POST PROCESSING OF AEROSOL DATASET RETRIEVED FROM SATELLITE DATA USING ROBUST ESTIMATION

Chi Li[1, 4], Yong Xue[*2, 3], Yingjie Li[3, 4], Jie Guang[3]

[1]Center for Earth Observation and Digital Earth, Chinese Academy of Sciences, No.9 Dengzhuang South Road, Haidian District, Beijing 100094, China
[2]Faculty of Computing, London Metropolitan University, 166-220 Holloway Road, London N7 8DB, UK
[3]State Key Laboratory of Remote Sensing Science, Jointly Sponsored by the Institute of Remote Sensing Applications of Chinese Academy of Sciences and Beijing Normal University, Institute of Remote Sensing Applications, Chinese Academy of Sciences, No. 20 Datun Road, Chaoyang District, Beijing 100101, China
[4]Graduate University of Chinese Academy Sciences, No. 19 Yuquan Road, Shijingshan District, Beijing, 100049, China
E-mails: {wslygr@163.com; yxue@irsa.ac.cn; liyingjie_rs2003@163.com; guangjier@163.com}

KEY WORDS: Aerosol Optical Depth (AOD), Homogeneity, Robust Estimation

ABSTRACT: It is widely accepted that aerosols are spatially correlative. On the retrieved AOD (aerosol optical depth) map, large gradients in the aerosol concentrations are unexpected yet fairly commonplace, making quality control indispensable. During the AOD retrieval at 10 km resolution, one pixel at 10-km resolution is composed of 100 original pixels at 1-km resolution. For these pixels in 1-km resolution, there are outliers harming the homogeneity of the 10 x 10 'box' and making the estimation of the pixel's TOA (top of atmosphere) reflectance difficult, which lies an intensive influence on the quality of AOD results. In this paper, we apply the IGG (Institute of Geodesy and Geophysics) scheme of robust estimation to detecting and removing outliers and to acquiring the TOA reflectance of the whole pixels. By visually comparing and quantitatively validating the AOD results of our method with that of immediate averaging and residual test algorithms, we preliminary elicit that robust estimation performs best in the quality improvement of AOD results.

1. INTRODUCTION

The atmospheric aerosol affects the radiative transfer in both direct and indirect ways. Accumulating evidences have been released to indicate that aerosol is an important factor of climate (King et al., 1999; Gao et al., 2009), leading prevalent aerosol researches in recent years. Aerosol optical depth (AOD) is a basic measurement of aerosol, the retrieval of which falls into two methods: Ground survey and satellite remote sensing (RS). Compared to the insufficiency of ground survey, using RS to retrieve AOD may work more efficiently and boost the possibility of learning the large area aerosol in real-time (Gao et al., 2009). Therefore a range of algorithms have been designed to retrieve AOD using RS images derived from different satellites and sensors. In the last few years, AOD retrieval has been much evolved depending on the next generation of satellite imagers such as the Moderate Resolution Imaging Spectrometer (MODIS) due to the enhanced number of narrow and well-chosen spectral channels (King et al., 1999). Based on these advantages, Synergetic Retrieval of Aerosol Properties (SRAP-MODIS) model can be used for various ground surfaces, including for high reflective surface (Tang et al., 2005).

Aerosol is also spatially correlative as widely accepted, leaving the spatial distribution of AOD continuous. But, unexpected large gradients and 'severe' AOD values in the aerosol concentrations are fairly common after the quantitatively calculation of AOD from satellite data. As we see in Figure 1 (a), the spatial distribution of AOD values are not that smooth or continuous, and obviously one extreme high (red) value (declared in red circles) is surrounded by many low (blue) values. None of these abnormities is favorable, making quality control of the retrieval necessary.

The focus of our investigation is on the aerosol retrieval of 10-km resolution using SRAP-MODIS model. Since the SNR (Signal to Noise Ratio) of one single pixel of MODIS data is not high enough to make the pixel sufficiently sensitive to characterize aerosol, retrieval of 10-km resolution will better reduce the noise (Remer et al., 2006). In our case one pixel at 10-km resolution is composed of 100 original pixels at 1-km resolution. Cloudy pixels are ditched out at first as SRAP-MODIS can only be accomplished on scenes without cloud contamination. In the remaining pixels there are still outliers with extreme TOA reflectance, harming the homogeneity of the 10 x 10 'box' and making the estimation difficult, at last laying an intensive influence on the quality of AOD results.

