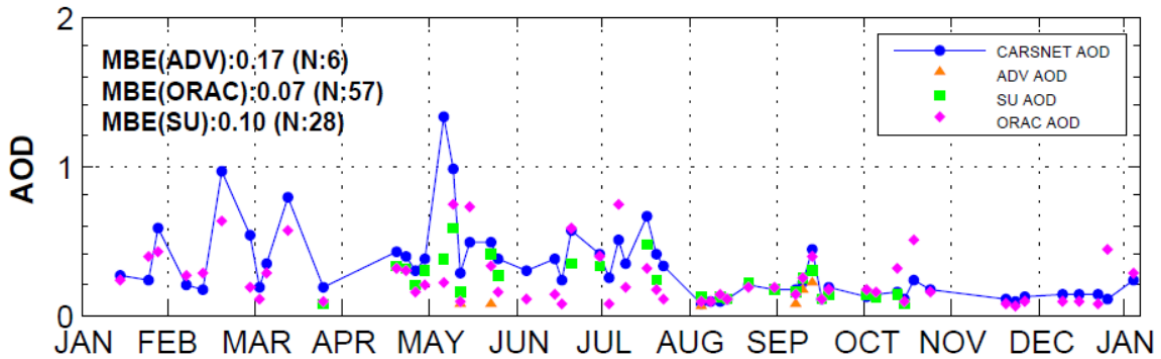


Fig. 15. is the time series comparison of AATSR AOD with CARSENT AOD over the site of Dunhuang in 2008.

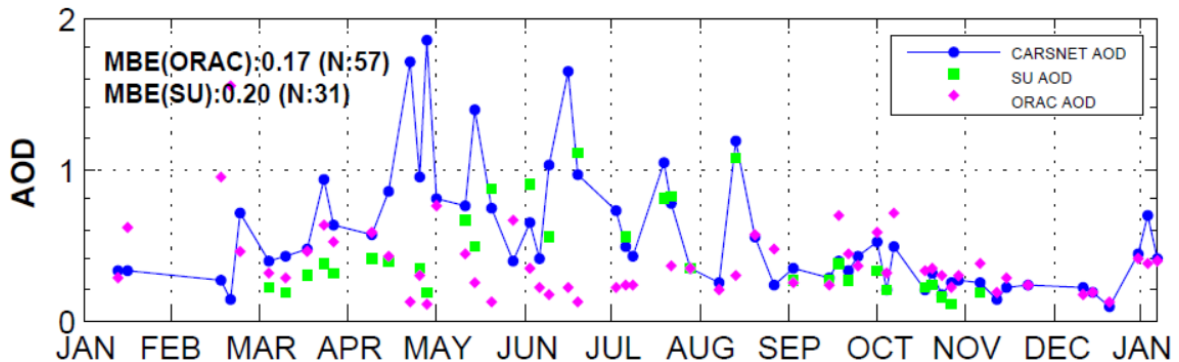


Fig. 16. is the time series comparison of AATSR AOD with CARSENT AOD over the site of Tazhong in 2008.

5.3 Inter-comparison

In conclusion, the SU algorithm ~~has good performances on~~ performs well in calculating AOD over different land covers from March to October. Slight underestimation occurs over barren ground or sparsely vegetation at different times, and there are no

obvious features ~~of in terms of~~ precision ~~on in the~~ time series over grasslands. For complex land surfaces where ~~the dominated dominant~~ land cover is vegetation, the SU algorithm ~~has extremely good performance on is extremely effective in~~ estimating AOD. In the last section, we draw a conclusion that the SU algorithm underestimates AOD over mainland ~~of~~ China in 2008; probably because the ~~dominated dominant~~ land cover in ~~the~~ western China is barren or sparsely vegetation, over which the SU algorithm underestimates AOD more severely.

The ADV algorithm underestimates AOD ~~in at most of the selected sites we selected~~. We categorize these sites into four classes according to ~~the~~ MBEs of different sites: Class_1 (MBE<0.1), Class_2 (0.2>MBE>0.1), Class_3 (0.3>MBE>0.2), ~~and~~ Class_4 (MBE>0.3). The ADV algorithm underestimates AOD over all selected sites, leading to all ~~selected~~ MBEs ~~we select being~~ larger than 0. We make such categories for the purpose of assessing ~~the~~ contribution of different surfaces to AOD estimation. Only XiangHe of 2008 belongs to Class_1, ~~and~~ Linan, Shangdianzi, ~~and~~ SACOL ~~were are~~ classified into Class_2. Only Urumchi is in Class_3. Note that, even though Lanzhou and Datong were not selected due their location, they should be classified into Class_4.

Overall, the ADV algorithm underestimates AODs at all sites but at different levels, as ~~the categories we make above demonstrated by the above categories~~. Serious underestimation occurs over the sites in Class_3 and Class_4 in ~~the~~ western China, where the dominant ~~nted~~ land cover is a mixing of a ~~small portion of~~ urban area and a large portion of grasslands. For the sites in Class_2, ~~there're~~ differences ~~exist~~ between Beijing and SACOL. SACOL is ~~much like similar to~~ the sites in Class_3 and Class_4, the main land cover of which is grasslands. Over the sites in Class_1, the algorithm ~~has good performance performs well;~~ with high R and low MBE, but there ~~are're~~ no ~~common~~ features in ~~common on the terms of~~ surface ~~circumstance conditions~~.

The ORAC product collocates most pairs of all ~~of~~ these products. Most ~~collocated~~ pairs ~~of~~ the SU product and ADV product ~~collocate are distributed at occur in~~ March to October, but the collocated pairs of the ORAC product ~~distribute at occur during~~ each month over some sites in 2008. ~~Since Otherwise, more Because more~~ matches ~~mean suggest more greater~~ errors; for the ~~target of~~ determination of the ~~contribution of~~ “outlier” ~~contribution~~ to ~~the~~ overall performance of the ORAC algorithm, we introduce ~~the~~ ratio of ~~the~~ individual difference to average ~~the~~ differences for each site;

$$DR = \frac{|\tau_{AERO,i} - \tau_{sate,i}|}{(\sum_{i=1}^n |\tau_{AERO,i} - \tau_{sate,i}|) / n} \quad (11)$$

where DR<1; ~~indicating indicates~~ a “relatively good” match; ~~Where~~ >DR >1 ~~indicates a, it's a~~ “relatively poor” match; ~~Where and~~ DR >3 ~~is an, it's probably an~~ “outlier” (see Table 6).