Related measures are taken in other retrieving algorithms. For example, in the algorithm over ocean from MODIS proposed by NASA, after screening out all the cloudy and sediment pixels, the algorithm sorts the remaining pixels of the 400-pixels' 'box' at 500m resolution according to their $\rho_{0.86}$ value, discards the darkest and brightest 25%, eliminating residual cloud contamination, cloud shadows, or other unusual extreme conditions (outliers) in the box. Then the mean reflectance of the remaining pixels is calculated if there are considerable pixels left (Remer et al., 2006). However, this method is not suitable in our case as there are at most 100 pixels, instead of 400. Simply applying this method will only result in less residual pixels and worse estimation. On the other hand, if pixels in the box extremely deviate from each other, this method performs badly as remaining pixels are still inhomogeneous. We are willing to remove all the 'severe' pixels and make use of as many 'useful' pixels as possible at the same time.

Approaches to controlling the outlier influence fall into two broad categories: outlier identification and robust estimation. A problem with outlier identification is that the 'clean' subset is rarely known (Yang et al., 2002). This problem seldom matters in the theory of robust estimation. Robust estimation, aiming at resisting outliers and overcoming the ill-conditioning distribution, has been investigated for long time in statistics. Many methods have been proposed and applied in the field of Geodesy, the choice of which is based on the approximate distribution of the observations (Yang et al., 1999). In our case, the TOA reflectance of each pixel is not correlative to others, thus we can make use of the robust parameter estimation for independent observations into this case. Many methods are available, from which we choose the IGG (Institute of Geodesy and Geophysics) scheme to get the TOA reflectance of the whole pixels. Based on the theory of equivalent weight, the IGG scheme can detect and remove the outliers precisely and smoothly with a high breakdown point (Zhou, 1989).

In this paper we make use of four visible bands from TERRA and AQUA and the SRAP-MODIS model to retrieve AOD. To certify the advantage of robust estimation, we also tested the method of residual test to detect and remove outliers and compare the results of these two methods and that of immediate averaging with Aerosol Robotic Network (AERONET) observations.

2. MODELS

SRAP-MODIS model can be simply explained in the equations as follows (Tang et al., 2005; Xue et al., 1995):

$$\begin{cases} A_{j,\lambda_i} = \dfrac{(A'_{j,\lambda_i} b - a_j) + a_j (1 - A'_{j,\lambda_i}) e^{(a_j - b)\varepsilon[\tau_M^\lambda(\infty) + \tau_A^\lambda(\infty)]\sec\theta'_j}}{(A'_{j,\lambda_i} b - a_j) + b(1 - A'_{j,\lambda_i}) e^{(a_j - b)\varepsilon[\tau_M^\lambda(\infty) + \tau_A^\lambda(\infty)]\sec\theta'_j}} \\ \tau_M^\lambda(\infty) = 0.00879 \lambda_i^{-4.09} \\ \tau_A^\lambda(\infty) = \beta \lambda_i^{-\alpha} \\ K = \dfrac{A_{1,\lambda_i}}{A_{2,\lambda_i}} = \dfrac{A'_{1,2.13\mu m}}{A'_{2,2.13\mu m}} \end{cases} \quad (1)$$

where j = 1, 2 means the TERRA and AQUA observations close in time, λ_i = 1, 2, 3 means the three bands of 470, 550, 670nm; $a_j = \sec\theta_j$, b and ε are known constants. In SRAP-MODIS, the Angstrom wave exponent α is regarded as a constant and β is varied for two observations. The ratio K of the surface reflectance in the two observations is a constant assumed to be independent of the wavelength and decided by the TOA reflectance in 2.13μm where the atmospheric influence on radiative transfer is ignorable (Kaufman et al., 1997). Actually, it is difficult to get the analytical solution of the nonlinear equations above. However, an approximate numerical method can be obtained by means methods such as Newton iteration algorithm.

Robust M (maximum likelihood type) estimation (Huber, 1964) is the basic and most commonly used type of robust estimation, which has been widely studied and applied in geodesy. The IGG scheme based on the equivalent weight method (Zhou, 1989) is one type of M estimation. The main idea of the equivalent weight is to robustify the calculation procedures of the parameter estimation using a suitable weight function (Gui et al., 1998). As we know, the weight of a parameter is an indicator of its precision. Unfortunately under most circumstances we don't know the prior knowledge about the precision of each parameter. In the theory of robust estimation, this problem can be solved by using the weight function. A weight function can be defined as:

$$\overline{P} = P_0 \phi(v) / |v| \tag{2}$$

where P_0 is the original weight of one parameter before the weighed mean calculation, v is the residual of the same parameter after the weighed mean calculation, and $\overline{P}$ is the equivalent weight for the next calculation. In the IGG scheme, the weight function is designed as (Zhou, 1989):

$$\overline{P} = \begin{cases} 0 \,, |v| > k_2\sigma \\ P_0 \dfrac{k_1\sigma}{|v|} \,,\; k_1\sigma < |v| \le k_2\sigma \\ P_0 \,, |v| \le k_1\sigma \end{cases} \tag{3}$$

where σ is the standard deviation after each weighed mean calculation. Normally k_1 could be set near 1.5 and k_2 near 2.5. As the expected distribution of pixels is normality, the probability of $|v|>k_2\sigma$ is very low (lower than 0.01) if k_2 is more than 2.5. We take these observations as gross errors and ditch them out. Second, the observations fitting the function $k_1\sigma<|v|\le k_2\sigma$ (probability lower than 0.13) are also 'harmful' to the homogeneity but still useful, that's why we don't remove them but decrease their weight since the precision is lower. As to the third class of observations, their residuals are low and we took them as 'good' values. Thus their weights are reserved. As the iterative calculation goes on, σ and v will become smaller and smaller and the weight of the parameter will be close to present its real precision.

Initial weight of the calculation is 1 for each pixel. The iteration is not over until $|v|\le k_1\sigma$ is tenable for all remaining pixels. If at least 10 pixels are left, mean of all remaining pixels is calculated as the TOA reflectance of the box. Otherwise homogeneity of pixels in the box is too severe to be used in AOD retrieval and the whole 'box' is eliminated. The last thing to note is that in the SRAP-MODIS model 4 correlative bands are used in the retrieval of one pixel , thus 4 pixels in all bands are removed if a pixel in one band is taken as gross error ($\overline{P} = 0$).

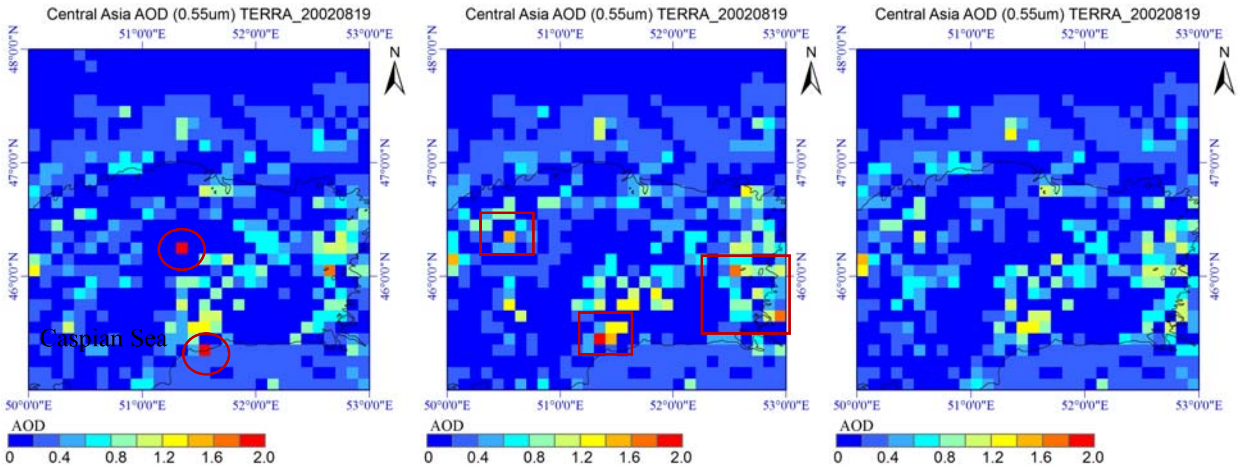
To better verify our algorithm, we also experiment the residual test method. Mean and standard deviation of the pixels are calculated at the beginning. Then we simply take pixels fitting $|v|>1.5\sigma$ as 'severe' and ditch them out. Mean of all remaining pixels is taken to present the box if at least 10 pixels are left.