There are no obvious “possible outliers” in Ejina shown in Fig. 15. Most of ~~the~~ DRs are in ~~at the~~ range ~~from of~~ 0 to 3; only two DRs are larger than 3, and the maximum ~~al~~ (overestimation) is 5.112. The retrieved AOD in March is ~~a~~ possible “outlier”;

because it is overestimated, ~~but however, whereas~~ most are underestimated. Another two sites ~~with dominated~~ dominated by ~~land cover of~~ barren or sparsely ~~vegetation-vegetated land cover~~ are Dunhang (~~about approximately~~ 85%) and Tazhong (100%). The ~~circumstance conditions~~ in Tazhong ~~is are~~ complex, and there ~~are is~~ no obvious ~~laws relationship~~ between the CARSNET data and the ORAC AODs. Most of ~~the~~ DRs are less than 3, ~~and a total of~~ 8 DRs are larger than 3. ~~It's basically identified that~~
 5 ~~the one~~ The DR in February is ~~an~~ “outlier”; because the varying tendencies are different between the ORAC product and the ground-based data, ~~and only this points was~~ ~~indicative of~~ overestimation.

Table 6. ~~DR d~~ Distribution of DRs of specific sites.

Site	DR<1	1<DR<3	3<DR<5	5<DR	Total
Urumchi	47	40	2	1	90
Ejina	51	43	1	1	96
Tazhong	63	17	5	3	88
Dunhuang	57	31	1	2	91

The ORAC product has the largest coverage at the ~~cost expense of~~ ~~losing~~ accuracy, especially ~~exist of in the presence of~~
 10 “outliers”, and only the ORAC product ~~has collocated~~ ~~collocate~~ validation pairs over some sites ~~at during~~ each month in ~~all~~ three years. The ORAC algorithm underestimates AODs over Ejina, Tazhong and Dunhuang, but the “possible outliers” reduce the differences between the CARSNET data and the ORAC product. Xilinhot, Urumchi and SACOL ~~have share the~~ same main land cover of grasslands. The problem is that the underestimations over these sites are not at ~~one the same~~ level.

It is worth noting that the ORAC algorithm has ~~the ability in calculating to calculate~~ high AODs; however, most of ~~the AODs~~
 15 ~~of which have~~ DRs are larger than 3, indicating ~~that~~ the estimation of high AOD is unstable ~~with and has~~ large error, ~~even to reduce the whole~~ ~~reducing the overall~~ precision.

6. Seasonal characteristics of three algorithms

~~The mainland China, cross about 60 degree of longitude and 30 degree of latitude, is dominated by monsoon-driven climate. In such vast territory, there are big differences in climate pattern from western to eastern China. The main climate type in~~
 20 ~~eastern and eastern coastal China is monsoon climate. For western China far from the ocean, the climate type is hybrid of monsoon and continental climate. In dry seasons (winter, first half of spring, and last half of autumn), poor vegetation coverage, loosen surface and winds in most northern China regions make coarse particles (sea salt and desert dust) into aerosol. Fine particles from coal combustion in winter and soot from straw burning in autumn is also important source of aerosol. In rainy seasons (mainly in summer), high vegetation blocks dust blowing into aerosol and reduce surface reflectance at visible~~
 25 ~~wavelength. Table 7 shows the seasonal distribution of validation results of three algorithms. For the mainland China which is~~

located in Northern Hemisphere from 20°N to 55°N, the spring time starts from about March to May, the summer time starts from about June to August, the autumn time starts from about September to November and the winter time is about from December to February in next year.

5 **Table 7. Seasonal distribution of validation results of three algorithms.**

		<u>N</u>	<u>MSA</u>	<u>MAA</u>	<u>MBE</u>	<u>MAE</u>	<u>RMSE</u>	<u>RMB</u>	<u>KAPPA</u>
<u>AATSR</u>	<u>3Years</u>	<u>568</u>	<u>0.21</u>	<u>0.35</u>	<u>-0.13</u>	<u>0.14</u>	<u>0.2</u>	<u>0.61</u>	<u>0.38</u>
<u>ADV</u>	<u>Spring</u>	<u>186</u>	<u>0.26</u>	<u>0.41</u>	<u>-0.16</u>	<u>0.16</u>	<u>0.23</u>	<u>0.62</u>	<u>0.4</u>
	<u>Summer</u>	<u>164</u>	<u>0.16</u>	<u>0.29</u>	<u>-0.13</u>	<u>0.14</u>	<u>0.19</u>	<u>0.54</u>	<u>0.26</u>
	<u>Autumn</u>	<u>190</u>	<u>0.2</u>	<u>0.32</u>	<u>-0.12</u>	<u>0.13</u>	<u>0.17</u>	<u>0.62</u>	<u>0.36</u>
<u>AATSR</u>	<u>3Years</u>	<u>1085</u>	<u>0.29</u>	<u>0.31</u>	<u>0.02</u>	<u>0.15</u>	<u>0.25</u>	<u>0.93</u>	<u>0.44</u>
<u>ORAC</u>	<u>Spring</u>	<u>294</u>	<u>0.35</u>	<u>0.37</u>	<u>-0.02</u>	<u>0.18</u>	<u>0.3</u>	<u>0.95</u>	<u>0.5</u>
	<u>Summer</u>	<u>296</u>	<u>0.28</u>	<u>0.35</u>	<u>-0.07</u>	<u>0.17</u>	<u>0.26</u>	<u>0.79</u>	<u>0.4</u>
	<u>Autumn</u>	<u>265</u>	<u>0.23</u>	<u>0.22</u>	<u>0.01</u>	<u>0.1</u>	<u>0.16</u>	<u>1.04</u>	<u>0.43</u>
	<u>Winter</u>	<u>230</u>	<u>0.29</u>	<u>0.28</u>	<u>0.01</u>	<u>0.15</u>	<u>0.25</u>	<u>1.03</u>	<u>0.38</u>
<u>AATSR</u>	<u>3Years</u>	<u>715</u>	<u>0.29</u>	<u>0.4</u>	<u>-0.11</u>	<u>0.12</u>	<u>0.2</u>	<u>0.73</u>	<u>0.5</u>
<u>SU</u>	<u>Spring</u>	<u>222</u>	<u>0.3</u>	<u>0.42</u>	<u>-0.12</u>	<u>0.14</u>	<u>0.24</u>	<u>0.72</u>	<u>0.53</u>
	<u>Summer</u>	<u>241</u>	<u>0.32</u>	<u>0.43</u>	<u>-0.11</u>	<u>0.13</u>	<u>0.2</u>	<u>0.74</u>	<u>0.49</u>
	<u>Autumn</u>	<u>237</u>	<u>0.26</u>	<u>0.35</u>	<u>-0.09</u>	<u>0.11</u>	<u>0.16</u>	<u>0.74</u>	<u>0.49</u>