3. DATA AND STUDY AREA

In this paper, the date and time to validate our model are TERRA and AQUA data over Asia, from August 19 to 25, 2002, covering the area between 15°N - 60°N and 35°E - 150°E. Since AERONET provides a long-term, continuous and reliable database of aerosol optical depth used in the characterization and validation of satellite retrievals, AERONET measured data in the same time and area was also collected to do the preliminary validation, with level 1.5 and 2.0. In order to precisely validate our results, only AERONET observations acquired 30 minutes around the time of TERRA passing is used.

4. RESULTS AND VALIDATION

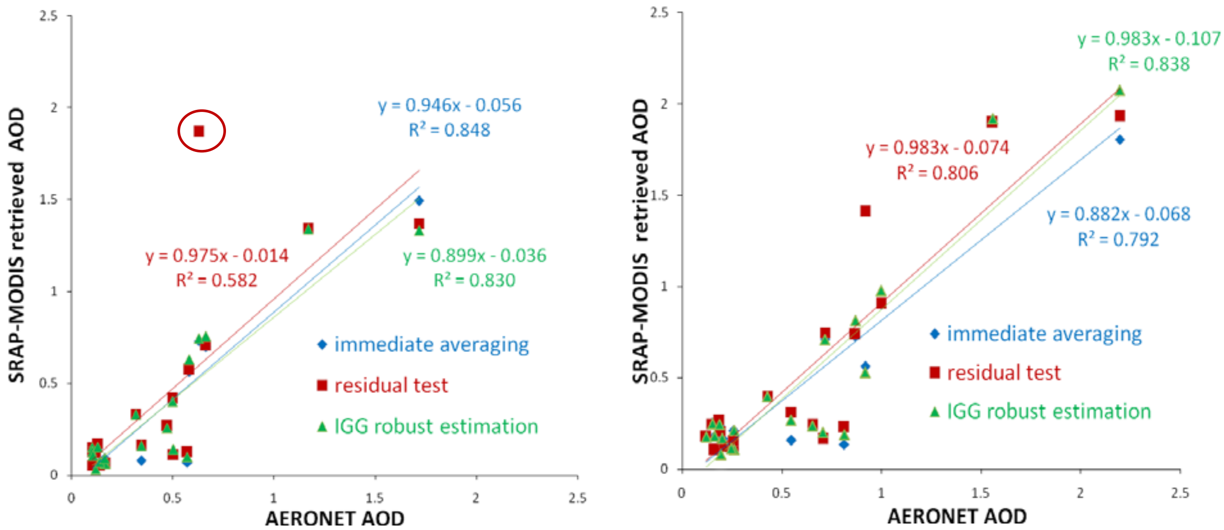
Results in band4 (550nm) of TERRA from SRAP-MODIS retrieval using the three methods discussed above to estimate the TOA reflectance of the box in Centre Asia (45°N - 48°N and 50°E - 53°E), August 19th, 2002 are presented here in Figure 1. We can easily see that extreme values (red color) at the northeast corner of Caspian Sea in Figure 1(a) disappears in the other two figures, suggesting that the AOD result in this point is improved to be more reasonable by applying two methods to controlling outliers. Still in Figure 1(b) we can see the distribution of AOD values is not continuous in the map (seen in red rectangles), while in Figure 1(c) we see smoother AOD distribution and smaller gradients in the aerosol concentrations, where AOD values tend to be closer to each other. According to these visual effects, the IGG scheme of robust estimation can best improve the data quality on the AOD maps.



(a) (b) (c)

Figure1. Retrieved AOD of TERRA in central Asia (0.55μm) using: (a) the immediate averaging (existing) method; (b) the residual test method ; (c)the IGG scheme of robust estimation.