Low mean of uncertainty (MUs) at 550nm means these retrievals are of high quality in Fig. 17. Most of Std_S are below 0.08, indicating high uniformity of ADV products (see Fig. 17). Most of collocated pairs of ADV AODs are concentrated below the 1-1 line and the RMB is 0.61, showing a tendency of underestimation. This kind of underestimation has an impact on ADV algorithm performances, for example, the RMS error is 0.19 in summer time, otherwise, the corresponding RMB is 0.54, which makes the KAPPA coefficient the smallest (0.26) than other seasons. The MBEs is from -0.12 in autumn to -0.16 in spring in Table 7, which means that the ADV algorithm tends to underestimate AOD in all seasons (except winter) over mainland China (See Figure 18). For monsoon climate, the main aerosol types in many parts of China are influenced by coarse particles (dust from Western China and sea salt from eastern coastal China) in spring time. The performance on calculating aerosol properties of mixture of coarse particles is best in spring time with highest KAPPA coefficient, even though there are some samples with high MUs and the RMS error is 0.23.

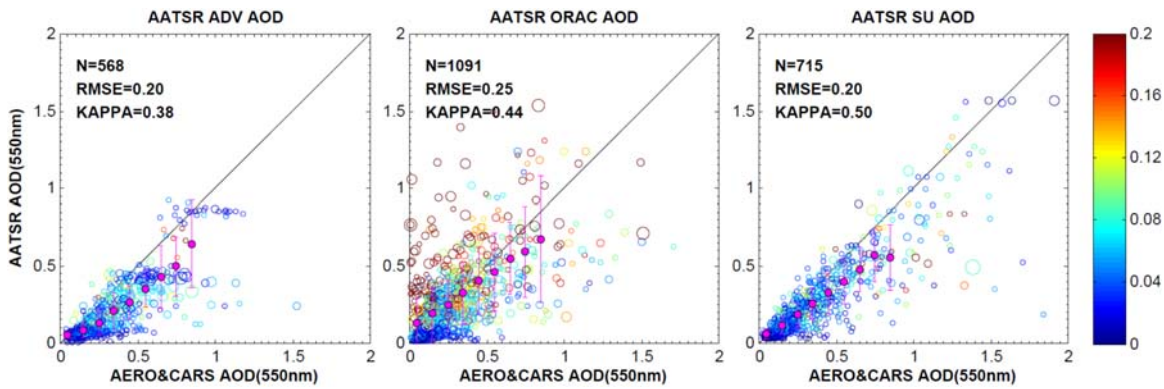
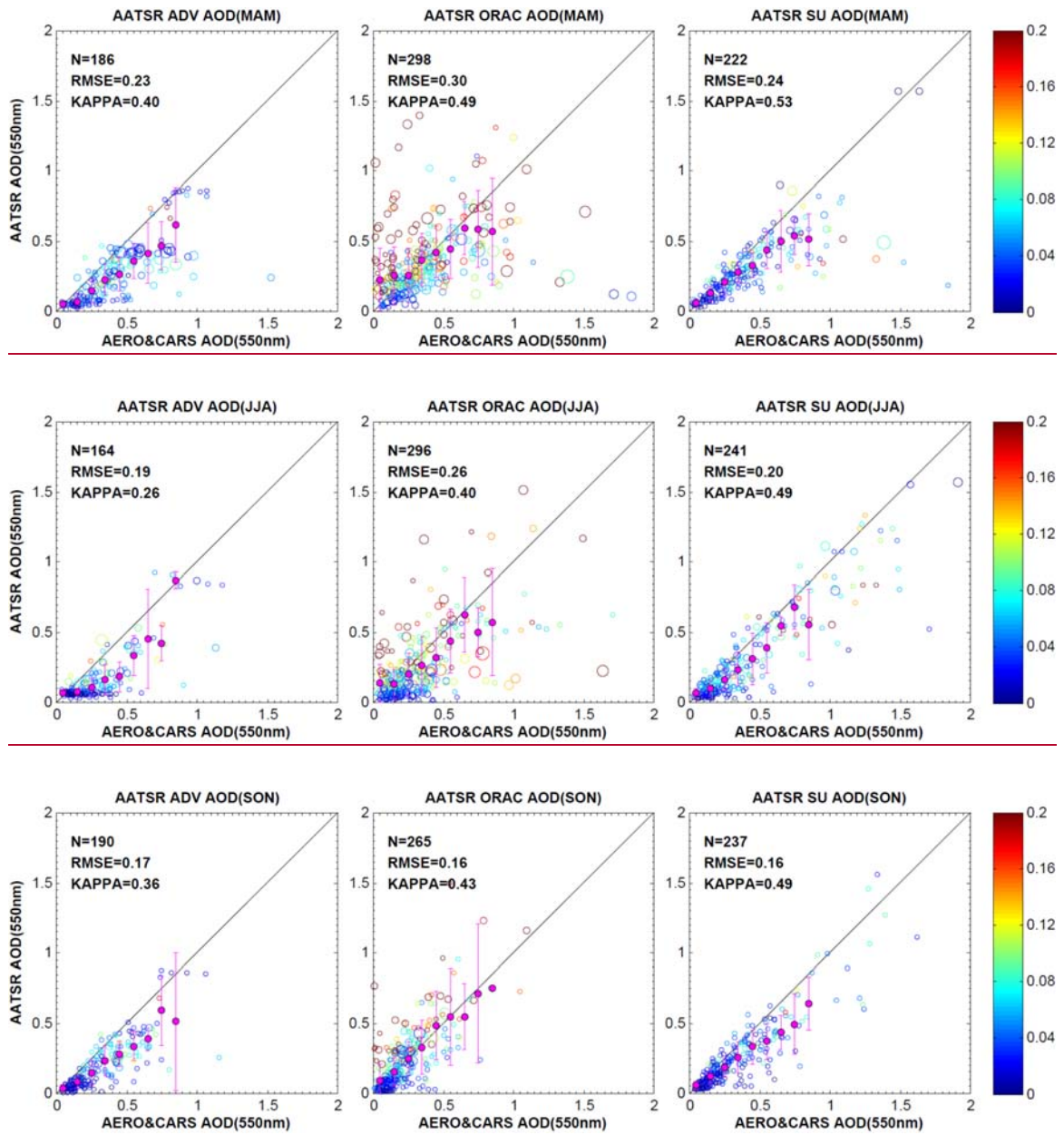


Fig. 17 Scatter plots of AATSR ADV, ORAC and SU L2 AOD products with ground-based data in China in 2007, 2008 and 2010. The black solid line represents 1-1 line. The magenta points are means for specific range of AERO&CARS AOD and the magenta lines are mean $\pm 2\sigma$ of retrievals at certain range. The areas and colours are determined by means of uncertainty (MU) dataset in AATSR L2 products and standard deviation of retrievals (Std S) in collocation frame of 50 km \times 50 km respectively.

The matches of the ORAC product collocated with reference data are distributed discretely at two sides of 1-1 line in Fig. 17. The best performance with high KAPPA coefficient of 0.5 is in spring with no underestimation, even though the RMS error is high about 0.30. The KAPPA coefficient in the autumn time is lower than in the spring time, even though most of evaluation metrics are better in the autumn. Note that, only ORAC product of these three products has been collocated enough matches (more than 30) with reference data in the winter time. The performance of ORAC in winter is between that in spring and autumn without obvious underestimation or overestimation. The limitation of ORAC algorithm is the stability in retrieve aerosol properties, as shown in Fig. 18, the magenta mean $\pm 2\sigma$ lines for each season at each range are longer than those for other two products.

The SU algorithm has better performances in three years, getting KAPPA coefficients of 0.50. Most retrievals in matches are of high quality collocated with reference data and most Std S are lower than 0.08, i.e. the sample quality is high and this coincides with assumption of aerosol properties uniformity in 50km \times 50km area. The best performance on retrieval is in the autumn, lowest RMS error of 0.16 and largest RMB of 0.74 in three seasons shown in Fig. 18. The magenta lines are similar with those of ADV product in corresponding seasons, showing same level of stability in retrieving AOD. The SU algorithm has no obvious differences in retrieving AOD in three seasons. One limitation of SU and ADV algorithms is less than 30 collocated matches in the winter time so that we can't evaluate its performance during that time.



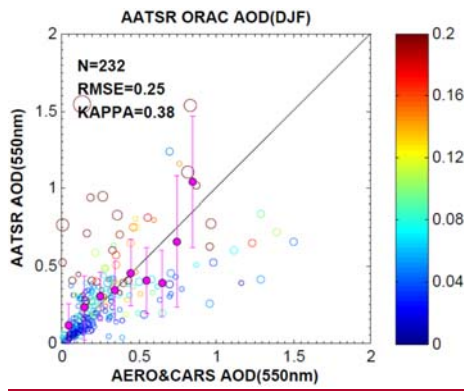


Fig. 18 Scatter plots of AATSR ADV, ORAC and SU L2 AOD products with ground-based data in China in the spring, summer, autumn and winter time of 2007, 2008 and 2010.

The latest MODIS Moderate Resolution Imaging Spectroradiometer (MODIS) Collection 6 (C6) product were released in 2013, including aerosol datasets produced by two “Dark Target” (DT) algorithms (one is for retrieving over ocean and the other is for retrieving over land) and “Deep Blue” algorithm for retrieving over bright or semi-arid surface (Levy et al., 2013). For over land, the DT algorithm uses an updated cloud mask to allow retrieval of heavy aerosol compared to algorithm employed in MODIS Collection 5. It is reported that MODIS C6 products (produced by three algorithms) are of high quality (Sayer et al., 2014). Here, we select both MODIS C6 DT and DB 10km × 10km merged dataset as reference data for cross-validation of AATSR L2 AOD products. The matches in Fig. 19 are randomly chosen from MODIS and AATSR collocated AOD datasets. The ADV AOD has lowest RMSE of 0.11. The SU algorithm has same performance with ORAC (similar RMSE and KAPPA) but with a little underestimation as the magenta line in Fig. 19.

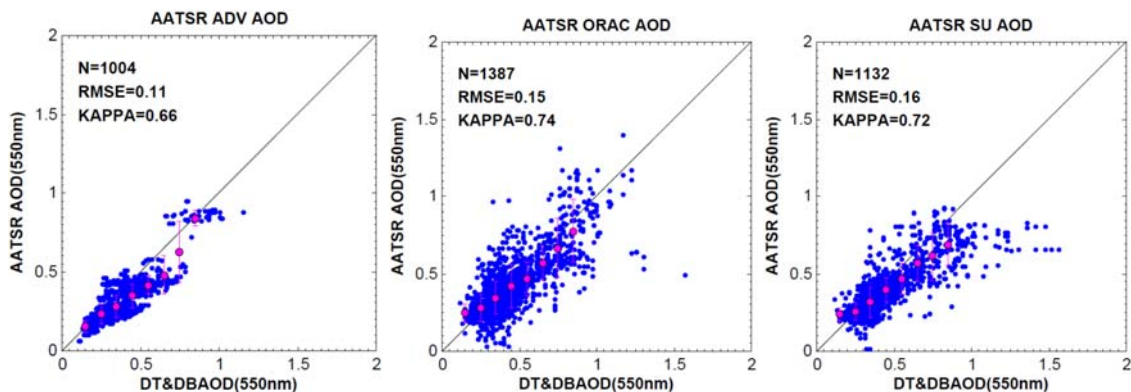


Fig. 19 Scatter plot of AATSR AOD and DT&DB AOD

7.6. Conclusions

These three algorithms (the SU algorithm, the ADV algorithm and the ORAC algorithm) ~~have display~~ different performances ~~on in~~ estimating AOD over mainland China in 2007, 2008 and 2010. However, ~~there's none of the algorithms one show an explicitly having~~ better performance than the other two ~~algorithms~~. The SU and ~~the~~ ADV products have higher accuracy over most ~~selected of sites we select~~ but less coverage, ~~whereas~~ the ORAC product has ~~more greater~~ coverage at the cost of accuracy.