Then retrieved AOD was compared with the AOD measured at several AERONET stations in the time and area we discussed in section 3. Since there is no value for band4 in the AERONET observations, we just employ the data in band 1 (670nm) and band 3 (470nm). The validation results are showed in Figure 2. First we easily conclude that the residual test method is not appropriate according to the apparently deviating retrieved AOD value (declared in the red circle) and poor R^2 (0.582 and 0.806) in both bands. In Figure 2 (a) we can see that the averaging method performs best, with R^2 (0.848) and slope (0.946) closest to 1, verifying the accuracy of SRAP-MODIS model. Meanwhile the results of IGG method is also acceptable, whose R^2 (0.830) and slope (0.899) are very close to those of the averaging method. At the same time, in Figure 2 (b) we see an obvious improvement in the results of IGG method (slope of 0.983 and R^2 of 0.838) compared to that of the averaging method (slope of 0.882 and R^2 of 0.792) whose accuracy is not as good in band 3 as in band1. In summary, the IGG method can preserve reliable and stable accuracy in both bands, and its quality improvement of SRAP-MODIS retrieved AOD is validated tentatively.



(a) (b)

Figure 2. Comparison of AERONET AOD with retrieved AOD using 3 methods mentioned above
(a) AOD at 0.67μm; (b) AOD at 0.47μm.

5. CONCLUTION

An improvement of data quality in AOD retrieval using SRAP-MODIS is presented in this paper. For the concern of homogeneity in the 100-pixel box during retrieving AOD of 10-km resolution, we tried the IGG scheme of robust estimation to discover the outliers and make the best estimation of the TOA reflectance of the box. Compared to the results of immediate averaging and residual test, we preliminary conclude that this method performs well in the retrieval and data quality is best improved.

We must admit that the progress showed in the quantitative validation is not so evident, partially due to the lack of useful AERONET observations in areas with harsh surface homogeneity. Further specific validation is expected especially in such areas. Moreover our method hasn't been applied in considerable number of different areas and dates, neither have more robust estimators been tested, leaving us fairly much further work in the future.

6. ACKNOWLEDGMENTS

This work was supported in part by the Ministry of Science and Technology, China under Grants No. 2010CB950802, 2008AA12Z109 and 2010CB950803. The MODIS data come from NASA Goddard Space Flight Center, Level 1 and Atmosphere Archive and Distribution System (LAADS Web). The ground measurement data for validation are from 38 AERONET sites in Asia. The authors thank all the principal investigators and their staff for providing the data and products used in this investigation.

7. REFERENCES

Gao, Y. L., Xia, L. H., Wang, F., Huang, J., 2009. The progress and prospect of remote sensing for aerosol optical depth. Joint Urban Remote Sensing Event, 1-3, pp. 677-682.

Gui, Q., and Zhang, J., 1998. Robust biased estimation and its applications in geodetic adjustments. Journal of Geodesy, 72, pp. 430- 435.

Huber, P. J., 1964. Robust estimation of a location parameter. Annals of Mathematical Statistics, 35, pp. 73-101.

Kaufman, Y. J., Wald, A. E., Remer, L. A., Gao, B. C., Li, R. R., and Flynn, L., 1997. The MODIS 2.1μm channel-correlation with visible reflectance for use in remote sensing of aerosol. IEEE Transaction on Geoscience and Remote Sensing, 35, pp.1286-1298.

King, M.D., Kaufman, Y.J., Tanré, D., and Nakajima, T., 1999. Remote Sensing of Tropospheric Aerosols from Space: Past, Present, and Future. Bulletin of the American, 80(11), pp. 2229-2259.

Remer, L. A., Tanré, D., and Kaufman, Y. J., 2006. Algorithm for remote sensing of tropospheric aerosol from MODIS: collection 5. atbd_mod_02.pdf. (from http://modis.gsfc.nasa.gov/data/atbd/atmos_atbd.php.)

Tang, J., Xue, Y., Yu, T., and Guan, Y., 2005. Aerosol optical depth determination by exploiting the synergy of Terra and Aqua MODIS. Remote Sensing of Environment, 94(13), pp. 327-334.

Xue, Y., and Cracknell, A. P., 1995. Operational bi-angle approach to retrieve the Earth surface albedo from AVHRR data in the visible band. International Journal of Remote Sensing, 16(3), pp. 417-429.

Yang, Y., Cheng, M. K., Shum, C. K., Tapley, B. D., 1999. Robust estimation of systematic errors of satellite laser range. Journal of Geodesy, 73, pp. 345-349.

Yang, Y., Song, L., Xu, T., 2002. Robust estimator for correlated observations based on bifactor equivalent weights. Journal of Geodesy, 76, pp. 353-358.

Zhou, J., 1989. Classical theory of errors and robust estimation (in Chinese). Acta Geodaetica et Cartographica Sinica, 18, pp. 115-120.