All ~~of~~ these algorithms tend to underestimate AOD to some degree. ~~The u~~ Underestimation ~~gets becomes~~ more severe with ~~increase of increasing~~ AOD or aerosol loading. The method of grouping helps to ~~identify find more, especially~~ “possible outliers” in different regions of aerosol loading.

The precision of ~~the SU algorithm and~~ ADV algorithms is at ~~the same level over different surfaces~~. However, the ~~difference is that~~ SU product has more ~~strict~~ quality control than ~~the~~ ADV product, and it eliminates AODs to make ~~the~~ MBE less than 0.10 over different sites (de Leeuw et al. 2013). Over grasslands and barren vegetation, the SU ~~has displays good a strong~~ performance with slight underestimation (MBE < 0.10). ~~The limitationss~~ of underestimation and applicability of ~~the~~ ADV are more obvious over such sites. For complex surfaces ~~mixed of with~~ two or more land cover ~~types~~, the performances of these three algorithms are at ~~the same level~~. Note that, Lanzhou and Datong are different from other sites, even though the main land cover ~~of them type~~ is grasslands. All ~~of~~ these algorithms ~~have~~ underestimated AOD at a high level, perhaps because these algorithms are ~~not un~~ sensitive to absorptive aerosols.

Only the ORAC product ~~exists shows~~ “possible outliers” identified by equation (7), which ~~substantially decreases its~~ accuracy ~~a lot. Almost, the~~ The most obvious feature of ~~the~~ “possible outliers” is that the retrieved AODs are higher than the ground-based measurements.

As reference data, AERONET L2 data have some limitations, including the distribution and number of sites in mainland China. Most ~~of~~ sites of AERONET ~~are distributed~~ in eastern China and ~~the~~ coastal region of China for special experimental use, ~~leading to we can't get enough as a result, sufficient~~ reference data ~~cannot be obtained~~ to validate ~~the~~ AOD product. The CARSNET data ~~make~~ up for this shortage, because there ~~are're~~ more CARSNET sites in China, especially in ~~western~~ China, where ~~hardly any few~~ AERONET sites ~~had built have been constructed~~. Limited both by reference data and satellite retrievals, most co-allocated pairs ~~are distributing occur~~ in March to November, and few ~~are distributing occur~~ in winter (December, January and February).

Acknowledgements

This work was supported in part by the Ministry of Science and Technology (MOST) ~~of~~ China under ~~G~~grant Nos. 2013CB733403 and 2013AA122801, ~~and by~~ the National Natural Science Foundation of China (NSFC) under ~~G~~grant Nos. 41471306 and GF 30-Y20A02-9003-15/16, and ~~the~~ EU/FP7 MarcPolo project (Grant Agreement ~~Number No.~~ 606953). Part

of the work ~~is done~~ ~~was conducted~~ in preparation for the Aerosol_cci project (ESA-ESRIN project AO/1-6207/09/I-LG), from which three AATSR AOD products were provided. The data for uncertainty analysis and validation came from ~~thirty-four~~ ³⁴ AERONET sites and eight CARSNET ~~datasites~~. We thank the PIs, investigators and their staff for establishing and maintaining the data for this study.

5 References

- Adhikary, B., Kulkarni, S., Dallura, A., Tang, Y., Chai, T., Leung, L. R., Qian, Y., Chung, V., Ramanathan, C. E., and Carmichael, G. R.: A regional scale chemical transport modelling of Asian aerosols with data assimilation of AOD observations using optimal interpolation technique, *Atmos. Environ.*, 42, 8600-8615, doi:10.1016/j.atmosenv.2008.08.031, 2008.
- 10 Andreae, M. O., and Rosenfeld, D.: Aerosol-cloud-precipitation interactions. Part 1. The nature and sources of cloud-active aerosols, *Earth-Sci. Rev.*, 89, 13-41, doi:10.1016/j.earscirev.2008.03.001, 2008.
- Bevan, S., North, P., Los, S., and Grey, W.: A global dataset of atmospheric aerosol optical depth and surface reflectance from AATSR, *Remote Sens. Environ.*, 116, 199-210, doi:10.1016/j.rse.2011.05.024, 2012.
- ~~Bloch, D. A., and Kraemer, H. C., 2×2 Kappa coefficients: Measures of agreement or association, *Biometrics*, 45: 269-287, doi: 10.2307/2532052, 1989.~~
- 15 ~~doi: 10.2307/2532052, 1989.~~
- Bilal, M., J. E. Nichol, and Chan, P. W.: Validation and accuracy assessment of a Simplified Aerosol Retrieval Algorithm (SARA) over Beijing under low and high aerosol loadings and dust storms, *Remote Sens. Environ.*, 153, 50-60, doi:10.1016/j.rse.2014.07.015, 2014.
- ~~Bloch, D. A., and Kraemer, H. C., 2×2 Kappa coefficients: Measures of agreement or association, *Biometrics*, 45: 269-287, doi: 10.2307/2532052, 1989.~~
- 20 ~~doi: 10.2307/2532052, 1989.~~
- Che, H., Zhang, X., Chen, H., Damiri, B., Goloub P., Li, Z., Zhang, X., Wei, Y., Zhou, H., Dong, F., Li, D., and Zhou, T.: Instrument calibration and aerosol optical depth validation of the China Aerosol Remote Sensing Network, *J. Geophys. Res.* 114, doi:10.1029/2008JD011030, 2009.
- Che, Y., Xue, Y., Xu, H., Mikusauskas, R., and She, L.: The inter-comparison of AATSR aerosol optical depth retrievals from various algorithms, in Proceedings of the IEEE International Geoscience and Remote Sensing Symposium held in Milan, Italy, 27-31 July 2015, 2230-2233, doi:10.1109/IGARSS.2015.7326249, 2015.
- 25 Chu, D. A., Kaufman, Y. J., Ichoku, C., Remer, L. A., Tanre', D., and Holben, B. N.: Validation of MODIS aerosol optical depth retrieval over land, dust, *Geophys. Res. Lett.*, 29, MOD2-1–MOD2-4, doi: 10.1029/2001GL013205, 2002.
- Climate Change Initiative: <http://www.esa-aerosol-cci.org/>, last access: 15 January 2016.
- 30 Cohen, J. A.: A coefficient of agreement for nominal scales, *Educ. Psychol. Meas.*, 20, 37-46, 1960.
- de Leeuw, G., Holzer-Popp, T., Bevan, S., Davies, W. H., Desclotres, J., Grainger, R. G., Griesfeller, J., Heckel, A., Kinne, S., Klüser, L., Kolmonen, P., Litvinov, P., Martynenko, D., North, P., Ovigneur, B., Pascal, N., Poulsen, C., Ramon, D.,

- Schulz, M., Siddans, R., Sogacheva, L., Tanré, D., Thomas, G. E., Virtanen, T. H., von Hoyningen-Huene, W., Vountas, M., and Pinnock, S.: Evaluation of seven European aerosol optical depth retrieval algorithms for climate analysis, *Remote Sens. Environ.*, 162, 295-315, doi:10.1016/j.rse.2013.04.023, 2013.
- Flowerdew, R. J., and Haigh, J. D.: Retrieval of aerosol optical thickness over land using the ATSR-2 Dual-Look Satellite Radiometer, *Geophys. Res. Lett.*, 23, 351-354, 1996.
- 5 Foody, G. M.: Status of land cover classification accuracy assessment, *Remote Sens. Environ.*, 80, 185-201, doi:10.1016/S0034-4257(01)00295-4, 2002.
- Griggs, M.: Satellite Observation of Atmospheric Aerosols During the EOMET Cruise, *J. Atmos. Sci.*, 36, 695-698, 1979.
- Grey, W. M. E., North, P. R. J., Los, S. O., and Mitchell, R. M.: Aerosol Optical Depth and Land Surface Reflectance from Multiangle AATSR Measurements: Global Validation and Intersensor Comparisons, *IEEE T. Geosci. Remote*, 44, 2184-2197, doi:10.1109/TGRS.2006.872079, 2006.
- 10 Holben, B. N., Eck, T. F., Slutsker, I., Tanre', D., Buis, J. P., Setzer, K. A., Vermote, E., Reagan, J. A., Kaufman, Y. J., Nakajima, T., Lavenu, F., Jankowiak, I., and Smirnov, A.: AERONET—A Federated Instrument Network and Data Archive for Aerosol Characterization, *Remote Sens. Environ.*, 66, 1-16, doi:10.1016/S0034-4257(98)00031-5, 1998.
- 15 Holzer-Popp, T., de Leeuw, G., Griesfeller, J., Martynenko, D., Kluser, L., Bevan, S., Davies, W., Ducos, F., Deuze, J. L., Grainger, R. G., von Hoyningen-Huene, W., Kolmonen, P., Litvinov, P., North, P., Poulsen, C. A., Ramon, D., Siddans, R., Sogacheva, L., Tanre, D., Thomas, G. E., Vountas, M., Descloitres, J., Griesfeller, J., Kinne, S., Schulz, M., and Pinnock, S., A.: Aerosol retrieval experiments in the ESA Aerosol_cci project, *Atmos. Meas. Tech.*, 6, 1919-1957, doi:10.5194/amt-6-1919-2013, 2013.
- 20 ~~Levy, R. C., Remer, L. A., and Dubovik, O.: Global aerosol optical properties and application to Moderate Resolution Imaging Spectroradiometer aerosol retrieval over land, *J. Geophys. Res.*, 112, doi:10.1029/2006JD007815, 2007.~~
- ~~Levy, R. C., Remer, L. A., Kleidman, R. G., Mattoo, S., Ichoku, C., Kahn, R., and Eck, T. F.: Global evaluation of the Collection 5 MODIS dark target aerosol products over land, *Atmos. Chem. Phys.* 10, 10399-10420, doi:10.5194/acp-10-10399-2010, 2010.~~
- 25 ~~Li, M., and Zhang, L.: Haze in China: Current and future challenges, *Environ. Pollut.*, 189, 85-86, doi:10.1016/j.envpol.2014.02.024, 2014.~~
- Ichoku, C, D. A. Chu, Mattoo, S., Kaufman, Y. J., Remer, L. A., Tanre', D., Slutsker, I., and Holben, B. N.: A spatio-temporal approach for global validation and analysis of MODIS aerosol products, *Geophys. Res. Lett.*, 29, MOD1-1-MOD1-4, doi: 10.1029/2001GL013206, 2002.
- 30 Intergovernmental Panel on Climate Change (IPCC): Climate Change 2013: The Physical Science Basis, Contribution of Working Group I to the Fifth Assessment Report of the IPCC, edited by: Stocker, T. F., Qin, D., Plattner, G. K., Tignor, M., Allen, S. K., Boschung, J., Nauels, A., Xia, Y., Bex, V., an Midgley, P. M., Cambridge University Press, Cambridge, UK, New York, NY, USA, 2013.

- Kinne, S., Lohmann, U., Feichter, J., Schulz, M., Timmreck, C., Ghan, S., Easter, R., Chin, M., Ginoux, P., Takemura, T., Tegen, I., Koch, D., Herzog, M., Penner, J., Pitari, G., Holben, B., Eck, T., Smirnov, A., Dubovik, O., Slutsker, I., Tanre, D., Torres, O., Mishchenko, M., Geogdzhayev, I., Chu, D. A., and Kaufman, Y.: Monthly averages of aerosol properties: a global comparison among models, satellite data, and AERONET ground data, *J. Geophys. Res.*, 108, doi:10.1029/2001JD001253, 2003.
- 5 Kahn, R. A., Gaitley, B. J., Martonchik, J. V., Diner, D. J., and Crean, K. A.: Multiangle Imaging Spectroradiometer (MISR) global aerosol optical depth validation based on 2 years of coincident Aerosol Robotic Network (AERONET) observations, *J. Geophys. Res.*, 110, doi:10.1029/2004JD004706, 2005.
- Kahn, R. A., Li, W. H., Moroney, C., Diner, D. J., Martonchik, J. V., and Fishbein, E.: Aerosol source plume physical characteristics from Space-based multiangle imaging, *J. Geophys. Res.*, 112, doi:10.1029/2006JD007647, 2007.
- 10 Kahn, R. A., Nelson, D. L., Garay, M. J., Levy, R. C., Bull, M. A., Diner, D. J., Martonchik, J. V., Paradise, S. R., Hansen, E. G., and Remer, L. A.: MISR aerosol product attributes and statistical comparisons With MODIS, *IEEE T. Geosci. Remote*, 47, 4095-4113, doi: 10.1109/TGRS.2009.2023115, 2009.
- Kahn, R. A., Garay, M. J., Nelson, D. L., Levy, R. C., Bull, M. A., Diner, D. J., Martonchik, J. V., Hansen, E. L., Remer, L. A., and Tanre', D.: Response to "Ward Unified Satellite Climatology of Aerosol Properties. 3. MODIS versus MISR versus AERONET", *J. Quant. Spectrosc. Ra.*, 112, 901-909, 2011.
- 15 Kaufman, Y. J., Tanre', D., Remer, L. A., Vermote, E. F., Chu, A., Holben, B. N.: Operational remote sensing of tropospheric aerosol over land from EOS moderate resolution imaging spectroradiometer, *J. Geophys. Res.*, 102, 17051-17067, doi: 10.1029/96JD03988, 1997.
- 20 Kinne, S., O'Donnel, D., Stier, P., Kloster, S., Zhang, K., Schmidt, H., Rast, S., Giorgetta, M., Eck, T. F., and Stevens, B.: MAC-v1: A new global aerosol climatology for climate studies, *J. Adv. Model Earth Syst.*, 5, 704-740, doi:10.1002/jame.20035, 2013.
- Kokhanovsky, A. A., Curier, R. L., de Leeuw, G., Grey, W. M. F., Lee, K. H., Bennouna, Y., Schoemaker, R., and North, P. R. J.: The inter-comparison of AATSR dual-view aerosol optical thickness retrievals with results from various algorithms and instruments, *Int. J. Remote Sens.*, 30, 4525-4537, doi:10.1080/01431160802578012, 2009.
- 25 Kokhanovsky, A. A. and de Leeuw, G.: *Satellite Aerosol Remote Sensing Over Land*, Springer, Berlin, 2009.
- [Kolmonen, P., Sogacheva, L., Timo, H., Virtanen, de Leeuw, G., and Kulmala, M.: The ADV/ASV AATSR aerosol retrieval algorithm: current status and presentation of a full-mission AOD dataset, Int. J. Digit. Earth, doi: 10.1080/17538947.2015.1111450, 2015.](#)
- 30 [Levy, R. C., Mattoo, S., Munchak, L. A., Remer, L. A., Sayer, A. M., Patadia, F., and Hsu, N. C.: The Collection 6 MODIS aerosol products over land and ocean, Atmos. Meas. Tech., 6, 2989-3034, doi:10.5194/amt-6-2989-2013, 2013.](#)
- [Levy, R. C., Remer, L. A., and Dubovik, O.: Global aerosol optical properties and application to Moderate Resolution Imaging Spectroradiometer aerosol retrieval over land, J. Geophys. Res., 112, doi:10.1029/2006JD007815, 2007.](#)

- Levy, R. C., Remer, L. A., Kleidman, R. G., Mattoo, S., Ichoku, C., Kahn, R., and Eck, T. F.: Global evaluation of the Collection 5 MODIS dark-target aerosol products over land, *Atmos. Chem. Phys.* **10**, 10399-10420, doi:10.5194/acp-10-10399-2010, 2010.
- Li, M., and Zhang, L.: Haze in China: Current and future challenges, *Environ. Pollut.*, **189**, 85-86, doi:10.1016/j.envpol.2014.02.024, 2014.
- 5 ~~McMurry, P. H.: A review of atmospheric aerosol measurements, *Atmos. Environ.*, **34**, 1959-1999, doi:10.1016/S1352-2310(99)00455-0, 1999.~~
- Martonchik, J. V., Diner, D. J., Kahn, R., and Gaitley, B.: Comparison of MISR and AERONET aerosol optical depths over desert sites, *Geophys. Res. Lett.*, **31**, doi:10.1029/2004GL019807, 2004.
- 10 ~~McMurry, P. H.: A review of atmospheric aerosol measurements, *Atmos. Environ.*, **34**, 1959-1999, doi:10.1016/S1352-2310(99)00455-0, 1999.~~
- Misra, A., Jayaraman, A., and Ganguly, D.: Validation of MODIS derived aerosol optical depth over Western India, *J. Geophys. Res.*, **113**, doi:10.1029/2007JD009075, 2008.
- Mei, L., Xue, Y., Xu, H., Guang, J., Li, Y., Wang, Y., Ai, J., Jiang, S., And He, X.: Validation and analysis of aerosol optical thickness retrieval over land, *Int. J. Remote Sens.*, **33**, 781-803, doi:10.1080/01431161.2011.577831, 2012.
- 15 ~~North, P.R.J., Briggs, S.A., Plummer, S.E. and Settle, J.J., (1999). Retrieval of land surface bidirectional reflectance and aerosol opacity from ATSR-2 multiangle imagery, *IEEE T. GEOSCI. REMOTE*, **37**, 526-537, doi: 10.1109/36.739106, 1999.~~
- ~~Kolmonen, P., Sogacheva, L., Timo, H., Virtanen, de Leeuw, G., and Kulmala, M.: The ADV/ASV AATSR aerosol retrieval algorithm: current status and presentation of a full mission AOD dataset, *Int. J. Digit. Earth*, doi: 10.1080/17538947.2015.1111450, 2015.~~
- 20 Prins, E. M., Feltz, J. M., Menzel, W. P., Menzel, W. P., and Ward, D. E.: An overview of GOES-8 diurnal fire and smoke results for SCAR-B and 1995 fire season in South America, *J. Geophys. Res.*, **103**, 31821-31836, 1998.
- Rosenfield, G. H., and Fitzpatrick-Lins, K.: A coefficient of agreement as a measure of thematic classification accuracy, *Photogramm. Eng. Rem. S.*, **52**, 223-227, 1986.
- 25 Remer, L. A., Kaufman, Y. J., Tanré, D., Mattoo, S., Chu, D. A., Martins, J. V., Li, R.-R., Ichoku, C., Levy, R. C., Kleidman, R. G., Eck, T. F., Vermote, E., and Holben, B. N.: The MODIS aerosol algorithm, products, and validation, *J. Atmos. Sci.*, **62**, 947-973, doi: http://dx.doi.org/10.1175/JAS3385.1, 2005.
- ~~Stehman, S. V.: Selecting and interpreting measures of thematic classification accuracy, *Remote Sens. Environ.*, **62**, 77-89, doi:10.1016/S0034-4257(97)00083-7, 1997.~~
- 30 Samet, J. M., Dominici, F., Curriero, F. C., Coursac, I., And Zeger, S. L.: Fine particulate air pollution and mortality in 20 U.S. cities, 1987-1994, *New Engl. J. Med.*, **343**, 1742-1749, doi: 10.1056/NEJM200012143432401, 2000.

- Sayer, A. M., Munchak, L. A., Hsu, N. C., Levy, R. C., Bettenhausen, C., and Jeong, M.-J., MODIS Collection 6 aerosol products: Comparison between Aqua's e-Deep Blue, Dark Target, and "merged" data sets, and usage recommendations. *119(24)*, 13,965–13,989, 2014 (DOI: 10.1002/2014JD022453)
- 5 Stehman, S. V.: Selecting and interpreting measures of thematic classification accuracy, *Remote Sens. Environ.*, 62, 77-89, doi:10.1016/S0034-4257(97)00083-7, 1997. Torres, O., Bhartia, P. K., Herman, J. R., Sinyuk, A., Ginoux, P., and Holben, B.: A long-term record of aerosol optical depth from TOMS observations and comparison to AERONET measurements, *J. Atmos. Sci.*, 59, 398-413, doi: [http://dx.doi.org/10.1175/1520-0469\(2002\)059<0398:ALTROA>2.0.CO;2](http://dx.doi.org/10.1175/1520-0469(2002)059<0398:ALTROA>2.0.CO;2), 2002.
- 10 Takahashi, Y., Higashi, M., Furukawa, T., and Mitsunobu, S.: Change of iron species and iron solubility in Asian dust during the long-range transport from western China to Japan, *Atmos. Chem. Phys.*, 11, 11237-11252, doi:10.5194/acp-11-11237-2011, 2011.
- Thomas, G. E., Poulsen, C. A., Curier, R. L., de Leeuw, G., Marsh, S. H., Carboni, E., Grainger, R. G., and Siddans, R.: Comparison of AATSR and SEVIRI aerosol retrievals over the Northern Adriatic, *Q. J. Roy. Meteor. Soc.*, 133, 85-95, doi: 10.1002/qj.126, 2007.
- Thomas, G. E., Carboni, E., Sayer, A. M., Poulsen, C. A., Siddans, R., and Grainger, R. G.: Oxford-RAL Aerosol and Cloud (ORAC): Aerosol Retrievals from Satellite Radiometers, In: *Aerosol Remote Sensing Over Land*, edited by: Kokhanovsky, A. and de Leeuw, G., Springer, Berlin, 2009.
- 15 Torres, O., Bhartia, P. K., Herman, J. R., Sinyuk, A., Ginoux, P., and Holben, B.: A long-term record of aerosol optical depth from TOMS observations and comparison to AERONET measurements, *J. Atmos. Sci.*, 59, 398-413, doi: [http://dx.doi.org/10.1175/1520-0469\(2002\)059<0398:ALTROA>2.0.CO;2](http://dx.doi.org/10.1175/1520-0469(2002)059<0398:ALTROA>2.0.CO;2), 2002.
- 20 Tzanis, C., and Varotsos, C.A.: Tropospheric aerosol forcing of climate: a case study for the greater area of Greece, *Int. J. Remote Sens.*, 29, 2507-2517, 2008.
- Varotsos, C., Ondov, J., Tzanis, C., Ozturk, F., Nelson, M., Ke, H., and Christodoulakis, J.: An observational study of the atmospheric ultra-fine particle dynamics, *Atmos. Environ.*, 59, 312-319, 2012.
- Varotsos, C.: Airborne measurements of aerosol, ozone, and solar ultraviolet irradiance in the troposphere, *J. Geophys. Res.*, 110, D09202, doi:10.1029/2004JD005397, 2005
- 25 Veefkind, J. P., de Leeuw, G., and Durkee, P. A.: Retrieval of aerosol optical depth over land using two-angle view satellite radiometry during TARFOX, *Geophys. Res. Lett.*, 25, 3135-3138, 1998.
- 30 ~~Takahashi, Y., Higashi, M., Furukawa, T., and Mitsunobu, S.: Change of iron species and iron solubility in Asian dust during the long range transport from western China to Japan, *Atmos. Chem. Phys.*, 11, 11237-11252, doi:10.5194/acp-11-11237-2011, 2011.~~
- Xue, Y., Xu, H., Guang, J., Mei, L., Guo, J., Li, C., Mikusauskas, R., and He, X.: Observation of an agricultural biomass burning in central and east China using merged aerosol optical depth data from multiple satellite missions, *Int. J. Remote Sens.*, 35, 5971-5983, doi:10.1080/2150704X.2014.943321, 2014.